ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Republic of the Philippines
North Eastern Mindanao State University
Main Campus
Tandag City, Surigao del Sur
Graduate School

Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao Del Sur

Marvin Guillarte Minguillan

A Graduate Thesis Presented to the Faculty of Graduate School, North Eastern Mindanao State University
Rosario, Tandag City, Surigao Del Sur

In Partial Fulfilment of the Requirements for the Degree of Master of Science in Computer Science

Publication Date: 2026/06/30

How to Cite: Marvin Guillarte Minguillan (2026) Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao Del Sur. *International Journal of Innovative Science and Research Technology*, 11(6), 1787-1867. https://doi.org/10.38124/ijisrt/26jun615

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPROVAL SHEET

This Thesis entitled Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement In DepEd Surigao del Sur, prepared and submitted by Marvin G. Minguillan in partial fulfillment of the requirements for the degree Master of Science in Computer Science is hereby recommended for approval and acceptance.

AIMEE F. ROMERO — Censor

NAP NICHOLE GREG S. SALERA, PhEdD-TM — Adviser

APPROVED in partial fulfillment of the requirements for the degree Master of Science in Computer Science by the Oral Examination Panel with a rating of PASSED.

ESMAEL V. MALIBERAN, DIT

Chairman

BORN CHRISTIAN ISIP, DTE

Member

MYELINDA BALDELOVAR, DIT

Member

CHERLY B. SARDOVIA, PhD-TM

Member

ACCEPTED and APPROVED in partial fulfillment of the requirements for the degree of Master of Science in Computer Science

JENNIFER M. MONTERO, EdD

Dean, Graduate School

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

ACKNOWLEDGMENT

The completion of this study would not have been possible without the generous support, guidance, and encouragement of individuals who willingly shared their time, knowledge, and expertise throughout this journey.

First and foremost, I express my heartfelt gratitude to my thesis advisor, Nap Nicole Greg S. Salera, PhEdD-TM**, for his** unwavering guidance, patience, and invaluable insights. His continuous supervision greatly influenced the development and success of this study. Working under his mentorship has truly been both an honor and a privilege, and his dedication to my academic growth has left a lasting impact on me.

I would also like to extend my sincere appreciation to the panel of examiners, led by Esmael V. Maliberan, DIT**,** together with Dr. Born Christian Isip, DTE**,** Myelinda Baldelovar, DIT**,** and Mrs. Cherly Sardovia, PhD-TM**,** for their constructive feedback, insightful suggestions, and thoughtful recommendations. Their expertise, encouragement, and understanding have been instrumental in refining and strengthening this research.

My deepest thanks go to my questionnaire validators, Mr. Melbrick Evallar, Mr. Michael Estal**,** and Dr. Bryan Guibijar, for ensuring the validity and relevance of the research instrument. I am equally grateful to the IT validators, Mr. Kristian Dior Layno**,** Mr. Eugene Saarenas**,** and Mr. Christopher Jabagat**,** who also served as a checker for the sentiment analysis component of this study. Their technical expertise and support significantly contributed to the accuracy and reliability of the findings.

I would also like to sincerely acknowledge Dr. Ezekiel Pelayo, my statistician, for his guidance and expertise in the statistical treatment and analysis of the data. His assistance was crucial in ensuring the rigor and credibility of this research.

My gratitude also goes to Lorenzo O. Macasocol, PhD, CESO V, for granting me permission to conduct this study in the Department of Education, Surigao del Sur Division. His support made the data collection process possible and smooth.

Above all, I offer my deepest gratitude to God for His guidance, strength, and wisdom, which sustained me in completing this endeavor.

M.G.M

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

DEDICATION

This thesis is wholeheartedly dedicated to the people who have been my source of strength, inspiration, and encouragement throughout this journey.

Above all, I offer this work to Almighty God, with deepest gratitude for His countless blessings, wisdom, strength, and guidance. His presence has sustained me through every challenge and has made this achievement possible.

I lovingly dedicate this work to my parents and family, whose unconditional love, prayers, sacrifices, and unwavering support have been the foundation of my perseverance. Their belief in me gave me the courage to continue, even during the most difficult times. This accomplishment is as much theirs as it is mine.

I also extend this dedication to my friends and loved ones, whose encouragement, understanding, and constant presence lifted me during long nights and moments of doubt. Their support reminded me that I was never alone in this journey.

This work stands as a humble expression of my deepest gratitude to all of you. Thank you for being part of this meaningful milestone in my life.

M.G.M

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

ABSTRACT

The manual processing of Client Satisfaction Measurement (CSM) data in government service offices has long been associated with encoding errors, fragmented consolidation, and delayed reporting, which undermine the transparency and accountability standards required under Republic Act No. 11032. This study employed a developmental research design to develop and evaluate a web-based automated CSM system integrated with a Bidirectional Encoder Representations from Transformers (BERT)-based sentiment analysis component within the Department of Education (DepEd) Division of Surigao del Sur. The study involved 61 participants through purposive sampling, composed of 50 client-respondents who directly transacted with the division offices and provided feedback through the online CSM platform, eight key informants who oversee CSM implementation and frontline service delivery, and three IT experts who conducted technical evaluation of the system's quality attributes.

System functionality was confirmed through a structured ten-test-case plan, all of which passed successfully. User evaluation using the ISO 25010 Software Quality Standard yielded Very Great Extent ratings across all six quality indicators, with Security obtaining the highest mean score of 4.77, followed by Maintainability at 4.75, Efficiency at 4.60, Usability at 4.35, Reliability at 4.33, and Functionality at 4.29. The BERT model, evaluated on 28 valid feedback items drawn, achieved an overall accuracy of 71.43%, with per-class F1-Scores of 82.76% for Positive, 60.00% for Negative, and 58.82% for Neutral sentiments. Thematic analysis of key informant responses identified five substantive dimensions affirming that the system supports ARTA-compliant reporting. These findings establish that automated CSM platforms can substantially improve accuracy, timeliness, and regulatory compliance in public sector service evaluation.

Keywords: *Client Satisfaction Measurement, Sentiment Analysis, BERT, Web-Based System, Data Aggregation, Artificial Intelligence.*

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

TABLE OF CONTENTS

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

LIST OF FIGURES

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

LIST OF TABLES

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

CHAPTER ONE
THE PROBLEM AND ITS SETTING

➢ *Introduction*

The present study focuses on the development and implementation of an automated system for analyzing, aggregating, and reporting Customer Satisfaction Measurement (CSM) data using a BERT-based sentiment analysis algorithm. In many public service institutions, particularly within government agencies, customer satisfaction surveys are widely used as a mechanism for evaluating service delivery and ensuring compliance with transparency and accountability standards. However, despite the availability of digital tools, many organizations still rely on semi-manual processes for collecting, consolidating, and interpreting survey responses. This creates inefficiencies in data handling, delays in report generation, and limitations in extracting meaningful insights from qualitative feedback. There is therefore a need to explore more advanced, automated approaches that can transform raw feedback into real-time, data-driven reports that support decision-making and service improvement.

Existing literature highlights the growing role of sentiment analysis and machine learning in improving customer feedback systems. Sentiment analysis, as part of natural language processing, enables the automatic classification of textual feedback into meaningful categories such as positive, negative, or neutral sentiments, thereby reducing the need for manual interpretation (Alqurafi & Alsanoosy, 2024). Studies have shown that traditional manual analysis of feedback is time-consuming and prone to errors, resulting in delayed reporting and reduced responsiveness to customer concerns (Omas-as & Encarnacion, 2024). More advanced approaches, particularly those using transformer-based models such as BERT, demonstrate significantly higher accuracy and efficiency in handling large volumes of textual data, with performance levels reaching up to 95% accuracy in sentiment classification (Rahman & Maryani, 2024). Furthermore, sentiment analysis has been recognized as a critical tool for improving customer satisfaction, enhancing service quality, and enabling organizations to make data-driven decisions based on real-time feedback (Li et al., 2024). These studies collectively support the potential of automated sentiment analysis systems in addressing the limitations of traditional feedback mechanisms.

Despite these advancements, gaps remain in the practical application of these technologies within local public service settings, particularly in the context of government compliance systems such as ARTA (Anti-Red Tape Authority) and Customer Satisfaction Measurement frameworks. In many agencies, including those in Surigao del Sur Division, feedback is still collected through manual or semi-digital means, such as paper-based surveys or basic online forms, which require manual consolidation and interpretation. This process often leads to delays in report generation, inconsistencies in data interpretation, and underutilization of qualitative feedback. As noted in institutional studies, delays in data evaluation and consolidation directly affect the timeliness and effectiveness of reporting systems (Omas-as & Encarnacion, 2024). In the local context, clients frequently experience dissatisfaction due to slow feedback processing, lack of real-time reporting, and limited visibility of how their responses influence service improvements. These persistent issues highlight the need for a more efficient, automated, and intelligent system that can process large volumes of feedback accurately and generate timely reports aligned with ARTA standards and digital governance initiatives.

The significance of this study lies in its contribution to improving public service delivery through the integration of advanced machine learning techniques into customer feedback systems. By utilizing a BERT-based sentiment analysis algorithm, the study aims to enhance the accuracy, efficiency, and reliability of CSM reporting while reducing the time and effort required for data processing. This aligns with the broader movement toward digital governance, where technology is used to improve transparency, efficiency, and citizen engagement in public services (Hien et al., 2024). The proposed system is expected to provide real-time insights into customer satisfaction, support evidence-based decision-making, and strengthen compliance with ARTA requirements. Ultimately, this study contributes to the frontier of knowledge by bridging the gap between theoretical advancements in sentiment analysis and their practical application in local government service delivery, thereby promoting more responsive, data-driven, and citizen-centered public administration.

➢ *Theoretical Framework*

This study is anchored on the premise that the manual collection, consolidation, and interpretation of Client Satisfaction Measurement data in DepEd Surigao del Sur can be improved through an automated and theoretically grounded system. The proposed system integrates structured satisfaction ratings with BERT-based sentiment analysis of qualitative client feedback to generate accurate, efficient, and ARTA-compliant real-time reports. To support this claim, the study adopts one main theory and three supporting theories, each directly linked to a specific variable, process, or component of the system. The Expectancy–Disconfirmation Theory serves as the central theoretical foundation, while the SERVQUAL Model, the Technology Acceptance Model, and Sentiment Analysis using Natural Language Processing and BERT function as supporting frameworks that operationalize satisfaction measurement, system utilization, and automated data analysis.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

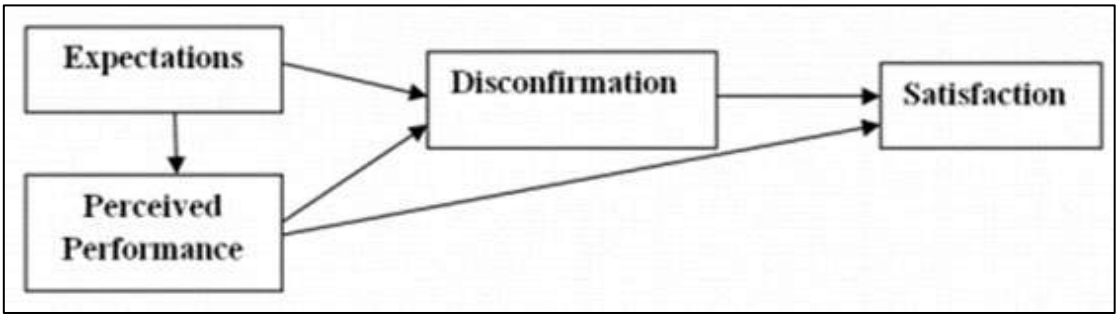


Fig 1 Expectancy-Disconfirmation Theory (Oliver, R. L. (1980)

The Expectancy–Disconfirmation Theory (Oliver, 1980), as illustrated in figure 1, provides the primary theoretical justification for understanding client satisfaction as the central construct of the study. The theory explains that satisfaction is a cognitive evaluation resulting from the comparison between expectations and perceived service performance (Oliver, 1980; Yüksel & Yüksel, 2001). When perceived performance meets or exceeds expectations, satisfaction occurs, while dissatisfaction results when performance falls below expectations. In this study, this theory directly explains the dependent variable, which is client satisfaction. The structured ratings collected through the Client Satisfaction Measurement instrument represent explicit evaluations of service performance, while textual feedback expresses the experiential outcomes of the same evaluative process. The outputs of the automated system therefore reflect aggregated disconfirmation outcomes that are theoretically interpretable. This positions the Expectancy–Disconfirmation Theory as the conceptual basis for interpreting both quantitative and qualitative satisfaction data within the system.

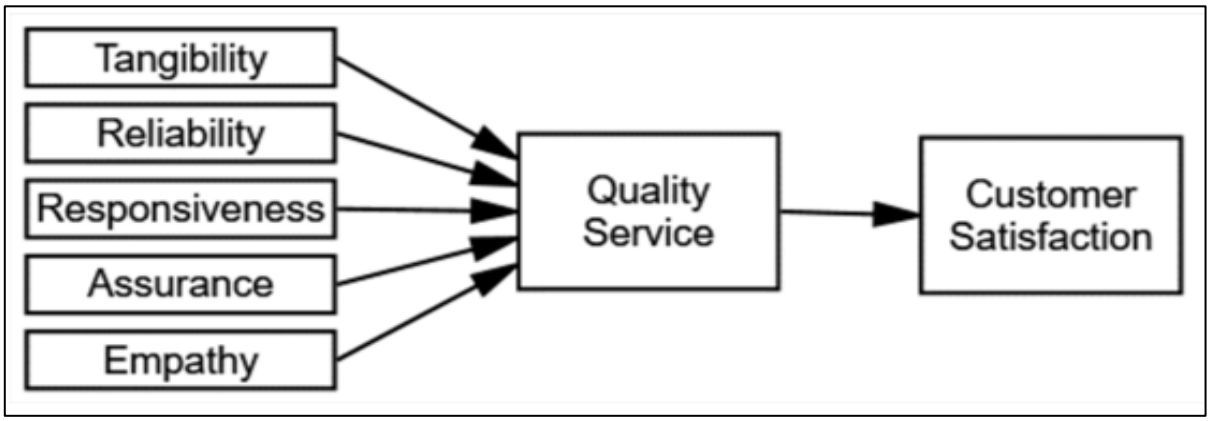


Fig 2 SERVQUAL Model of Parasuraman, A., Zeithaml, V. A., & Berry,L. (1988)

To operationalize how satisfaction is measured, the SERVQUAL Model, as shown in figure 2, provides the theoretical structure for the input component of the system. The model identifies five dimensions of service quality, namely tangibles, reliability, responsiveness, assurance, and empathy, which define how service performance is evaluated (Parasuraman et al., 1988; Arli et al., 2024). In this study, these dimensions form the basis of the structured survey instrument used in the Client Satisfaction Measurement tool. Each survey item is aligned with a specific dimension, allowing the system to collect standardized and quantifiable data. This situates SERVQUAL within the input layer of the system, where client responses are organized into measurable indicators. The use of SERVQUAL enables the system to aggregate and compare satisfaction scores across offices, time periods, and service types in a consistent manner. However, while the model effectively structures numerical ratings, it does not capture the meaning behind client experiences expressed in open-ended responses. This limitation highlights the need for a mechanism that can interpret qualitative feedback.

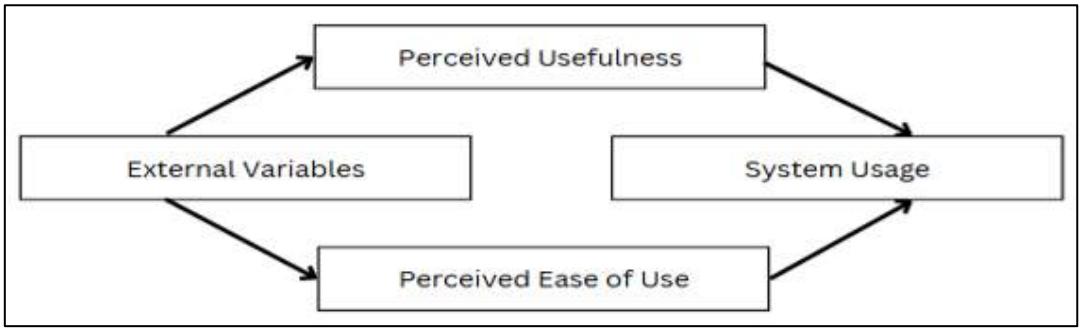


Fig 3 Technology Acceptance Model (Davis, 1989)

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

While SERVQUAL ensures that satisfaction is measured systematically, the effectiveness of the proposed system also depends on its usability and adoption within the organization. This aspect is explained by the Technology Acceptance Model as sown in figure 3, which posits that system use is influenced by perceived usefulness and perceived ease of use (Davis, 1989; Chuttur, 2009). In this study, TAM is directly linked to the system utilization component, particularly the interaction of DepEd personnel with the automated platform. The system is designed to support efficient data collection, automated processing, and real-time reporting, which enhances perceived usefulness by improving job performance. At the same time, user-friendly interfaces and simplified processes support perceived ease of use. This ensures consistent system adoption, which is necessary for reliable data generation and reporting. However, while TAM explains how the system is adopted and used, it does not address how the collected feedback is analyzed. This analytical requirement is addressed by the final supporting theory.

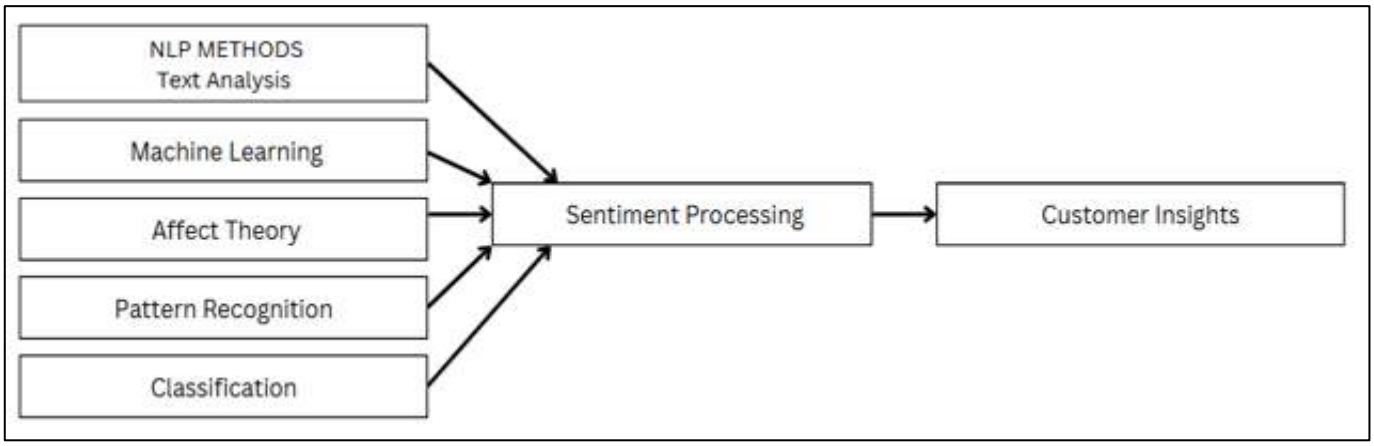


Fig 4 Sentiment Analysis by Tomkins (1982)

Sentiment Analysis using Natural Language Processing and BERT as shown in figure 4 provides the analytical mechanism for processing qualitative client feedback. Natural Language Processing enables the extraction of meaning from textual data, while BERT offers a context-aware model capable of accurately classifying sentiments into positive, neutral, or negative categories (Devlin et al., 2019; Mozumder et al., 2024; Rahman & Maryani, 2024). In this study, this framework is directly linked to the data processing component of the system, where open-ended client responses are transformed into structured sentiment outputs. This addresses the limitation of SERVQUAL by capturing the emotional and contextual aspects of client experiences that are not reflected in numerical ratings. More importantly, sentiment analysis establishes a direct connection to the Expectancy–Disconfirmation Theory. The sentiment classifications generated by BERT are interpreted as indicators of disconfirmation outcomes. Positive sentiment reflects that expectations were met or exceeded, while negative sentiment reflects unmet expectations. This ensures that qualitative outputs are not treated as separate results but are aligned with the theoretical foundation of satisfaction.

A critical feature of this framework is the integration of structured ratings and textual sentiment outputs within the system. The structured ratings derived from the SERVQUAL-based survey instrument provide a quantitative representation of client satisfaction, reflecting the extent to which service performance meets expectations. In contrast, the textual feedback analyzed through BERT-based sentiment analysis provides a qualitative interpretation of client experiences by capturing the emotional tone and contextual meaning of their responses. These two forms of data are processed simultaneously and converge in the reporting stage of the system. The numerical ratings indicate the level of satisfaction, while the sentiment classifications explain the underlying reasons behind those ratings. This integration ensures that the system produces outputs that are both statistically measurable and contextually meaningful, thereby strengthening the interpretation of satisfaction in accordance with the Expectancy–Disconfirmation Theory.

These theories form a coherent and focused framework in which each component contributes to a unified system. The Expectancy–Disconfirmation Theory explains the formation of satisfaction as the central variable. The SERVQUAL Model ensures that satisfaction is measured through structured and validated dimensions within the input component. The Technology Acceptance Model ensures that the system is usable and adopted within the organization, supporting consistent operation. Sentiment Analysis using Natural Language Processing and BERT enables the automated interpretation of qualitative feedback within the processing component. Each theory addresses a specific limitation of the others, resulting in a system that is theoretically grounded, operationally efficient, and analytically comprehensive. This integrated framework supports the development of an automated Client Satisfaction Measurement system capable of producing accurate and real-time reports, thereby addressing the limitations of the existing manual process and fulfilling the objectives of the study.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Conceptual Framework*

The study adopts a system-oriented conceptual framework that illustrates the transformation of raw client feedback into actionable insights through an automated data aggregation and reporting system. The framework integrates structured survey responses and unstructured textual feedback within a unified processing environment supported by artificial intelligence and database technologies. This framework is designed to directly address the identified problems related to inefficiency, accuracy, automation, and compliance in client satisfaction reporting.

The system begins with the input component**,** which consists of Client Satisfaction Measurement (CSM) data collected through an online survey platform embedded within the Human Resource Information System (HRIS). This includes both quantitative responses using a Likert scale and qualitative open-ended feedback. The integration of the CSM module within the HRIS facilitates a more efficient and centralized approach to data collection, reducing reliance on manual processes and improving accessibility and consistency of information across transactions.

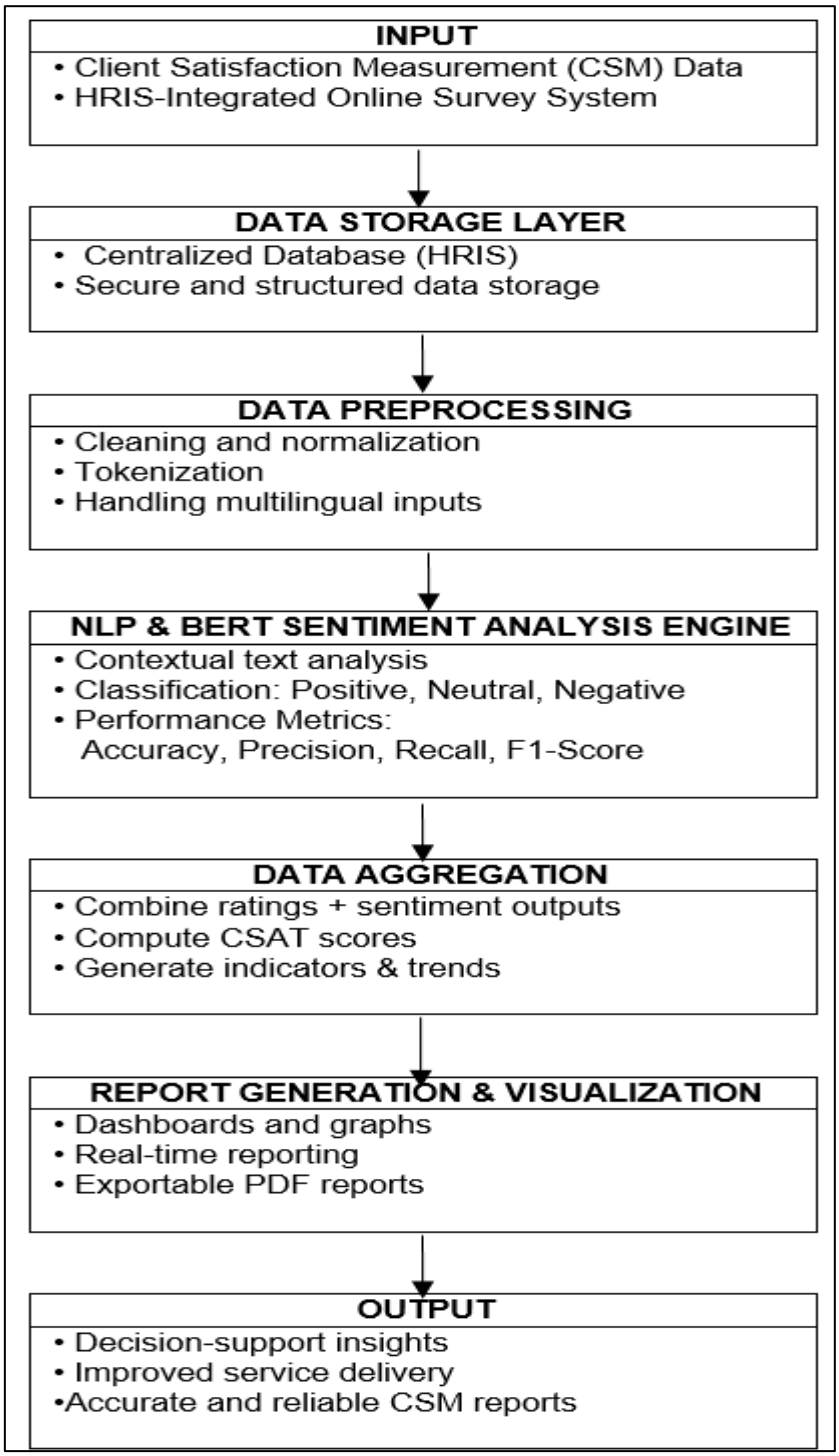


Fig 5 Conceptual Framework of the Study

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Following data collection, the system proceeds to the data storage layer**,** where all responses are securely stored in a relational database. This centralized storage ensures data integrity, traceability, and efficient retrieval of information. By organizing data within a structured database environment, the system minimizes redundancy and delays associated with manual consolidation, thereby supporting faster and more reliable data processing.

The next stage involves data preprocessing**,** which prepares raw data for analysis through cleaning, normalization, tokenization, and handling of multilingual inputs. This stage plays a critical role in improving data quality by removing inconsistencies and irrelevant elements that may affect analysis. Through this process, the system enhances the reliability and consistency of the dataset, ensuring that subsequent analytical procedures produce accurate and dependable results.

After preprocessing, the system performs sentiment analysis using Natural Language Processing (NLP) and Bidirectional Encoder Representations from Transformers (BERT)**.** The BERT model enables deep contextual understanding of textual feedback by analyzing relationships between words within a sentence. This allows the system to classify feedback into positive, neutral, or negative sentiments with a higher level of accuracy compared to traditional methods. The use of an advanced machine learning model supports automated analysis and real-time interpretation of qualitative data, significantly improving the efficiency of report generation.

To further ensure the robustness of the analytical process, the system incorporates performance evaluation metrics**,** including Accuracy, Precision, Recall, and F1-Score. These metrics provide quantitative measures of the model's effectiveness in classifying feedback, allowing continuous validation and improvement of the system's performance. The inclusion of these evaluation mechanisms strengthens the reliability of the generated insights.

The processed data then moves to the data aggregation stage**,** where structured survey ratings and sentiment outputs are combined. This integration enables the computation of overall satisfaction scores and the identification of patterns such as satisfaction trends and sentiment distribution. By merging numerical and textual data, the system produces a more comprehensive assessment of service quality, supporting deeper analysis and interpretation.

Finally, the system generates outputs through the report generation and visualization layer**,** which includes real-time dashboards, graphical representations, and downloadable reports. These outputs provide timely and data-driven insights that support decision-making and service improvement. Moreover, the automation of reporting processes enhances transparency, responsiveness, and consistency in client satisfaction monitoring, aligning with established standards for public service delivery.

Overall, the conceptual framework presents a continuous and systematic process where client feedback is efficiently collected, validated, analyzed, and transformed into actionable information. Each component of the system contributes to improving operational efficiency, analytical capability, data accuracy, and compliance, thereby supporting the overall objective of enhancing client satisfaction measurement and reporting.

➢ *Statement of the Problem*

The main objective of the study is to develop Web-based real-time automation of data aggregation and reporting using Bidirectional Context Understanding. Specifically, the study seeks to answer the following:

- What technological solutions can be used to streamline the collection, storage, and processing of survey data to reduce the time and effort required for data consolidation and analysis?
- What algorithm can be used to effectively and efficiently generate real-time reports and analyze the data automatically?
- How can the proposed technological solutions enhance the accuracy and reliability of CSM reports through automated data entry and validation?
- How can customer feedback be classified using the BERT (Bidirectional Encoder Representations from Transformers) model, and how can the performance of the classification task be measured using Accuracy, Precision, Recall, and F1-Score?
- How can the system ensure compliance with ARTA standards for client satisfaction reporting?

➢ *Scope and Limitation of the Study*

This study was conducted to develop and evaluate a web-based system that automates the aggregation, analysis, and reporting of Client Satisfaction Measurement (CSM) data within the Department of Education (DepEd) Division of Surigao del Sur. The general purpose of the study is to address inefficiencies in manual data consolidation and to improve the accuracy, timeliness, and reliability of client satisfaction reporting through the integration of database technologies and artificial intelligence-driven analytics.

The subject matter of the study focuses on the automation of client feedback processing, specifically involving structured survey data and unstructured textual responses. It covers key topics such as data collection through an online platform, centralized data storage, preprocessing of multilingual responses, sentiment analysis using a transformer-based model, automated data aggregation, and real-time reporting through dashboards. The study also examines how these technological components contribute to streamlining public service monitoring and enhancing decision-making processes in a government setting.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

The study was conducted within the DepEd Division of Surigao del Sur, which serves as the primary locale and source of data. The system is implemented in selected frontline service offices within the division, including the Human Resource Management Office, Records or Receiving Section, Administrative Office, Budget and Accounting Office, School Governance and Operations Division (SGOD), Curriculum Implementation Department (CID), and other service units that regularly interact with clients and collect satisfaction feedback. The data used in the study belong to the division and are generated through client transactions facilitated by these offices.

The population of the study consists of three groups of respondents. The first group includes internal users such as administrative personnel, Human Resource staff, and designated CSM focal persons who are responsible for managing and analyzing client feedback. The second group includes external respondents, specifically clients or stakeholders who access services from the identified offices and provide feedback through the survey system. The third group is the client-respondents who have actual transactions in the division office. These respondents form the basis of the dataset used for system testing and evaluation.

The data for this study were gathered over a specific period covering the implementation and testing phases of the system, spanning several months within the academic or fiscal year in which the study was conducted. This period includes the collection of survey responses, system processing, and evaluation of outputs to ensure that the generated reports reflect actual operational conditions.

The system developed in this study operates through a structured process. Data are collected through an online survey integrated into the Human Resource Information System (HRIS), capturing both quantitative ratings and qualitative feedback. These data are stored in a centralized database and undergo preprocessing to ensure consistency and usability. A sentiment analysis module, powered by a BERT-based model, classifies textual responses into positive, neutral, or negative sentiments, consistent with studies demonstrating the effectiveness of transformer-based approaches in text classification . The processed data are then aggregated and analyzed to compute satisfaction scores and identify trends, which are subsequently presented through real-time dashboards and reports.

Despite its contributions, the study has several limitations. The implementation is confined to the DepEd Division of Surigao del Sur and selected offices, which may limit the generalizability of the findings to other divisions or institutions. The evaluation is also based on a specific group of users and respondents, which may not fully capture all possible variations in system usage.

A significant limitation lies in the sentiment analysis component of the system. While a BERT-based model is utilized, its performance may be affected by the size and quality of the training dataset, which may not be sufficiently extensive to capture all linguistic variations. Additionally, language variation such as code-switching between English, Filipino, and local dialects may reduce classification accuracy. The model may also have difficulty interpreting context-specific expressions, including idiomatic language, sarcasm, and culturally nuanced statements, which are inherently complex for automated systems to process. These constraints highlight the challenges of applying natural language processing techniques in multilingual and context-sensitive environments.

Furthermore, the study does not include integration with external government systems beyond the HRIS, nor does it incorporate advanced functionalities such as predictive analytics or automated recommendation systems. The overall effectiveness of the system is also dependent on the accuracy, completeness, and consistency of the data provided by users. Additionally, the developed system does not incorporate a looping mechanism, meaning it cannot automatically retrain or update its classification model based on newly collected feedback data. The sentiment analysis component operates on a fixed trained model, and any refinement of classification performance would require manual retraining with an updated dataset outside the system environment. Nevertheless, the study provides a practical and context-specific solution for improving client satisfaction data management within the identified setting.

➢ *Significance of the Study*

This part of the study describes the contributions made to the following entities:

- *The DepEd Surigao del Sur Division.*

The findings of this study may provide empirical basis for enhancing decision-making processes within the division. By examining the application of an automated system for client satisfaction measurement, the study may inform the development of data-driven approaches to resource allocation, policy formulation, and service improvement. Furthermore, the proposed system may contribute to improved operational efficiency by reducing manual processes associated with data consolidation and analysis. However, such improvements are contingent upon appropriate system implementation, user acceptance, and institutional support.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

- *Clients of DepEd Surigao del Sur.*

The study may have implications for improving the responsiveness of public services. Through the integration of automated data processing and feedback analysis, it is possible that client concerns can be identified and addressed more systematically. This, in turn, may contribute to more timely and evidence-based service adjustments. Nevertheless, the extent of these benefits may depend on the consistency of data collection and the effective use of generated insights in decision-making.

- *Computer Science and Information Systems.*

From an academic perspective, this study contributes to the field of Computer Science by demonstrating the application of machine learning and natural language processing techniques, particularly transformer-based models such as BERT, in the context of public-sector data analytics. Prior studies have shown that BERT-based models can significantly improve the accuracy of sentiment classification and enable more precise interpretation of textual feedback . In this regard, the study extends existing work by situating these techniques within a government service environment, thereby contributing to the growing body of knowledge on intelligent information systems and automated decision-support tools. It may also provide a reference framework for integrating structured and unstructured data processing within web-based systems.

- *Public-Sector Information Systems and Governance.*

The study may also offer insights into the role of artificial intelligence in enhancing public administration processes. The integration of automated analytics in client satisfaction measurement aligns with emerging perspectives on AI-enabled governance, which emphasize improved efficiency, transparency, and data-driven policymaking . As such, the findings may inform future initiatives aimed at digital transformation in public institutions, particularly in aligning service delivery mechanisms with regulatory standards such as the Anti-Red Tape Act. However, the applicability of the system may vary depending on institutional capacity, infrastructure readiness, and data governance practices.

- *IT Practitioners and System Developers.*

For practitioners, the study may serve as a technical reference for designing and implementing automated survey analysis systems. It provides insights into system architecture, data preprocessing, and the integration of sentiment analysis models within real-time reporting platforms. The study may also highlight considerations related to scalability, data quality, and model performance in real-world applications.

- *Future Researchers.*

The study contributes to the academic discourse on automated client satisfaction measurement by providing a conceptual and methodological foundation for future investigations. It may serve as a basis for further research on improving model accuracy, expanding multilingual sentiment analysis, or exploring other machine learning techniques for public-sector applications. Additionally, future studies may build upon this work by examining system adoption, user experience, and long-term impact on organizational performance.

➢ *Definition of Terms*

To ensure that the terminology in the analysis was understood clearly, the researcher conceptually and operationally described the following:

- *Accuracy.*

A performance metric that determines how correctly a classification model predicts outcomes. In this study, it is measured as the percentage of CSM textual feedback correctly classified by the BERT model into positive, negative, or neutral categories compared with the manually labeled test dataset.

- *Algorithm.*

A defined sequence of computational steps used to solve a problem or perform a task. In this study, it refers to the set of programmed procedures embedded in the system to process survey inputs, compute satisfaction scores, perform sentiment classification, and generate automated reports.

- *API Integration.*

A mechanism that allows different software systems to communicate and exchange data. In this study, it is used to connect the CSM survey platform with the database and analytics module, enabling automatic transmission of client responses into the system without manual intervention.

- *ARTA Compliance.*

The adherence to the standards set by the Anti-Red Tape Authority in delivering public services. In this study, it is reflected in the system's ability to provide timely, transparent, and efficient client satisfaction reports that support DepEd's service monitoring and decision-making processes.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

- *Automated Data Entry.*
  A system feature that captures and records data without manual encoding. In this study, it refers to the direct storage of client responses submitted through the online CSM platform into the database in real time.

- *Automated Reporting.*
  The generation of reports by a system without human intervention. In this study, it pertains to the automatic creation of graphical and numerical summaries of client satisfaction and sentiment results displayed in the dashboard.

- *BERT (Bidirectional Encoder Representations from Transformers).*
  A deep learning-based language model designed to understand contextual relationships in text. In this study, it is used to analyze open-ended CSM responses and classify them into sentiment categories based on contextual meaning.

- *Client Satisfaction Measurement (CSM).*
  A feedback mechanism used to gather client perceptions of service quality. In this study, it serves as the primary data collection tool embedded in the web-based system, capturing both quantitative ratings and qualitative comments from clients of the DepEd Division Office.

- *Classification Task.*
  The process of assigning data into predefined categories. In this study, it refers to the categorization of textual client feedback into positive, negative, or neutral sentiments using the trained BERT model.

- *CSAT (Customer Satisfaction) Algorithm.*
  A computational method used to quantify satisfaction levels. In this study, it refers to the formula applied by the system to compute overall satisfaction scores from Likert-scale responses, which are then displayed in the reporting dashboard.

- *Customer Feedback.*
  Information provided by clients regarding their service experience. In this study, it consists of numerical ratings and textual responses collected through the CSM platform and used as input data for aggregation and sentiment analysis.

- *Customer Satisfaction.*
  The level of fulfillment perceived by clients based on service experience. In this study, it is quantified using computed mean scores from CSM ratings and interpreted alongside sentiment classification results.

- *Data Aggregation.*
  The process of collecting and combining data from multiple sources into a unified dataset. In this study, it refers to the consolidation of all CSM responses into a centralized database for analysis and reporting.

- *Data Preprocessing.*
  The preparation of raw data before analysis. In this study, it involves cleaning, normalizing, and transforming textual feedback, including tokenization and removal of irrelevant characters, before inputting the data into the BERT model.

- *Digital Governance.*
  The use of digital technologies to enhance public service delivery and management. In this study, it is demonstrated through the system's role in supporting DepEd's data-driven decision-making using automated client satisfaction analytics.

- *Efficiency.*
  The ability to accomplish tasks with minimal time and resources. In this study, it is evaluated based on the reduction in time required for data collection, processing, and report generation compared to manual methods.

- *Evaluation Metrics.*
  Standards used to assess the performance of a model. In this study, these include accuracy, precision, recall, and F1-score used to evaluate the effectiveness of the BERT sentiment classification.

- *F1-Score.*
  A metric that combines precision and recall to provide a balanced measure of model performance. In this study, it is computed to assess how well the BERT model balances correct sentiment predictions and missed classifications.

- *Machine Learning Classification.*
  An automated method of categorizing data based on learned patterns. In this study, it refers to the use of the BERT model to classify textual feedback into sentiment categories.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

- *Natural Language Processing (NLP).*
  A field of artificial intelligence that focuses on analyzing human language. In this study, it is applied through BERT to interpret and process textual responses from clients.

- *Precision.*
  A metric that measures the correctness of positive predictions made by a model. In this study, it represents the proportion of correctly predicted sentiment labels among all predicted instances.

- *Real-Time Analytics Engine.*
  A system component that processes data immediately after input. In this study, it refers to the module that analyzes incoming CSM responses and updates computed results in the database instantly.

- *Real-Time Reporting.*
  The ability to generate updated outputs immediately after processing data. In this study, it refers to the instant display of satisfaction scores and sentiment results in the system dashboard.

- *Recall.*
  A metric that measures the ability of a model to identify all relevant instances. In this study, it represents the proportion of actual sentiment instances correctly identified by the BERT model.

- *Reliability.*
  The consistency of a system in producing stable results. In this study, it is measured through user evaluation based on ISO 25010, focusing on consistent system performance during repeated use.

- *Sentiment Analysis.*
  A computational technique used to determine the emotional tone of text. In this study, it refers to the classification of client feedback into positive, negative, or neutral sentiments using the BERT model.

- *System Functionality.*
  The degree to which a system performs its intended tasks. In this study, it is evaluated based on how effectively the system collects, processes, analyzes, and reports client satisfaction data.

- *System Security.*
  The protection of system data from unauthorized access. In this study, it is assessed based on the system's ability to safeguard client information during data collection, storage, and processing.

- *System Usability.*
  The ease with which users can interact with a system. In this study, it is measured usingISO 25010 indicators based on respondents' experience in navigating the system and interpreting outputs.

- *Technological Solutions.*
  The application of technology to solve operational problems. In this study, it refers to the integration of a web-based survey platform, database system, and BERT model to automate client satisfaction analysis.

- *Tokenization.*
  A process in NLP that breaks text into smaller units. In this study, it refers to converting textual feedback into tokens required as input for the BERT model during preprocessing.

- *Validation Mechanism.*
  A process that ensures data accuracy and completeness. In this study, it refers to system controls that check required fields and proper input formats before accepting survey responses.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

CHAPTER TWO
REVIEW OF RELATED LITERATURE AND STUDIES

This section presents a concrete structure of relevant literature and studies relevant to the researcher's research area. It also includes prior and existing studies related to the concepts under study. It also enumerates, describes, summarizes, objectively evaluates, and clarifies previous research.

➢ *Related Literature*

- *Foreign*

Across international settings, digital transformation has increasingly shaped the way organizations gather, manage, and interpret service-related data. In public administration and service management, digital government initiatives have been associated with improvements in transparency, responsiveness, and citizen engagement. Hien et al. (2024) explained that digital governance structures create public value by strengthening service quality, accountability, and participation, while Indar et al. (2024) observed that online public service systems improve ease of access, speed of response, and administrative consistency. Taken together, these discussions suggest that digital infrastructures are not merely technological add-ons but strategic mechanisms for improving institutional performance.

This movement toward digitalization also affects how feedback data are collected and organized. Instead of relying on fragmented or paper-based procedures, many institutions have shifted toward integrated and automated collection systems that reduce procedural delays and improve the flow of information. Alqurafi and Alsanoosy (2024) noted that automated customer feedback tools make it easier for organizations to collect and classify large volumes of reviews without relying on time-consuming manual reading, while Indar et al. (2024) emphasized that online systems become more effective when they are designed to support simplicity, usefulness, and speed. In this sense, the literature points to automation as a practical response to the limitations of traditional feedback mechanisms.

As automated systems evolve, attention is increasingly given to how they process incoming data and generate outputs in real time. Real-time data processing is important because service institutions do not only need to gather information; they must also interpret it quickly enough to support immediate action and timely reporting. Hien et al. (2024) stressed that digital governance becomes more meaningful when it enables data-driven policymaking and faster institutional response, whereas Alqurafi and Alsanoosy (2024) underscored the role of automated analysis in converting raw textual feedback into structured information that can support service improvement. These viewpoints reinforce the idea that speed and interpretability are central to contemporary feedback systems.

The issue of data accuracy and consistency further strengthens the relevance of automated platforms. Automated systems are expected not only to accelerate reporting but also to improve the dependability of stored information through validation and preprocessing. Kokab et al. (2022) explained that sentiment analysis systems require careful preprocessing, feature extraction, and classification in order to avoid information loss and misclassification, while Göker (2025) emphasized that unstructured customer reviews demand systematic transformation into analyzable formats before reliable classification can occur. These discussions imply that the quality of system outputs is highly dependent on how well the input data are cleaned, organized, and validated.

The growing reliance on artificial intelligence has also expanded the analytical reach of customer satisfaction systems, especially in handling unstructured feedback. Transformer-based models have become important in this regard because they capture contextual relationships that older methods often miss. Li et al. (2024) demonstrated that BERT and its variants are capable of producing highly competitive results in sentiment analysis tasks because they interpret words in relation to surrounding context, while Göker (2025) similarly found that BERT-based models outperform several deep learning alternatives in classifying customer reviews. Viewed together, these works show that contextual language models are particularly suitable for environments where textual feedback needs to be translated into meaningful decision support.

In connection with this, performance evaluation has become a necessary part of automated sentiment classification. It is not enough for a model to produce labels; its output must also be assessed using reliable and accepted standards. Li et al. (2024) explained that classification performance must be examined through metrics such as accuracy and class-based measures to determine whether the model is balanced across categories, whereas Kokab et al. (2022) highlighted the importance of precision, recall, F1-score, and related indicators in determining how well a sentiment classifier handles diverse datasets. These measures are especially relevant in studies involving public-facing systems because classification outcomes may influence reporting quality and administrative interpretation.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Beyond analysis itself, regulatory alignment remains an important concern in automated service systems. Digital tools are more valuable when they support institutional accountability and make reporting processes more transparent and standardized. Hien et al. (2024) argued that digital public systems strengthen public trust when they improve transparency and accountability, while Indar et al. (2024) noted that online service systems become more useful when they are consistent, understandable, and responsive to stakeholder needs. Thus, the foreign literature collectively suggests that automated client satisfaction systems become most effective when they combine efficient data processing, reliable analytics, and alignment with governance standards.

- *Local*

In the Philippine setting, the use of digital systems in public administration has become increasingly associated with efforts to streamline bureaucratic procedures and improve service delivery. Local literature points out that automation can reduce delays, eliminate redundant manual work, and support more organized internal operations. Alberto (2025) discussed how process streamlining in DepEd division offices depends on digital platforms, policy alignment, and institutional standardization, while Muñoz (2024) explained that artificial intelligence and digital governance can improve public administration by supporting more informed and responsive decision-making. These discussions establish a strong local context for understanding why automated reporting systems are becoming necessary in government offices.

The movement toward automation is closely linked to the need for better storage and retrieval of institutional data. In public service environments, scattered or inconsistent datasets can easily weaken reporting quality and delay administrative response. Hernandez (2025) emphasized that service access and client experience are better understood when institutions rely on organized, data-driven reporting frameworks, while Alberto (2025) noted that centralized digital systems help maintain consistency across offices and improve overall administrative coordination. This shows that the literature treats data storage not simply as a technical requirement but as an organizational condition for reliable reporting.

Local scholars have also underscored the importance of accuracy and validation in automated environments. Since client satisfaction reports often inform administrative action, the integrity of data entry and processing becomes especially important. Alberto (2025) observed that one of the continuing challenges in public-sector administration is inconsistency in implementation across offices, which can weaken the reliability of records and reporting outputs. In a related way, Muñoz (2024) argued that AI-supported systems must be accompanied by ethical and organizational safeguards to ensure that generated information remains trustworthy and usable. These viewpoints indicate that accuracy is not achieved through digitization alone but through the careful design of validation and monitoring mechanisms.

Another prominent concern in the local literature is the growing use of artificial intelligence and natural language processing in extracting insights from user-generated data. In multilingual environments such as the Philippines, automated interpretation of textual feedback requires models capable of handling code-mixed and low-resource language contexts. Maceda et al. (2023) found that a fine-tuned multilingual BERT model effectively analyzed Philippine code-mixed social media data, while Tomas et al. (2024) demonstrated that BERT-based approaches remain useful even in low-resource Tagalog classification tasks. These contributions are especially relevant to systems that analyze open-ended feedback, since they support the feasibility of using BERT in a local service-reporting context.

The literature also recognizes that sentiment-oriented systems must be evaluated systematically to ensure that their classifications are meaningful and defensible. Laureta et al. (2025) emphasized that transformer-based NLP models can effectively detect both sentiment and profanity in Filipino social media comments, while Goma et al. (2024) demonstrated that BERT-based sentiment models can achieve strong performance in classifying social media content related to public-health topics. These findings do not merely support the use of automated classification; they also imply that metrics such as accuracy, precision, recall, and F1-score remain necessary in determining whether classification outputs are suitable for organizational use.

At the level of service governance, local literature further shows that automation supports transparency, accountability, and institutional responsiveness. Hernandez (2025) argued that data-driven dashboards improve the monitoring of service access and satisfaction trends, while Muñoz (2024) maintained that intelligent systems in governance can strengthen policy implementation when they are paired with clear structures for ethical and responsible use. Such discussions reinforce the relevance of automated client satisfaction systems in Philippine public institutions, especially where compliance-oriented reporting and citizen-centered service delivery are expected.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Related Studies*

- *Foreign*

Recent international studies provide concrete empirical support for the use of automated and intelligent systems in customer satisfaction analysis. A growing number of investigations show that organizations increasingly rely on machine learning models to estimate satisfaction levels, classify sentiments, and generate insights from large-scale feedback data. Mozumder et al. (2024) found that advanced machine learning techniques significantly improved the analysis of customer satisfaction drivers in the fintech industry, while Manderscheid and Lee (2024) demonstrated that predictive approaches can replicate survey response distributions in a way that strengthens customer satisfaction estimation. These results indicate that customer feedback analysis is moving beyond simple tabulation toward predictive and classification-oriented frameworks.

The usefulness of automated systems becomes even clearer when real-time or large-volume reporting is required. Manderscheid and Lee (2024) showed that automated satisfaction prediction can reduce bias arising from low survey response rates and can produce more representative performance information, whereas Mozumder et al. (2024) noted that customer feedback modeling becomes more meaningful when advanced systems are used to identify the distinct factors affecting satisfaction. These studies suggest that automated reporting is not only faster but also more analytically useful when compared with traditional summary-based methods.

A similar pattern appears in studies focusing on sentiment analysis of textual data. Rahman and Maryani (2024) reported that BERT outperformed several traditional and deep learning models in analyzing customer feedback and yielded highly accurate sentiment classifications. These findings strengthen the case for using transformer-based and machine learning models in contexts where open-ended comments are an important source of service evaluation.

The relevance of BERT becomes more evident in studies that directly compare it with other alternatives. Abello et al. (2023) found that BERT-based frameworks improved tweet sentiment analysis by identifying contextual meanings more accurately, while Li et al. (2024) showed that BERT and its variants remain strong performers in customer-review classification tasks involving multiple sentiment categories. These studies confirm that BERT is not simply another classification tool; it is an architecture particularly suited for contextual interpretation in sentiment analysis.

The issue of performance measurement is also central in the foreign studies. Rahman and Maryani (2024) evaluated sentiment models through accuracy and other standard metrics to determine reliability, whereas Abello et al. (2023) highlighted the importance of examining precision, recall, and F1-score in assessing whether the model captures sentiment consistently across categories. These empirical findings mirror the standard evaluative practices expected in machine learning research and reinforce the need to examine classification performance in a comprehensive way.

In addition to model performance, customer satisfaction studies have also emphasized the broader service implications of automation. Arli et al. (2024) found that the dimensions of service quality significantly influence satisfaction, loyalty, and reuse intention in online delivery systems, while Mahdjoub et al. (2026) demonstrated that hybrid NLP architectures improve the prediction of customer satisfaction by combining traditional and deep learning techniques. The convergence of these findings suggests that automated analysis systems do not merely classify feedback; they also strengthen the strategic value of customer satisfaction data for service enhancement.

Other foreign studies further show that machine learning contributes to better organizational transparency and service monitoring. Mozumder et al. (2024) observed that advanced classification frameworks offer actionable insight into customer experiences, while Mahdjoub et al. (2026) emphasized that layered NLP techniques make feedback processing more adaptable and more useful for customer relationship management. These results are particularly important for service systems that aim to generate reports not only for internal review but also for institutional accountability.

- *Local*

Within the Philippine context, empirical studies also show that automated and digital systems strengthen service monitoring and feedback management in educational and public institutions. Aliazas et al. (2022) found that design-based feedback processes improved teaching effectiveness by making student responses more systematically available to decision-makers, while Estoria-Maquiling (2023) demonstrated that service quality implementation is closely tied to satisfaction outcomes in DepEd frontline services. These studies reveal that organized feedback systems play a significant role in improving institutional responsiveness.

The same trend is evident in studies examining digital service use and customer satisfaction in educational environments. Cervantes et al. (2024) found that digital payment utilization among DepEd employees was significantly related to customer satisfaction, while Doria (2025) demonstrated that a data-driven survey model integrating sentiment and prescriptive analytics provided richer and more actionable insights than traditional satisfaction tools. These findings show that digital platforms become more valuable when they are designed not only to collect data but also to convert it into usable institutional intelligence.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Local studies also underscore the importance of unified and automated feedback systems in reducing delays and improving report management. Omas-as and Encarnacion (2024) proposed a unified feedback management system with text analytics and sentiment analysis, showing that automation reduces the burden of manual evaluation and consolidation. In another context, Legal (2023) found that mobile-based selected school services improved responsiveness to stakeholders by making services more accessible and efficient. Together, these findings suggest that system automation can directly support better reporting and faster feedback cycles.

The application of text analytics and machine learning is also becoming more visible in Philippine studies. Taruc and Dela Cruz (2024) used natural language processing to quantify the effectiveness of student organization activities through emotional response analysis, while Sy et al. (2024) applied BERTopic and transformer-based modeling to analyze stakeholder perceptions toward the Universal Access to Quality Tertiary Education program. These studies illustrate that automated text processing is increasingly being used not only for general sentiment detection but also for extracting patterns and thematic insights from stakeholder feedback.

Several local studies further highlight the relevance of multilingual and context-sensitive language models. Although centered on different datasets, Taruc and Dela Cruz (2024) and Sy et al. (2024) both demonstrate that sentiment and topic-oriented systems can be adapted to Philippine settings where qualitative feedback plays an important evaluative role. This is important because public-sector and educational feedback often contains local expressions, mixed language use, and context-specific meanings that require more advanced analytical treatment.

The role of reporting and compliance is likewise evident in local empirical work. Quiatchon (2024) found that implementation of Republic Act 11032 was significantly associated with customer satisfaction, indicating that efficient and transparent processes affect how clients evaluate public service. In a related educational context, Estoria-Maquiling (2023) observed that service quality standards influence satisfaction levels when implementation is monitored and assessed systematically. These findings suggest that automated satisfaction systems may also help institutions meet governance and compliance expectations.

Studies on service usability and system accessibility also provide useful insight. Legal (2023) demonstrated that service access improved when digital tools were designed around user needs, while Omas-as and Encarnacion (2024) argued that stakeholder satisfaction is strengthened when feedback systems are responsive, automated, and capable of generating meaningful reports. This implies that system effectiveness depends not only on technical sophistication but also on how well the tool supports the user's experience.

- *Synthesis*

The reviewed literature and studies collectively highlight the growing role of digital transformation, automation, and artificial intelligence in improving data management and customer satisfaction analysis. Across both foreign and local sources, there is a consistent emphasis on the shift from manual and fragmented processes toward integrated and automated systems that support efficient data collection, storage, and reporting. The literature underscores the importance of real-time data processing, structured data management, and validation mechanisms in ensuring the accuracy and reliability of information. In addition, the studies demonstrate that machine learning and sentiment analysis models, particularly those based on contextual language understanding, are effective in analyzing unstructured feedback and generating meaningful insights. Furthermore, both bodies of work emphasize the significance of performance evaluation metrics in validating classification outputs, as well as the need for transparency, accountability, and compliance in public service systems.

From the reviewed readings, it becomes evident that the effectiveness of automated systems is not solely dependent on the presence of technology but on how well these technologies are integrated into organizational processes. The researcher understands that data quality, system design, and user interaction all play critical roles in ensuring that automated systems produce reliable and actionable outputs. The literature suggests that the transition to automation must be accompanied by proper validation, standardization, and evaluation mechanisms to maintain the integrity of data and reports. Moreover, the insights drawn from both foreign and local studies indicate that artificial intelligence, particularly in the form of advanced natural language processing models, provides a powerful tool for interpreting complex textual data. However, its effectiveness relies heavily on proper training, evaluation, and alignment with the specific context in which it is applied.

In relation to the present study, the reviewed literature provides a strong foundation for developing a web-based system that integrates automated data aggregation, real-time reporting, and BERT-based sentiment analysis. It highlights the importance of combining technological solutions, machine learning, and evaluation metrics to address challenges in client satisfaction measurement. Moreover, it emphasizes the need for data accuracy, system reliability, and compliance with regulatory standards, which guide the development of a more efficient and responsive reporting system.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

CHAPTER THREE
RESEARCH METHODOLOGY

This section of the study presents the research methods, research design, research locale, research instrument/s, data gathering procedure, algorithm design, materials used in the study, as well the statistical tools used in the treatment and analysis of data.

- *Research Design*

This study employed a developmental research design**,** which is appropriate for investigations that aim to design, develop, and systematically evaluate a technological system. Developmental research focuses on the creation of a product, process, or system and the rigorous assessment of its effectiveness, usability, and applicability in real-world contexts. According to Richey and Klein (2007)**,** developmental research is concerned with the systematic study of design, development, and evaluation processes with the goal of establishing an empirical basis for the creation of instructional and technological products. Similarly, Creswell and Creswell (2018) emphasize that the selection of a research design should be aligned with the purpose of the study, particularly when it involves both the development of an innovation and the evaluation of its outcomes.

The use of a developmental research design in this study is appropriate because the primary objective is to develop an automated Client Satisfaction Measurement (CSM) analysis system that integrates data aggregation, real-time reporting, and sentiment analysis using artificial intelligence. The design allows the researcher to systematically construct the system while simultaneously evaluating its performance and acceptability. This approach is consistent with the increasing adoption of intelligent systems in public administration, where automation enhances efficiency, transparency, and evidence-based decision-making in service delivery .

To ensure methodological rigor and to address the panel's concern regarding clarity, the developmental research design of this study is structured into four distinct but interrelated components: system development, quantitative user evaluation, qualitative interview analysis, and BERT model validation**.** Each component is treated as a separate phase to avoid the blending of technical development and evaluation processes.

The first component, system development**,** focuses on the design and implementation of the automated system. This phase includes requirements analysis, system architecture design, database development, API integration, and the creation of dashboards for real-time reporting. The system is designed to capture both structured and unstructured client feedback from the CSM platform and process these data through automated workflows. This phase ensures that the system is functional, scalable, and aligned with operational needs in public service environments.

The second component, quantitative user evaluation**,** examines the acceptability and performance of the developed system using measurable indicators. Respondents, composed of intended users such as DepEd personnel and stakeholders, evaluate the system based on established software quality dimensions, including functionality, reliability, usability, efficiency, maintainability, and security. This approach aligns with the principle that system effectiveness must be evaluated through user perception and measurable performance indicators, particularly in public sector services where client satisfaction is a critical outcome variable..

The third component, qualitative interview analysis**,** is conducted to explore the experiences, perceptions, and challenges encountered by users in interacting with the system. Semi-structured interviews are utilized to gather in-depth data, which are then analyzed using thematic analysis. According to Creswell and Creswell (2018)**,** qualitative approaches are essential for understanding participants' experiences and providing contextual insights that cannot be captured through numerical data alone. This phase complements the quantitative findings by offering a deeper understanding of system usability, responsiveness, and practical implementation in real-world settings.

The fourth component, BERT model validation**,** focuses on evaluating the performance of the sentiment analysis module integrated into the system. A fine-tuned Bidirectional Encoder Representations from Transformers (BERT) model is used to classify textual feedback into positive, negative, or neutral sentiments. The model is trained and tested using labeled datasets, and its performance is assessed using standard evaluation metrics such as accuracy, precision, recall, and F1-score. Recent studies have shown that BERT-based models achieve high levels of accuracy in sentiment classification tasks due to their ability to capture contextual relationships in text data, making them suitable for analyzing large-scale customer feedback . This phase ensures that the artificial intelligence component of the system produces reliable and meaningful analytical outputs.

Overall, the developmental research design provides a comprehensive framework that integrates system creation and multi-level evaluation while maintaining clear distinctions between each methodological component. By separating system development, quantitative evaluation, qualitative analysis, and model validation, the study ensures methodological clarity, avoids overlap of processes, and enhances the reliability and validity of the findings. This structured approach supports the development of a robust, user-centered, and data-driven system for improving client satisfaction measurement in the public sector.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Research Locale*

This study was conducted in the Department of Education (DepEd) Surigao del Sur Division, located in the Caraga Region, Philippines. The division serves a geographically diverse population composed of coastal, upland, and urban communities, making it a dynamic setting for examining public service delivery and client satisfaction mechanisms. As a government agency mandated to provide efficient, transparent, and citizen-centered services, the division regularly interacts with various stakeholders, including teachers, school heads, parents, and external clients. These interactions generate substantial volumes of feedback data, particularly through the implementation of the Client Satisfaction Measurement (CSM) system, which is a key performance indicator aligned with national governance standards.

The locale is highly relevant to the present study as it actively adopts digital transformation initiatives aimed at improving service delivery. The integration of online platforms, feedback kiosks, and automated reporting systems reflects the division's commitment to modernizing its operations. Digital governance initiatives have been shown to enhance efficiency, transparency, and citizen engagement in public services, thereby improving overall service quality and satisfaction. In the context of DepEd Surigao del Sur, the increasing reliance on technology-based feedback systems provides an appropriate environment for implementing automated data aggregation, sentiment analysis, and machine learning models such as BERT to analyze client feedback.
Moreover, the selection of this locale was driven by existing operational challenges observed within the division's current feedback management practices. Despite the presence of digital tools, the process of consolidating, analyzing, and reporting client satisfaction data remains partially manual, leading to delays, inefficiencies, and potential inaccuracies in reporting. Similar issues have been identified in institutional settings where manual evaluation of feedback is time-consuming and prone to errors, thereby affecting timely decision-making and service improvement. These persistent concerns highlight the need for an automated and intelligent system capable of real-time data processing and analysis.

Fig 6 Location Map of the Study

The DepEd Surigao del Sur Division was therefore purposively selected as the research locale due to its readiness for digital innovation, availability of client feedback data, and the presence of practical challenges that necessitate technological intervention. The findings of this study are expected to directly benefit the division by providing a data-driven, automated solution for customer satisfaction measurement, while also offering a replicable model for other government agencies pursuing digital transformation and improved public service delivery.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Research Respondents/ Key Informants*

This study employed purposive sampling, a non-probability technique in which the researcher deliberately selects participants based on specific criteria to maximize the relevance and depth of information gathered (Creswell & Creswell, 2018; Patton, 2015; Tongco, 2007). This approach is appropriate for evaluative and mixed-methods research where participants must possess direct experience with or specialized knowledge of the phenomenon under study. Accordingly, three groups were identified: (1) client-respondents who directly interacted with the automated CSM system as service recipients; (2) key informants who oversee CSM implementation and frontline service delivery; and (3) IT experts who provided technical evaluation of the system's quality attributes.

Table 1 Research Respondents/ Key Informants

Participant Group	**Number**
Client-Respondent	50
Key Informants	8
IT Expert/ Sentiment Analysis Validators	3
Total	**61**

The fifty (50) client-respondents were selected using purposive sampling, a technique that allows the researcher to deliberately choose participants who meet specific criteria relevant to the study (Creswell & Creswell, 2018). In this study, the respondents were actual clients who transacted with the DepEd Surigao del Sur Division Office from February to March 2026 and provided feedback through the online CSM platform. Their transactions encompassed a range of frontline services, including liquidation of funds, submission of human resource (HR) reports, approval of program works, follow-up of applications and promotions, and procurement of resource materials for teaching, among others. It is important to clarify that some of these clients may also be internal personnel, such as teaching or non-teaching staff, who acted as service recipients during their transactions. Thus, they were treated as clients in the context of the study, since the evaluation focused on their experience as service users rather than their organizational roles.

The inclusion criteria for the client-respondents were: (1) must have completed a transaction with any frontline service office during the specified period; (2) must have accessed and submitted feedback through the online CSM platform; and (3) must have provided complete and valid responses. The exclusion criteria included: (1) individuals who did not complete the feedback form; (2) responses with missing or inconsistent data; and (3) individuals who did not transact within the defined timeframe. These criteria ensured that only respondents with actual, relevant, and usable experience were included, making them appropriate evaluators of the system.

The use of these respondents is justified because they had direct and recent interaction with both the service process and the feedback system, allowing them to evaluate the system’s usability, efficiency, and responsiveness based on real experience. In public service evaluation, actual service users are considered the most reliable source of data since their perceptions reflect real service encounters and satisfaction levels.

The eight (8) key informants were selected using purposive sampling, specifically criterion-based selection, wherein participants are chosen based on their knowledge and involvement in the subject of the study (Creswell & Poth, 2018). The inclusion criteria for key informants were: (1) must be a department head, unit head, or designated personnel involved in frontline service delivery or CSM implementation; (2) must have at least one (1) year of experience in handling or supervising related processes; and (3) must be willing to participate in the study. The exclusion criteria included personnel without direct involvement in service delivery or those with insufficient experience in the CSM process. These informants were appropriate evaluators because they possess process-level understanding and administrative insight, enabling them to identify operational strengths, weaknesses, and system requirements beyond what is observable by clients alone.

The three (3) IT experts were selected through expert sampling, a form of purposive sampling used when specialized knowledge is required (Etikan, I, 2016). The inclusion criteria for IT experts were: (1) must have professional experience in system development, software engineering, or database management; (2) must have at least two (2) years of relevant experience; and (3) must have prior involvement in system evaluation or development. The exclusion criteria included individuals without formal technical expertise or those with no experience in system assessment. These experts were appropriate evaluators because they provided objective and technical assessment of the system’s quality attributes, such as functionality, performance efficiency, security, and maintainability.

The sampling procedure involved identifying and selecting respondents based on these inclusion and exclusion criteria. For the client-respondents, the researcher extracted valid entries from the online CSM database within the specified timeframe. For key informants, individuals were identified based on their roles and responsibilities within the division. For IT experts, selection was based on professional qualifications and technical expertise.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

The sample sizes were considered sufficient for the purposes of the study. The fifty (50) client-respondents were adequate as they represented actual system users within a defined operational period, ensuring experience-based and relevant data. Faulkner (2003) and Virzi (1992) established that samples of 20 or more users sharing similar task contexts are sufficient to capture critical usability issues and yield reliable performance assessments. The eight (8) key informants were sufficient to achieve data saturation, as Guest, Bunce, and Johnson (2006) confirmed that saturation is typically reached between six and twelve interviews among participants with similar roles. Creswell and Poth (2018) further affirm that depth of experience, rather than quantity, ensures the richness of qualitative findings. The three (3) IT experts were appropriate for technical validation, consistent with Polit and Beck (2006) and Lynn (1986), who recommend a minimum of three to five expert evaluators for content validation studies.

- *Research Instrument*

This study utilized three (3) research instruments to gather the necessary data: (1) an adopted and validated survey questionnaire for the Online Client Satisfaction Survey based on ISO 25010:2011 Software Quality Standard of North Eastern Mindanao State University (appendix A), (2) a researcher-made interview guide questionnaire for key informants (appendix B), and (3) a sentiment analysis validation tool for evaluating the performance of the BERT-based classification model (appendix C). These instruments were carefully selected and developed to ensure alignment with the objectives of the study, particularly in assessing system effectiveness, extracting qualitative insights, and validating automated sentiment classification.

The first instrument was an adapted survey questionnaire designed to measure the effectiveness of the online client satisfaction survey system used by research respondents who had transactions with the DepEd Surigao del Sur Division. The instrument was anchored on the ISO 25010:2011 Software Quality Model form North Eastern Mindanao State University and subjected with content validation by the IT expert. The tool evaluated software systems based on characteristics such as functionality, reliability, usability, efficiency, maintainability, and security. The questionnaire consisted of structured items rated using a Likert scale to determine the extent of system effectiveness. This tool enabled the quantitative assessment of the automated customer feedback system, particularly in relation to real-time data aggregation and reporting.

The second instrument was a researcher-made interview guide questionnaire intended for key informants, particularly representatives from different departments within the division. This instrument contained open-ended questions that explored the informants' experiences, perceptions, and insights regarding the implementation of the automated client satisfaction system, including its compliance with ARTA standards and its effectiveness in improving service delivery. The interview guide was designed to elicit in-depth qualitative data that complemented the survey findings and provided contextual understanding of system utilization and challenges.

The third instrument was the sentiment analysis validation tool used to evaluate the performance of the BERT-based sentiment classification model. This tool involved a set of client satisfaction comments that were manually labeled by three (3) IT experts (human labeling) and compared with the system-generated labels. The tool was used to compute evaluation metrics such as Accuracy, Precision, Recall, and F1-Score, which are standard measures in assessing classification performance. This instrument ensured that the automated sentiment analysis component of the system was valid, reliable, and aligned with actual human judgment in interpreting customer feedback.

For face and content validation, all instruments were subjected to expert evaluation prior to data collection. The validation process involved two groups of experts based on the nature of each instrument. The survey questionnaire and sentiment analysis validation tool were evaluated by IT experts with expertise in system development, software quality standards, and sentiment analysis. The interview guide questionnaire, being a qualitative instrument, was evaluated by qualitative research experts with backgrounds in research methodology, social science inquiry, and educational or public service assessment. The inclusion criteria for the validators were: (1) at least a master's degree in a relevant field such as Information Technology, Computer Science, Education, or related disciplines; (2) demonstrated experience in research, system development, or instructional evaluation; and (3) familiarity with survey instrument design, sentiment analysis, or public service delivery systems. The validators evaluated the instruments in terms of clarity, relevance, organization, and alignment with the objectives of the study. Their comments and recommendations were systematically reviewed and incorporated, resulting in improvements in item construction, wording, and overall instrument coherence to ensure validity and appropriateness for data collection.

To establish validity and reliability, pilot testing was selectively conducted prior to the actual data gathering. The adopted survey questionnaire did not undergo pilot testing since it was based on the established ISO 25010:2011 Software Quality Standard, which has been widely validated and utilized in evaluating software systems. However, pilot testing was conducted for the researcher-made semi-structured interview guide questionnaire to assess the clarity, relevance, and comprehensibility of the questions. This process involved administering the interview guide to a small group of participants similar to the intended key informants. Feedback from the pilot testing led to refinements in question phrasing, sequencing, and structure to ensure that the items effectively elicited meaningful and accurate responses.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

For the sentiment analysis validation tool, reliability was ensured through inter-rater agreement among the IT experts who performed manual labeling of customer feedback. The consistency between human-labeled data and system-generated classifications was evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score. The results demonstrated that the BERT-based model produced reliable and consistent classifications, thereby confirming the robustness of the automated sentiment analysis component of the system.

➢ *Data Gathering Procedure*

The data gathering procedure of the study was conducted systematically to ensure clarity in the separation of system deployment, data collection, and model development processes. The procedure focused primarily on the collection of client satisfaction data for system evaluation, while a separate dataset preparation process was employed for the BERT model training and validation.

The study was conducted within the DepEd Surigao del Sur Division from February 2026 to March 2026, which represents the actual period during which client feedback data were collected through the deployed system. Prior to this, preparatory activities such as system development, testing, and initial dataset preparation for machine learning were undertaken to ensure readiness for implementation. This alignment ensures consistency between the study period and the operational status of the system.

The data gathering process began with the deployment of the Client Satisfaction Measurement (CSM) module**,** which was integrated into the existing Human Resource Information System (HRIS). The system was made accessible to clients transacting with various offices in the division. Upon completion of their transactions, clients were invited to provide feedback through an online survey platform embedded within the system. The survey instrument consisted of two components: (1) structured responses using a Likert-scale format to measure service quality indicators, and (2) open-ended responses that allowed clients to express their experiences and suggestions in textual form. This approach aligns with public sector service quality assessment practices, where both quantitative and qualitative feedback are essential for comprehensive evaluation .

The respondents included clients who had actual transactions with the division offices during the specified period. Only those who voluntarily completed the online feedback form were included in the dataset. The system automatically captured and stored all responses in a centralized database, eliminating the need for manual encoding and ensuring data accuracy and completeness. This automated data collection process supports efficiency and transparency in public service evaluation, consistent with digital governance practices that emphasize real-time data utilization for decision-making .

It is important to clarify that two distinct datasets were utilized in the study**.** The first dataset consisted of the actual client feedback collected from February to March 2026, which was used exclusively for system evaluation and analysis of client satisfaction**.** The second dataset was used for BERT model training and validation, which included preprocessed and annotated textual data. This training dataset may include a combination of collected feedback and supplementary labeled data to ensure sufficient volume and variability required for machine learning tasks. This distinction is necessary because sentiment analysis models, particularly transformer-based models such as BERT, require adequately sized and labeled datasets to achieve reliable performance .

The collected data for system evaluation were then prepared for analysis by organizing the structured responses into statistical formats and storing textual feedback for subsequent processing. No algorithmic processing was performed at this stage, as the purpose of this phase was purely data collection. The technical procedures involving preprocessing, tokenization, and sentiment classification using BERT were conducted separately and are discussed under the system and algorithm components of the methodology.

➢ *System Implementation*

- *System Model of the Online CSM*

Figure 7 presents the system model of the Online Client Satisfaction Measurement (CSM), which operates as a sentiment analysis pipeline using the Bidirectional Encoder Representations from Transformers (BERT) model to classify client feedback into Positive, Neutral, or Negative sentiment categories. The process begins when textual feedback submitted by clients through the online CSM platform is collected and stored as input data. These responses served as the primary dataset for sentiment classification and automated report generation.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

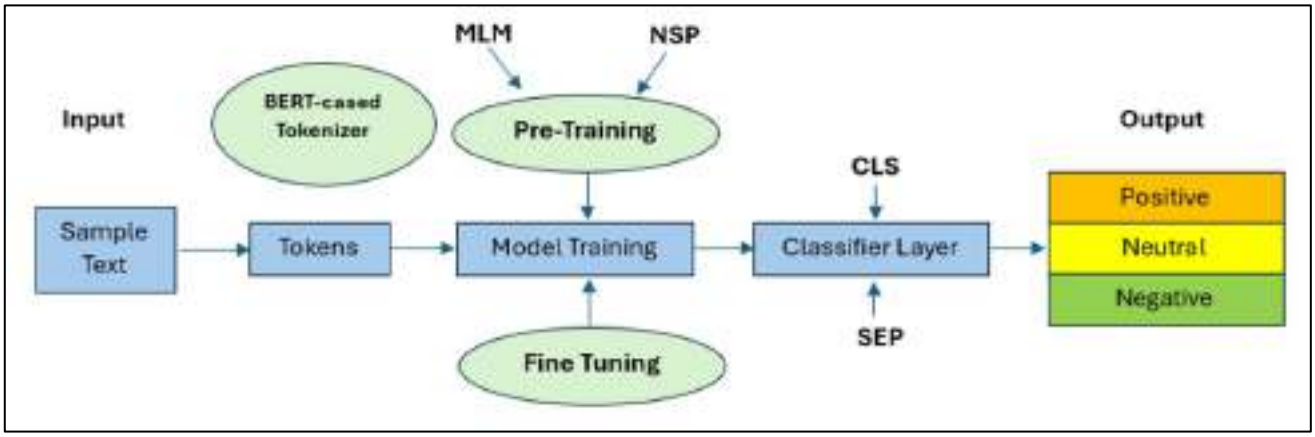


Fig 7 System Model of Online CSM

Prior to model training, the collected textual data underwent preprocessing to improve data quality and consistency. During this stage, unnecessary characters such as punctuation marks, symbols, extra spaces, and irrelevant special characters were removed. Text normalization was performed by converting all words into lowercase and standardizing frequently used terms. Considering that client responses included multilingual inputs such as English, Filipino, and local dialects, the preprocessing stage harmonized these variations to ensure consistent textual representation. This process minimized noise and improved the interpretability of the input data before model training.

After preprocessing, the cleaned text was processed using BERT's tokenizer**,** which converted textual input into tokenized representations. The tokenizer segmented the text into tokens and inserted special tokens such as [CLS] to represent the entire sequence and [SEP] to indicate sentence boundaries. Each toke**n** was transformed into numerical representations, including input IDs and attention masks, which served as inputs to the BERT model. This process enabled the model to capture contextual relationships among words within each feedback entry.

Following tokenization, the dataset was partitioned into three subsets using a 30-50-20 ratio. Specifically, 30% of the 350-entry corpus (105 entries) was allocated for model training, 50% (175 entries) for validation, and the remaining 20% (70 entries) for testing. The training subset was used to fit the BERT classifier parameters on the labeled feedback corpus. The validation subset monitored model convergence, inspected class behavior, and guided tuning decisions during fine-tuning. The testing subset was held out for final performance evaluation on unseen examples. This three-way partitioning prioritized the validation set to enable closer monitoring of model behavior and to reduce the risk of overfitting given the limited size of the corpus. The complete dataset framing report, statistical summary, and illustrative workflow used in this process are presented in appendix H.

The model training stage utilized a pre-trained BERT model**,** which was fine-tuned using the labeled CSM dataset. A classification layer was added on top of the BERT architecture**,** and the model parameters were adjusted through supervised learning. During training, the model iteratively updated its weights based on the training data to improve its ability to classify feedback into Positive, Neutral, and Negative categories. Fine-tuning allowed the model to adapt its contextual understanding specifically to client satisfaction data, which is consistent with established BERT-based sentiment analysis approaches .

After training, the model processed validation inputs through the classification layer. The output representation from the [CLS] token was used as the aggregate representation of each feedback entry and passed to the classifier to generate the final sentiment label. The system then assigned each feedback entry to one of the three sentiment categories based on contextual analysis.

This system model demonstrates how the Online CSM integrates preprocessing, tokenization, dataset partitioning, and BERT fine-tuning to produce automated sentiment classification. The process enables efficient analysis of textual responses and supports real-time, data-driven reporting of client satisfaction.

- *System Development Approach*

The system was developed using the Agile Software Development Life Cycle, which allows iterative development and continuous refinement of system components.

During the planning phase, stakeholder requirements were gathered, and the scope of the system was defined. The existing Client Satisfaction Measurement process was analyzed to identify inefficiencies and determine system requirements.

The analysis phase focused on documenting system workflows, including feedback collection, sentiment classification, and report generation. This phase also established how textual inputs would be processed through the integrated BERT-based sentiment analysis component.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

In the design phase, the system architecture was developed to ensure efficient integration of all components. The database schema was structured using MySQL to support data storage and retrieval. The user interface was designed to provide ease of use, while the Laravel framework supported backend processing, system logic, and security.

During development, the system components were implemented, including the online feedback submission module, automated data aggregation module, and sentiment analysis engine. The BERT-based classification mechanism was integrated into the system to enable automated processing of textual feedback.

➢ *System Testing, Technical Validation, Deployment, and Maintenance*

The system underwent comprehensive testing to evaluate functionality, reliability, and performance. Unit testing was conducted for individual modules, integration testing verified the interaction among system components, and system testing evaluated the overall workflow. Performance testing assessed system responsiveness under varying usage conditions.

To support controlled and repeatable technical validation of the sentiment analysis component, the researcher utilized data seeding within the Laravel framework. Data seeding is a built-in feature that allows the database to be populated with predefined records using PHP-based seed classes stored in the database/seeders directory. The default *Database Seeder* class was used to execute custom seed scripts and control the insertion of test data.

Using this approach, the researcher generated and inserted 350 textual feedback entries drawn from a seeded random sample of actual CSM survey records collected between February and March 2026 into the system database. These seeded entries consisted of varied expressions representing positive, neutral, and negative sentiments and included multilingual content such as English, Filipino, and local language inputs. The generated dataset served as simulated client feedback to test the system's ability to process diverse textual inputs.

After insertion, the seeded data were automatically processed by the system through the preprocessing, tokenization, and BERT-based classification pipeline. The system generated corresponding sentiment outputs, which were stored and aggregated for reporting. This procedure enabled the researcher to evaluate the functionality of the sentiment analysis component within the actual application environment.

It is important to note that the seeded data were used specifically for system testing and technical validation to ensure consistent and repeatable evaluation conditions. This approach allowed the researcher to verify system performance without relying solely on real-time user inputs.

Following successful testing and validation, the system was deployed in the Department of Education, Division of Surigao del Sur. Maintenance procedures and update mechanisms were established to ensure continuous functionality, system security, and adaptability to future requirements.

- *System Architecture*

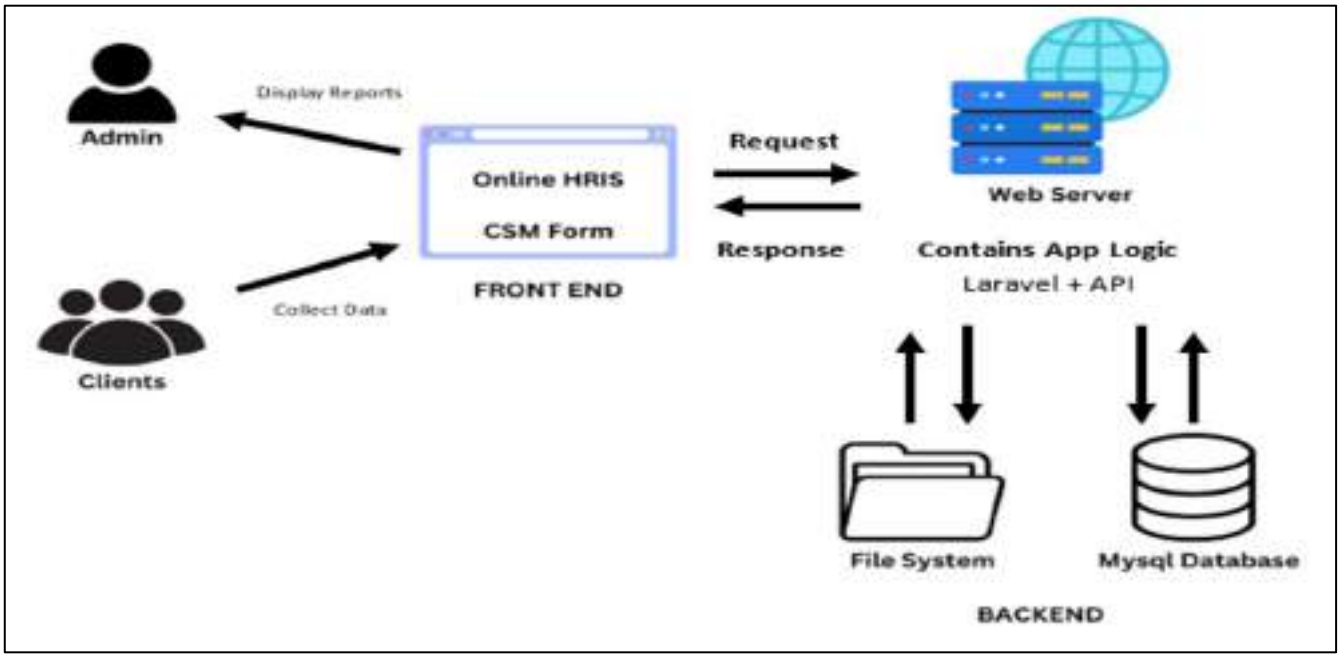


Fig 8 System Application Architecture

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

The automated system consists of interconnected components that facilitate efficient data processing and reporting. Users, including clients, administrators, and analysts, interact with the system through a web-based interface to submit feedback and access reports. All responses and processed sentiment data are stored in a centralized database.

An artificial intelligence processing module applies Natural Language Processing techniques and the BERT model to analyze textual feedback. The processed outputs are forwarded to the data aggregation and analysis engine, which structures the data and computes Customer Satisfaction indicators aligned with the SERVQUAL model.

The automated reporting module generates dashboards and downloadable reports that present real-time insights. The system also incorporates Anti-Red Tape Authority compliance principles to ensure transparency, accountability, and efficient service delivery. This architecture supports real-time processing and informed decision-making.

- *Case Diagram*

As shown in Figure 9, the system supports multiple user roles, including clients, system administrators, and analysts.

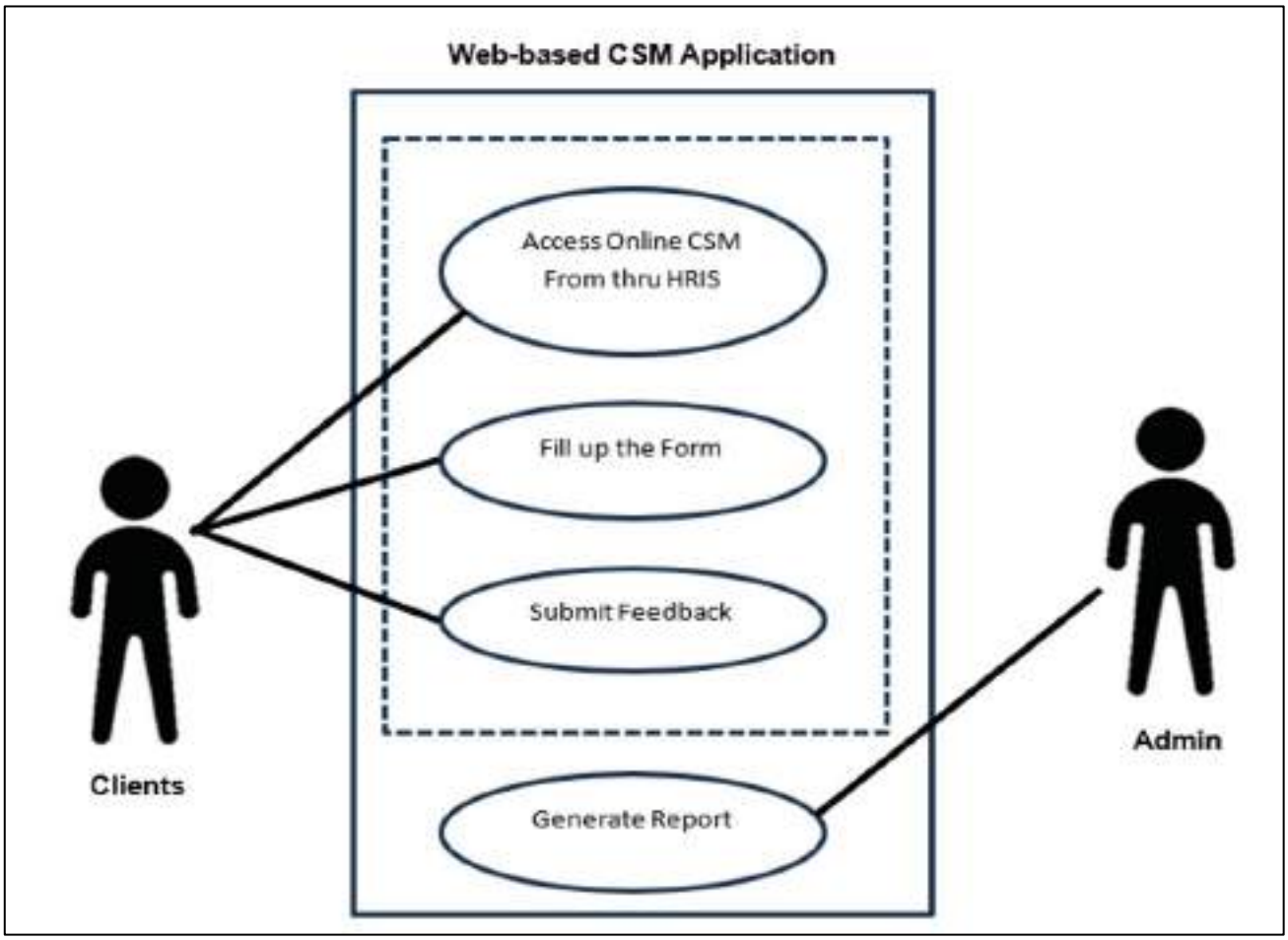


Fig 9 Case Diagram

Clients submit feedback through the Online CSM Form, which is processed by the sentiment analysis engine using NLP and BERT-based classification. The processed data are stored in the system database for further analysis.

The System Administrator manages user access, monitors system performance, and ensures compliance with ARTA standards. Analysts access the Automated Reporting Module to generate reports and evaluate client satisfaction trends. The system enables real-time processing, reducing manual workload and improving accuracy in feedback analysis.

- *Data Flow Diagram (DFD)*

The Data Flow Diagram (DFD) as shown in Figure 10 presents the automated client satisfaction measurement system's data movement starting from client inputs, leading to report creation and safeguarding data consistency. Clients generate survey responses using Online CSM Form Processing (Process 1.0), validating and standardizing the data for automation readiness. The data proceeds to AI Sentiment Analysis (Process 2.0), which uses BERT and NLP techniques to classify sentiments while extracting significant insights from the data.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

An algorithm establishes the processed sentiment information as feedback entries in the Feedback Database (D1) to operate as the structured data's central storage and retrieval point. The system utilizes Report Generation (Process 3.0) to create meaningful reports showing changes in client satisfaction patterns and sentiment direction. Statistics from our system help system administrators create reports that they employ as tools to develop operations and enhance service quality.

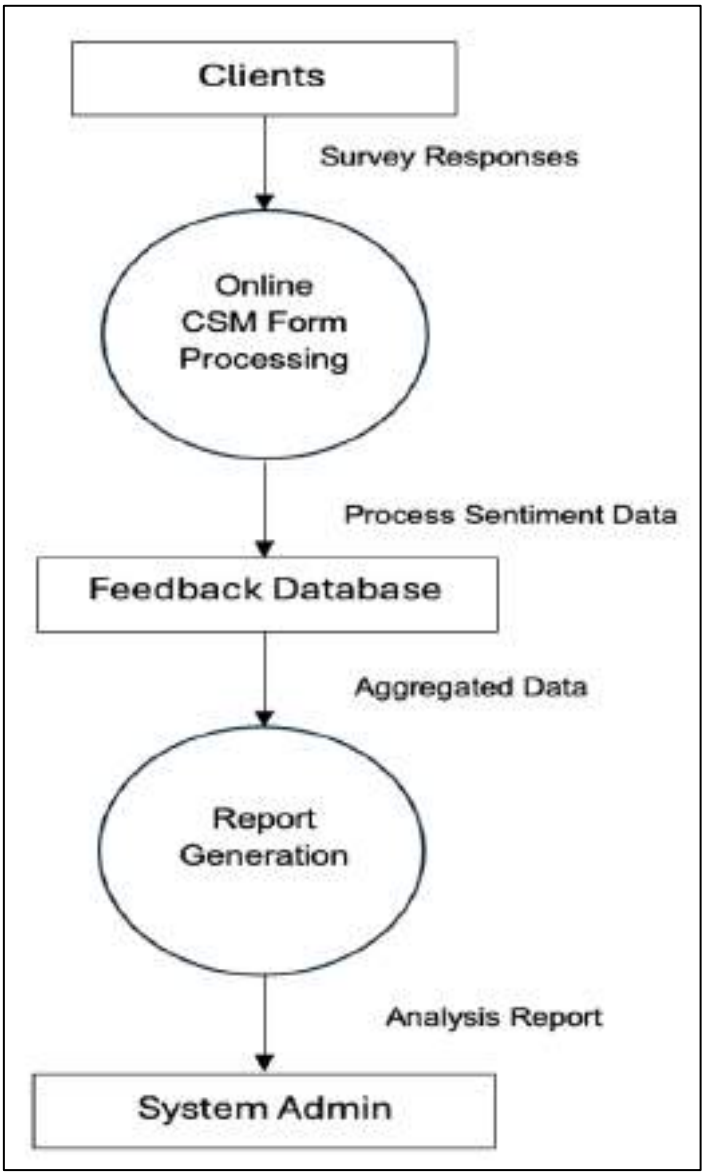


Figure 10. Data Flow Diagram

The Yourdon notation expresses the DFD through rectangles depicting external entities, including Clients and System Admin, as well as circular processes and directional arrow representations of data transfers. The organized system enables correct feedback evaluation and computerized automation to maximize operational speed. The system improves high-quality service delivery at DepEd Surigao del Sur while helping administrators use data to enhance decision-making and service improvement initiatives.

- *Test Plan*

A test plan was developed to systematically evaluate the system's reliability, functionality, usability, and performance. Unit testing was conducted to verify the correctness of individual modules, while integration testing ensured that interconnected components operated cohesively and accurately. These procedures confirmed that the system's core functions, data flow, and component interactions were properly executed.

System testing further validated the system's performance across actual workflow scenarios, whereas user acceptance testing determined its alignment with stakeholder requirements. In addition, performance testing assessed the system's stability, responsiveness, and capacity to sustain expected peak usage under varied operational conditions. Overall, the testing process established the system's readiness for reliable and effective deployment.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Table 2 Online CSM System Test Plan

Test Case ID	Test Scenario	Expected Result
TC-01	User Access	Successful account creation and access
TC-02	Login Authentication	Correct login credentials grant access
TC-03	Feedback Submission	The system records and confirms submitted feedback in real-time.
TC-04	Sentiment Analysis Processing	The system accurately classifies sentiment as positive, neutral, or negative using BERT-based NLP.
TC-05	Report Generation	The system generates structured, downloadable reports with real-time data visualization.
TC-06	Data Aggregation & Accuracy	The system consolidates feedback accurately and ensures consistency in the database.
TC-07	AI Processing Performance	The system processes and categorizes feedback within predefined response times.
TC-08	Compliance with ARTA Standards	The system generates reports and feedback insights in compliance with ARTA requirements.
TC-09	System Security & Access Control	Unauthorized access attempts are blocked, and user roles function correctly.
TC-10	System Load & Performance	The system handles high user traffic and data processing without performance degradation.

- *Ethical Considerations*

This study adhered to established ethical standards to ensure that the rights, dignity, and welfare of all participants and key informants were protected throughout the research process. The conduct of the study was guided by the principles of the Belmont Report, which include respect for persons, beneficence, and justice. These principles served as the foundation in ensuring that the research was carried out responsibly and with integrity.

Respect for persons was observed by ensuring voluntary participation and informed decision-making among all respondents. Prior to data collection, informed consent was obtained from teaching and non-teaching personnel who utilized the client satisfaction survey kiosk, as well as from key informants who participated in the interviews. The participants were clearly informed about the purpose of the study, the procedures involved, and their rights as respondents. Participation was entirely voluntary, and respondents were given the freedom to withdraw at any point without any negative consequences. Anonymity was maintained by not requiring personal identifiers, and confidentiality was ensured by presenting the results in summarized form.

The principle of beneficence was applied by ensuring that the study posed no harm to the participants. The procedures were carefully designed to avoid any form of physical, psychological, or social risk. Special attention was given to the handling of feedback data, particularly those that may contain sensitive opinions. The use of automated sentiment analysis and data aggregation was implemented with caution to prevent misinterpretation or misuse of responses. All collected data were securely stored and accessed only by the researcher. The study aimed to contribute positively by improving the efficiency of client satisfaction measurement and enhancing service delivery in the DepEd Surigao del Sur Division.

Justice was ensured through the fair selection and treatment of participants. Respondents were chosen based on their actual engagement with the client satisfaction system, ensuring that those directly involved were given equal opportunity to participate. There was no form of discrimination in the selection process, and all participants were treated equitably. The inclusion of IT experts in evaluating the sentiment analysis model was based on their qualifications and expertise, ensuring fairness and credibility in the validation process.

Furthermore, specific protocols were followed regarding data retention and disposal. All collected data will be retained securely for a period of three (3) years after the completion of the study for verification and academic purposes. After this period, all digital files will be permanently deleted using secure data erasure methods, and any printed materials will be shredded to prevent any possibility of data recovery or unauthorized disclosure.

In addition, ethical responsibility in the use of artificial intelligence and sentiment analysis was carefully considered. The study ensured transparency in the processing and classification of textual data, and the algorithm was used strictly as an analytical tool rather than a decision-making authority. Measures were taken to minimize algorithmic bias by applying consistent preprocessing and validation techniques. This ensured that the interpretation of client feedback remained fair, accurate, and aligned with the actual responses of participants.

Overall, the study maintained ethical integrity by implementing concrete measures for anonymization, secure data storage, controlled access, responsible data usage, defined retention policies, and proper disposal procedures. These practices ensured that participant rights were protected and that the research was conducted in a responsible, transparent, and ethically sound manner.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Data Analysis*

The qualitative data gathered in this study were analyzed using a thematic analysis approach, which is appropriate for identifying, organizing, and interpreting patterns of meaning from interview responses. This method was applied exclusively to the fifth Statement of the Problem, which examined how the system ensures compliance with ARTA standards for client satisfaction reporting. Thematic analysis allowed the researcher to systematically explore the perceptions of key informants regarding the system's capacity to address the reporting requirements of Republic Act No. 11032.

The analysis followed a structured coding framework. Initially, all responses were transcribed and subjected to data familiarization, where the researcher repeatedly reviewed the textual data to gain a comprehensive understanding of the participants' insights. This was followed by open coding, wherein significant statements and recurring ideas related to ARTA compliance, transparency, accountability, and standardized reporting were identified. These initial codes were generated inductively from the data to ensure that the findings reflected the actual experiences and perspectives of the key informants.

Subsequently, axial coding was employed to cluster related codes into broader categories aligned with the fifth Statement of the Problem. Codes concerning transparency, standardization of reporting formats, timeliness of submission, and adherence to government service standards were organized into themes on ARTA compliance and regulatory alignment. Finally, selective coding was conducted to refine these categories into core themes that directly addressed how the system supports ARTA-compliant client satisfaction reporting. The emerging themes were further validated against the responses of all eight key informants to ensure consistency and credibility of the qualitative findings.

➢ *Statistical Treatment*

The data collected from the online client satisfaction survey and textual customer feedback were analyzed using descriptive and model evaluation techniques to assess the level of customer satisfaction, determine the effectiveness of automated sentiment classification using the BERT model, and provide empirical bases for improving the Client Satisfaction Measurement (CSM) system. The following statistical treatments were applied:

- *Mean.*

The mean was used to determine the level of customer satisfaction based on the responses gathered from the online survey. It provided a numerical representation of the overall perception of clients regarding the services of the organization. The computed mean scores were interpreted using an appropriate Likert scale to describe the extent of satisfaction. This statistical measure served as the basis for identifying areas of strength and aspects that require improvement in service delivery.

- *Accuracy, Precision, Recall, and F1-Score.*

These evaluation metrics were used to measure the performance of the BERT (Bidirectional Encoder Representations from Transformers) model in classifying customer feedback into sentiment categories such as positive, negative, and neutral. Accuracy determined the overall correctness of the classification model, while Precision measured the proportion of correctly predicted positive observations. Recall assessed the model’s ability to correctly identify actual positive cases, and the F1-Score provided a balanced measure of Precision and Recall. These metrics were derived from the confusion matrix, which includes true positive, true negative, false positive, and false negative values. The use of these statistical measures ensured a comprehensive evaluation of the model’s effectiveness in analyzing customer sentiments and enhancing the reliability of automated client satisfaction reporting.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

CHAPTER FOUR
PRESENTATION, ANALYSIS, AND INTERPRETATION OF DATA

This chapter presented the results of the study on the automated analysis of data aggregation and reporting of client satisfaction measurement in DepEd Surigao del Sur, wherein the data were systematically organized, analyzed, and interpreted using a mixed-methods approach based on responses gathered through the online Client Satisfaction Measurement (CSM) form, and the findings were examined in relation to the research objectives and relevant literature, covering the respondents' demographic profile, perceptions of technological solutions, effectiveness of algorithms, improvements in accuracy and reliability, classification of feedback using the BERT model, and compliance with ARTA standards.

- *Technological Solutions Used to Streamline the Collection, Storage, and Processing of Survey Data*

The automated Client Satisfaction Measurement (CSM) system developed in this study addresses the longstanding limitations of the manual survey process in the DepEd Surigao del Sur Division Office, particularly the reliance on paper-based forms, manual encoding of responses, and fragmented data consolidation across departments. To resolve these inefficiencies, the system integrates a comprehensive set of technological solutions that cover the entire feedback lifecycle, from data collection and validation to storage, analysis, reporting, and security. The complete system interface and its functional components are presented in appendix E.

The first technological solution concerns data collection. The system employs a web-based survey platform accessible through online kiosks stationed within the Division Office, enabling clients to directly input their satisfaction feedback at the point of service immediately after completing a transaction. Rather than relying on paper forms that require subsequent manual encoding by administrative staff, clients interact with a digital interface that guides them through structured survey items covering service quality dimensions such as responsiveness, reliability, and staff courtesy. The interface is designed to be visually clean, straightforward, and navigable with minimal instruction, ensuring that clients across varying levels of digital literacy can complete the survey without difficulty. This design approach directly reduces one of the most common barriers to survey participation in public service settings, which is the perceived complexity and inconvenience of the feedback process. The accessibility and usability of this solution are consistent with the findings of Omas-as et al. (2024), who demonstrated that unified digital feedback platforms significantly improve stakeholder participation and satisfaction analysis in educational institutions. Cervantes et al. (2024) similarly confirmed that perceived ease of use and usefulness are the most critical factors driving sustained adoption of digital systems among DepEd employees.

Building on the collection mechanism, the system processes submitted data through real-time automated recording, which captures and stores each client response in the database immediately upon submission. This eliminates the multi-step manual process of collecting paper forms, encoding responses into spreadsheets, and consolidating data from multiple offices, which previously introduced delays and increased the likelihood of encoding errors. With automated recording, data are available for administrative review as soon as they are submitted, and authorized personnel can access the information through the admin dashboard using filters for office, service type, date range, and transaction category, allowing for precise and efficient data retrieval. To further ensure the integrity of each submission, the system enforces automated validation mechanisms at the point of entry. These include required fields that prevent clients from skipping survey items, input type restrictions that limit responses to valid formats, dropdown menus that standardize options, and real-time error prompts that notify clients of incomplete or inconsistent entries before the form is submitted. This means that only complete and valid responses are accepted into the database, substantially reducing the need for post-collection data cleaning. Mozumder et al. (2024) confirmed that AI-driven automation significantly enhances operational efficiency and reduces processing delays in customer satisfaction systems, while Rahman and Maryani (2024) demonstrated that automated processing mechanisms improve the reliability and consistency of feedback data collected at scale.

Once responses are validated and recorded, they are stored in a centralized relational database that serves as the single, unified repository for all CSM data across the Division Office. Unlike the previous system where data were stored in separate files or spreadsheets maintained by individual offices, the centralized database consolidates all records in a structured format that supports consistent data organization, prevents duplication, and ensures that information is retrievable at any point without depending on the availability of specific personnel. The database is designed with a relational structure that links records by office, service type, respondent category, and submission period, enabling administrators to perform cross-office comparisons and longitudinal analyses with relative ease. Furthermore, the architecture is built to be scalable, meaning it can accommodate increasing volumes of data as the system is expanded to cover more offices, service units, or time periods without requiring fundamental restructuring. Aliazas et al. (2024) found that centralized, integrated database systems improve transparency and operational efficiency in Philippine educational institutions, particularly when the database is designed to complement and extend existing institutional workflows rather than replace them entirely.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

With all data consolidated in a unified repository, the system is further equipped to generate automated summary reports and visual data presentations directly from the stored information through its administrative dashboard. Rather than requiring staff to manually extract, organize, and format data into reports, the system automatically computes satisfaction ratings, aggregates responses by office and service type, and presents the results through charts, graphs, and summary tables that are updated in real time as new submissions are recorded. Administrators can select specific filters to generate reports by period, office, or service category, and the outputs are exportable in formats that align with the reporting requirements of the Anti-Red Tape Authority (ARTA), ensuring compliance without additional manual formatting. This reporting capability directly addresses one of the most time-consuming aspects of the existing CSM process, which is the preparation of consolidated reports for submission to oversight bodies. Doria (2025) emphasized that embedding analytics and automated reporting tools within survey systems substantially improves reporting efficiency and enables evidence-based decision-making, particularly in institutions where data must be regularly submitted to regulatory agencies.

Across all of these functional components, the system incorporates a layered set of security and privacy safeguards to protect the confidentiality and integrity of client feedback data throughout the entire data lifecycle. These safeguards include user authentication through secure login credentials, role-based access control that limits system privileges based on the user's position and responsibilities, data encryption for both stored and transmitted information, secure server storage, and automated backup procedures that protect against data loss. These mechanisms collectively ensure that only authorized personnel can access administrative functions, view identifiable data, or generate and export reports, thereby preventing unauthorized access and potential misuse of sensitive client information. The system was developed in compliance with the Data Privacy Act of 2012, and its security features reflect the principle that client trust is a prerequisite for sustained participation in feedback systems. Parkavi et al. (2023) demonstrated that privacy-preserving design can be implemented in machine learning-based systems without compromising analytical performance, while Aliazas et al. (2024) confirmed that robust security features are among the most significant factors contributing to user trust and adoption of automated systems in Philippine educational and government institutions.

Through the integration of these solutions, the automated CSM system provides the DepEd Surigao del Sur Division Office with a comprehensive, end-to-end digital platform that resolves the inefficiencies of the manual process by automating each stage of the feedback workflow. The result is a system that not only reduces the administrative burden associated with data collection and reporting but also produces more accurate, complete, and timely satisfaction data that can meaningfully support service improvement and regulatory compliance.

- *CSAT Algorithm Effectively and Efficiently Generates Real-Time Reports and Automatically Analyzes Data*

The Customer Satisfaction (CSAT) algorithm was utilized in this study as the primary computational approach for effectively and efficiently generating real-time reports and automatically analyzing structured client feedback data. Through its capability to process rating-based responses, compute satisfaction scores, and consolidate results instantly, the algorithm enabled the transformation of raw survey inputs into meaningful and actionable information. This supports timely decision-making and significantly improves the efficiency of client satisfaction reporting.

To further strengthen the automated analysis process, the system integrates both CSAT computation for structured data and a BERT (Bidirectional Encoder Representations from Transformers) model for unstructured textual feedback**.** The CSAT computation facilitates automatic aggregation of numerical ratings into measurable satisfaction scores, allowing real-time generation of reports and visualization of trends. Studies have shown that automated and machine learning–driven approaches enhance customer satisfaction analysis by improving prediction accuracy and enabling efficient processing of large datasets (Mozumder et al., 2024; Manderscheid & Lee, 2024).

For textual feedback, the system utilizes BERT, a transformer-based deep learning model that captures contextual meaning in language. Unlike traditional models, BERT processes text bidirectionally, allowing it to interpret nuanced and context-dependent expressions. This enables the classification of open-ended responses into Positive, Neutral, or Negative sentiments, thereby complementing the numerical outputs generated by the CSAT algorithm. Empirical studies confirm that BERT consistently outperforms conventional machine learning techniques in sentiment classification and achieves higher accuracy in extracting insights from unstructured data (Rahman & Maryani, 2024)

Moreover, BERT-based approaches are highly applicable in multilingual contexts such as the Philippines, particularly in handling code-mixed and low-resource language data. Research shows that transformer-based models maintain strong performance across evaluation metrics such as accuracy, precision, recall, and F1-score even in linguistically complex environments (Maceda et al., 2023; Tomas et al., 2024; Laureta et al., 2025; Goma et al., 2024). This integration of structured and unstructured data processing ensures a more comprehensive and reliable automated reporting system

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Table 3 Online Client Satisfaction Survey Results

Indicators	Mean	Interpretation
Functionality	4.29	Very Great Extent
Reliability	4.33	Very Great Extent
Usability	4.35	Very Great Extent
Efficiency	4.60	Very Great Extent
Maintainability	4.75	Very Great Extent
Security	4.77	Very Great Extent

The results of the evaluation further demonstrate how the CSAT algorithm functioned within the system, as reflected in the clients' assessment of the kiosk . Among the indicators**,** security obtained the highest mean (4.77, Very Great Extent), while functionality obtained the lowest mean (4.29, Very Great Extent)**.**

The highest result in security indicates that the system was perceived as highly capable of protecting user data, ensuring confidentiality, and maintaining the integrity of responses during automated processing. This suggests that the CSAT algorithm, together with the integrated sentiment analysis component, operated within a secure environment that supports trustworthy real-time data analysis. This is critical because automated reporting systems rely heavily on data accuracy and protection. This finding is supported by Rahman and Maryani (2024), who emphasized that advanced machine learning models such as BERT enhance the reliability and integrity of customer feedback processing. Similarly, Alqurafi and Alsanoosy (2024) highlighted that automated sentiment analysis tools enable organizations to securely process large volumes of customer data while generating meaningful outputs.

The implication of this finding is that the effectiveness of the CSAT algorithm is reinforced by a secure system environment that promotes user trust and encourages honest participation. When respondents are confident in data privacy, the quality of feedback improves, leading to more reliable reports and better-informed decision-making.

On the other hand, the lowest result in functionality, although still interpreted as Very Great Extent, suggests that certain system features or performance aspects may require further enhancement. This indicates that while the CSAT algorithm effectively performs automated computation and reporting, there may be minor limitations in system responsiveness, flexibility, or user interaction. Functionality in automated systems involves not only correct computation but also adaptability and alignment with user needs.

This result may be attributed to the complexity of integrating multiple computational processes, including structured CSAT scoring and unstructured sentiment classification. Li et al. (2024) noted that BERT-based models require careful optimization and fine-tuning to ensure consistent performance across datasets. Similarly, Bashiri and Naderi (2024) explained that transformer-based sentiment analysis systems may encounter challenges when handling nuanced or ambiguous data despite their overall efficiency. These findings suggest that system performance depends not only on the algorithm itself but also on the quality of its implementation and optimization.

The implication of the lower functionality rating is the need for continuous system enhancement to fully maximize the capabilities of the CSAT algorithm. Improving interface design, system responsiveness, and processing efficiency can further enhance user experience and overall performance. This does not diminish the effectiveness of the algorithm but highlights that algorithmic efficiency must be supported by strong system design.

Overall, the contrast between the highest rating in security and the lowest rating in functionality provides a balanced understanding of the system's performance. The findings confirm that the CSAT algorithm is highly effective in generating real-time reports and automatically analyzing data when supported by a secure and reliable system environment. At the same time, continuous improvements in system functionality can further strengthen the efficiency, usability, and overall effectiveness of the automated client satisfaction measurement system.

- *Proposed Technological Solutions to Enhanced the Accuracy and Reliability of CSM Reports Through Automated Data Entry and Validation*

The accuracy and reliability of Client Satisfaction Measurement (CSM) reports in the automated system were evaluated through functional system testing conducted using a structured test plan. The system testing results are summarized in Table 2, which presents the ten test cases executed to verify that the system's core functions operate correctly and consistently. Each test case was designed to assess a specific functional component of the system, including data submission, validation, report generation, security, and processing performance, all of which directly influence the accuracy and dependability of the CSM reports produced.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Table 4 Online CSM System Test Plan Results

Test Case ID	Test Scenario	Expected Result	Status
TC-01	User Access	Successful account creation and access	Successful
TC-02	Login Authentication	Correct login credentials grant access	Successful
TC-03	Feedback Submission	The system records and confirms submitted feedback in real-time.	Successful
TC-04	Sentiment Analysis Processing	The system accurately classifies sentiment as positive, neutral, or negative using BERT-based NLP.	Successful
TC-05	Report Generation	The system generates structured, downloadable reports with real-time data visualization.	Successful
TC-06	Data Aggregation & Accuracy	The system consolidates feedback accurately and ensures consistency in the database.	Successful
TC-07	AI Processing Performance	The system processes and categorizes feedback within predefined response times.	Successful
TC-08	Compliance with ARTA Standards	The system generates reports and feedback insights in compliance with ARTA requirements.	Successful
TC-09	System Security & Access Control	Unauthorized access attempts are blocked, and user roles function correctly.	Successful
TC-10	System Load & Performance	The system handles high user traffic and data processing without performance degradation.	Successful

All ten test cases passed successfully, confirming that the system's core functions, including feedback submission, data aggregation, report generation, and security controls, operate correctly and consistently. These results provide empirical evidence of how the system's technological solutions enhance the accuracy and reliability of CSM reports.

The most direct contributor to report accuracy is the automated data entry mechanism, which was validated through TC-03 (Feedback Submission) and TC-06 (Data Aggregation and Accuracy). TC-03 confirmed that the system records and stores submitted client feedback in real time upon submission, eliminating the need for manual encoding by administrative staff. TC-06 further verified that the system consolidates feedback accurately and ensures consistency across the database, meaning that responses from multiple offices are stored and retrieved in a standardized, error-free format. Under the previous manual process, accuracy was compromised by encoding errors, missing entries, and inconsistent office-level consolidation. The successful execution of both test cases demonstrates that automated entry directly addresses these vulnerabilities by capturing data at the source and storing it without human intervention. This is consistent with findings by Aliazas et al. (2024), who confirmed that automated data capture systems significantly reduce encoding errors and improve data consistency in Philippine institutional settings.

Complementing automated entry, the system's built-in validation mechanisms further strengthen the completeness and dependability of collected data. The validation features, including required fields, input type restrictions, dropdown menus, and real-time error prompts, ensure that clients cannot submit incomplete or invalid responses. The successful result of TC-03 reflects this, as the system was confirmed to enforce submission rules before responses are accepted into the database. This produces cleaner datasets and reduces the need for post-collection data correction, which was one of the persistent limitations of the prior manual process. Rahman and Maryani (2024) established that automated preprocessing and validation processes are essential for ensuring the integrity and quality of feedback data in satisfaction measurement systems.

The accuracy of the reports generated by the system is further supported by TC-05 (Report Generation) and TC-08 (Compliance with ARTA Standards), both of which passed successfully. TC-05 confirmed that the system generates structured, downloadable reports with real-time data visualization, drawing directly from the validated and centralized database. TC-08 verified that these reports are generated in compliance with ARTA reporting requirements, ensuring that the outputs are not only accurate but also formally usable for regulatory submission without additional reformatting. Doria (2025) affirmed that embedding automated reporting tools within feedback systems substantially reduces reporting delays and improves the reliability of institutional outputs, particularly in agencies subject to external reporting mandates.

The security of the system also plays a significant role in sustaining the integrity of CSM data and reports. TC-09 (System Security and Access Control) confirmed that unauthorized access attempts are blocked and that role-based access controls function correctly, ensuring that only authorized personnel can view, edit, or export data. Data integrity cannot be guaranteed if the system is vulnerable to unauthorized access or alteration, making this test case particularly critical to the overall reliability of CSM outputs. Parkavi et al. (2023) demonstrated that privacy-preserving security measures are essential to maintaining the trustworthiness of feedback data in automated systems, while Rahman and Maryani (2024) emphasized that secure environments improve the credibility of machine-learning-generated outputs.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

TC-10 (System Load and Performance) confirmed that the system handles high user traffic and simultaneous data processing without performance degradation, ensuring that accuracy and reliability are maintained even under peak usage conditions. This is further supported by TC-07 (AI Processing Performance), which verified that the BERT-based sentiment analysis component processes and categorizes feedback within predefined response times. Although minor areas for improvement may still exist in system responsiveness and adaptability as usage scales over time, the successful results of both test cases confirm that the system performs reliably under realistic operational conditions. Li et al. (2024) and Bashiri and Naderi (2024) noted that continuous optimization remains important for transformer-based models to sustain classification accuracy across varying datasets.

Overall, the successful completion of all ten test cases confirms that the automated CSM system functions as intended across all critical operations, providing a dependable platform for accurate data entry, consistent validation, and reliable report generation across the DepEd Surigao del Sur Division Office.

- *Classification of Customer Feedback Using BERT and Evaluation of Model Performance*

This study evaluated the capability of the developed BERT-based sentiment classification model in categorizing customer feedback collected through the system. A total of 30 Customer Satisfaction Measurement (CSM) comments submitted by DepEd Surigao del Sur personnel were used as the basis for evaluation. These comments reflect actual sentiments expressed by clients regarding their transactional experiences with the division, including feedback on staff accommodation, document processing, waiting time, system accessibility, and overall service quality.

To establish the ground truth labels for model evaluation, three IT experts independently assigned a sentiment label to each comment, categorizing it as Positive, Neutral, or Negative. A majority voting scheme was then applied, wherein the label agreed upon by at least two of the three experts was adopted as the Final Human Label for each item. Two comments, Items 8 and 25, received conflicting labels from all three raters with no majority agreement and were consequently excluded from the evaluation. This exclusion ensured that only items with a stable and reliable ground truth were used in performance measurement, leaving 28 valid items for analysis, as shown in appendix F.

The BERT model processed each of the 28 valid feedback comments and generated a predicted sentiment label for each. These predictions were then compared against the Final Human Labels to assess how well the model performed across the three sentiment categories. The results of this comparison are presented in the confusion matrix in Table 5.

Table 5 Classification Matrix of BERT Model Predictions

Actual/Predicted	Positive	Neutral	Negative	Total
Positive	12	1	1	14
Neutral	3	5	0	8
Negative	0	3	3	6
Total	15	9	4	28

As shown in Table 5, of the 14 items with an actual Positive label, the BERT model correctly classified 12 as Positive, misclassified 1 as Neutral, and 1 as Negative. Of the 8 items with an actual Neutral label, 5 were correctly classified, while 3 were incorrectly predicted as Positive and none as Negative. Of the 6 items with an actual Negative label, 3 were correctly classified as Negative, while 3 were misclassified as Neutral and none as Positive. The diagonal values of the confusion matrix represent the correctly classified instances for each class, yielding 12, 5, and 3 correct predictions for the Positive, Neutral, and Negative classes, respectively.

The overall accuracy of the BERT model was computed using the sum of correctly classified instances divided by the total number of valid items. The 20 correct predictions out of 28 items yielded an overall accuracy of 71.43%. This means that the model correctly identified the sentiment of approximately 7 out of every 10 customer feedback comments in the evaluation dataset.

To provide a more detailed view of the model's behavior across each sentiment category, per-class performance metrics were computed using the one-vs-rest approach, wherein each class is evaluated as the positive class against all remaining classes. The Accuracy, Precision, Recall, and F1-Score for each class were derived from the True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) counts obtained from the confusion matrix. The complete computations are presented in appendix F, and the summary of results is shown in Table 6.

Table 6 Classification of Task Using Specific Matrix

Class	Accuracy	Precision	Recall	F1-Score
Positive	82.14%	80.00%	85.71%	82.76%
Neutral	75.00%	55.56%	62.50%	58.82%
Negative	85.71%	75.00%	50.00%	60.00%
Overall (Macro-avg)	71.43%	70.19%	66.07%	67.19%

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

- *Note:*

Metrics computed using the one-vs-rest approach. Overall row Note: Overall Accuracy = (12 + 5 + 3)/ 28 = 20 / 28 = 71.43%. Overall row for Precision, Recall, and F1-Score reports macro-averages across all three classes.

The highest classification performance was observed in the Positive class, which obtained a per-class accuracy of 82.14%, with a precision of 80.00%, recall of 85.71%, and F1-score of 82.76%. These balanced values indicate that the model was equally effective at predicting Positive sentiment correctly and at retrieving all actual Positive instances from the dataset, producing no significant bias toward over-prediction or under-detection. The strong and consistent performance in this class suggests that Positive feedback entries in the CSM dataset tend to contain clear and unambiguous linguistic markers such as expressions of gratitude, appreciation, and commendation that the BERT model was able to learn and recognize reliably during fine-tuning. Rahman and Maryani (2024) reported that BERT-based models achieve higher classification accuracy in positive sentiment categories because affirmative language tends to be more syntactically consistent and contextually explicit compared to other sentiment types. Singhal et al. (2023) similarly affirmed that transformer-based models such as BERT are particularly effective in capturing sentiment from contextually clear statements because their bidirectional attention mechanism aligns closely with how meaning is constructed in natural language. The practical implication of this result is that the developed system can reliably identify service strengths and favorable client experiences, which can support evidence-based recognition of well-performing offices and service areas within the division.

The Negative class recorded the highest per-class accuracy at 85.71%, with a precision of 75.00%, a recall of 50.00%, and an F1-score of 60.00%. The high per-class accuracy for this class is largely attributable to the relatively smaller proportion of Negative items in the dataset: correctly classifying the larger pool of non-Negative items contributed heavily to the accuracy figure. While the precision of 75.00% confirms that when the model predicted a Negative label, it was correct most of the time, the recall of 50.00% reveals that three out of six actual Negative comments were instead classified as Neutral. This discrepancy indicates that the model was conservative in assigning the Negative label, likely because several negative comments in the dataset were expressed indirectly, subtly, or in a mixed-language form rather than through explicit dissatisfaction markers. Al-Hawari and Ward (2020) similarly noted that automated sentiment systems may miss negative expressions that are implied rather than directly stated, particularly when complaints are framed politely or contextually. Campit (2023) further observed that negative sentiment classification in Filipino and multilingual feedback is especially susceptible to misclassification because critical sentiments are sometimes softened through cultural or linguistic conventions. The practical implication of this finding is significant for CSM reporting, as under-detection of Negative feedback may limit the developed system's capacity to flag service concerns that require prompt attention and intervention.

The Neutral class recorded the lowest per-class accuracy at 75.00%, with a precision of 55.56%, recall of 62.50%, and F1-score of 58.82%. These values collectively indicate that the model experienced the most difficulty with this sentiment category. The precision of 55.56% reveals that nearly half of the entries predicted as Neutral actually belonged to the Positive or Negative categories, while the recall of 62.50% shows that three out of eight actual Neutral items were misclassified into other categories. Neutral expressions are inherently difficult to classify because they frequently overlap with weakly positive or weakly negative statements, often lack explicit sentiment markers, and may consist of very short or ambiguous entries. Several of the Neutral items in the CSM dataset, including entries such as "none," "no comment," "n/a," and "ok," illustrate the challenge of assigning sentiment to feedback that contains little to no evaluative content. Parkavi et al. (2023) noted that sentiment detection systems consistently struggle with emotionally weak or ambiguous statements even when advanced machine learning techniques are applied, while Bilog (2020) emphasized that multilingual and short-text comments reduce the clarity of sentiment patterns that models rely on for classification. The Neutral category thus requires the most improvement, particularly through expanded and more consistently annotated training data, refined preprocessing strategies for brief or mixed-language entries, and possible clarification of annotation guidelines to improve inter-rater consistency among expert validators.

Across all three classes, the macro-averaged precision, recall, and F1-score were 70.19%, 66.07%, and 67.19%, respectively, alongside an overall accuracy of 71.43%. This level of performance indicates that the model correctly classified approximately seven out of every ten client feedback comments in the evaluation dataset, demonstrating a generally acceptable level of accuracy for an automated sentiment classification task operating on real-world, informal, and multilingual feedback data. The pattern across classes reflects the inherent distribution of sentiment difficulty in CSM data: Positive comments are typically the most explicitly worded, Negative comments vary in their degree of directness, and Neutral comments are the most linguistically ambiguous. Sadhu et al. (2024) and Omas-as et al. (2024) affirmed that advanced natural language processing models operating within this performance range can meaningfully reduce the manual effort required to analyze large volumes of institutional feedback while maintaining a reasonable level of classification reliability.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

The results collectively affirm that the BERT-based model is a useful and reasonably reliable component of the automated CSM system for classifying client feedback into Positive, Neutral, and Negative sentiment categories. However, the lower scores in the Neutral and Negative classifications indicate that continued improvement is necessary. Retraining the model with a larger and more diverse dataset, incorporating more examples of Filipino and regional dialect expressions, and refining the annotation guidelines used by the expert validators are recommended steps for enhancing classification performance over time. This is consistent with the position of Singhal et al. (2023), who emphasized that deep learning models become substantially more effective when trained on context-relevant and domain-specific data, and with Bilog (2020), who stressed the importance of accounting for the multilingual and culturally inflected nature of feedback in Philippine institutional settings.

- *The System Supports ARTA-Aligned Client Satisfaction Reporting*

The compliance of the automated Client Satisfaction Measurement (CSM) system with the standards of the Anti-Red Tape Authority (ARTA) under Republic Act No. 11032, also known as the Ease of Doing Business and Efficient Government Service Delivery Act of 2018, was assessed through the responses of key informants who are direct end users and administrators of CSM reports in the DepEd Surigao del Sur Division Office. Their responses consistently demonstrated that ARTA compliance is not merely a regulatory formality but a substantive governance obligation, and that the system's features directly address the specific requirements of ARTA across five dimensions: the criticality of compliance, the limitations of the current system, the standardization of reporting, the timeliness of submission, and the system features necessary for sustained compliance. To further operationalize this verification, an ARTA Compliance Checklist was developed and is presented in appendix G.

All key informants affirmed that meeting ARTA requirements is a critical obligation for their respective offices. Bethany, the Accountant, stated that "meeting the requirements of the Anti-Red Tape Authority (ARTA) when reporting client satisfaction is highly critical because it ensures compliance with national policies on transparency, efficiency, and accountability in public service." Erlyn, the HR Training Officer, similarly noted that "compliance with ARTA standards demonstrates that the department is committed to monitoring client satisfaction and continuously improving public service." Junrey, the Administrative Officer IV, described ARTA compliance as "very critical because it ensures transparency, accountability, and quality public service" and as a direct reflection of how effectively the department serves the public. Joielyn, the Administrative Officer IV for Records, grounded this further in institutional accountability by explaining that meeting ARTA requirements "is one of our agency's requirements for compliance and eligibility for PBB" and that it supports the continuous improvement of both internal and external services. Eppie, the Budget Officer, offered a resource-based perspective, noting that "Budget Office is very critical in terms of CSM because it entails funds and resources where timelines is considered." Michaella, the Administrative Officer II, likewise emphasized that "the CSM reports serve as one of the indicators of how well we deliver our services to the public, so it is important that these reports are accurate and submitted properly according to the guidelines set by ARTA." Marjun, the Planning Development Officer II, reinforced this by stating that "the law requires government offices to measure client satisfaction and improve service delivery" and that "following these standards also shows that our office is transparent and accountable to the public." These converging responses establish that all offices recognize ARTA-aligned CSM reporting as a governance requirement tied to public accountability, institutional performance, and regulatory eligibility.

Despite this recognition, the key informants identified several challenges in meeting ARTA standards under the current system. Erlyn explained that "one of the main challenges is the time required to manually collect, verify, and consolidate data from paper-based feedback forms" and that "the possibility of incomplete responses or misplaced forms may affect the completeness of the data," making it difficult to ensure consistent and timely report preparation. Marjun echoed this, stating that "the slow process of collecting and encoding the survey responses" is a challenge because "it takes time to organize the data and prepare the reports," and that "incomplete responses also make it difficult to prepare accurate reports." Bethany identified "manual data collection, time-consuming data encoding, and inconsistent participation from the clients" as recurring challenges that affect the efficiency, accuracy, and timeliness of ARTA-required reporting. Jovixson, the Planning Officer, was more direct, noting that "reports are manually encoded by the in-charge" and that "reports vary from office to office," which makes consolidation difficult. Michaella added a practical dimension, noting that "occasional technical issues such as slow system response, connectivity problems, or clients needing guidance on how to use the kiosk properly" can still produce incomplete data even in a partially digital environment. Eppie further noted that internet connectivity is a common problem and that when the internet is unavailable, the system goes offline without a backup system. These responses collectively confirm that ARTA compliance under the current setup is hindered by manual processes, fragmented consolidation, and unreliable data completeness, and that these are precisely the vulnerabilities the automated system is designed to address.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

The key informants were consistent in affirming that the automated CSM system can resolve these gaps by standardizing data collection and reporting in a manner aligned with ARTA requirements. Junrey stated that "an automated CSM system can standardize data collection and reporting formats" and that it "can automatically compute satisfaction ratings and generate reports aligned with ARTA requirements, ensuring compliance and accuracy." Erlyn explained that the system "can help ensure compliance with ARTA standards by automatically recording client responses, storing them in a centralized database, and generating reports based on standardized formats," thereby reducing errors and ensuring that data are consistently organized. Michaella further affirmed that "the system can be designed to follow the required format and indicators set by ARTA, making sure that the necessary information is captured from the start," which reduces the risk of missing data and makes the reporting process more consistent with ARTA guidelines. Joielyn stressed that "the automated CSM system should follow the required data needed for submission to ARTA and ensures that the approved enhanced CSM tool should be integrated and captured in the system." Jovixson summarized the compliance benefit of automation clearly: "with the automated CSM system, reports to be generated can be set to follow the ARTA standards. This way, we can ensure that whatever results are generated, it automatically follows the standards." These responses demonstrate that the system embeds ARTA compliance requirements directly into its data collection and reporting architecture, removing the inconsistency and variation that previously made compliance difficult to verify. This is supported by Aliazas et al. (2024), who found that automated systems embedded with standardized reporting protocols substantially improve consistency and reduce discrepancies across organizational units in Philippine government agencies.

The key informants also affirmed that automatic report generation is a decisive feature for ensuring the timeliness of ARTA-required submission. Michaella stated that "automatic report generation can greatly help in the timely submission of required reports" because "the system will already compile and summarize the responses, and staff will no longer need to manually compute or organize the data." Marjun similarly noted that "the system can quickly summarize the survey results and produce reports immediately," which "will help the office submit reports on time and avoid delays." Erlyn confirmed that "personnel will no longer need to manually compute or consolidate responses," and that "this will shorten the reporting process and allow the department to submit reports within the required deadlines." Junrey stated that "instead of manually compiling the data, the system can instantly generate the necessary reports once the responses are collected." Joielyn described the practical benefit most directly: "by simply generating a consolidated report in the system, you don't need to convey or ask each office to have their consolidated CSM result because the system itself can do it." Jovixson confirmed that "since the process is automated, creating reports can now be done swiftly," while Bethany added that automatic generation "streamlines the reporting process by reducing manual work and ensuring that data is processed quickly and accurately." These responses demonstrate that the automation of report generation eliminates the most significant source of reporting delay and positions offices to consistently meet ARTA submission deadlines. Alberto (2025) supports this finding, reporting that automated administrative systems reduce processing time and enhance workflow efficiency in government offices, thereby enabling timely compliance with regulatory reporting requirements.

The key informants also identified specific system features that would sustain ARTA compliance over time. Erlyn recommended that the system include "features such as automatic data recording, real-time data consolidation, standardized reporting templates, and secure data storage," together with "user-friendly interfaces for clients, data validation mechanisms to prevent incomplete responses, and monitoring dashboards that allow administrators to track satisfaction levels and generate reports easily." Michaella similarly stated that the system "should include automatic validation of responses to prevent incomplete submissions, a standardized reporting format based on ARTA guidelines, and automatic data consolidation for easier analysis," as well as "secure data storage and the ability to generate reports quickly." Bethany specified that the system should have "an ARTA-compliant survey questionnaire, automated data collection, and automatic report generation" as its core compliance-enabling features. Marjun added that the system "should have automatic data recording, validation of responses, and automatic report generation," as well as safe data storage and easy access to reports. Joielyn emphasized the importance of continuous monitoring by recommending that "each office should be notified of their weekly rating, so that they will be aware," underscoring the value of real-time office-level accountability rather than waiting until the end of the reporting cycle to review results. Jovixson introduced the dimension of inclusivity, stating that "to promote inclusivity, the system should include features that can assist persons that are differently abled, i.e., blind, deaf, etc.," which reflects an ARTA-aligned principle of universal accessibility in public service delivery. Eppie noted that "the automation recording of the CSM system results will make sure the departments remain compliant with ARTA" and further recommended that clients be required to respond to CSM questions before the issuance of a Certificate of Appearance, to achieve a higher and more reliable response rate. These responses affirm that the system's built-in validation, monitoring, accessibility, and security features collectively constitute the infrastructure for sustained ARTA compliance. This is consistent with the position of Cervantes et al. (2024), who found that digital adoption in DepEd-related settings improves when systems are both useful and easy to navigate, and with Parkavi et al. (2023), who emphasized that secure and privacy-conscious data handling is an essential condition for trustworthy institutional reporting.

Across all five dimensions, the key informants' responses provide direct evidence that the automated CSM system is designed and positioned to ensure ARTA-aligned client satisfaction reporting. The system's capacity to standardize survey indicators, consolidate responses automatically, generate reports in prescribed formats, and submit them within required timeframes collectively addresses the compliance gaps identified in the current manual process. To formally document and verify the extent of this compliance, an ARTA Compliance Verification Checklist is provided in appendix G, which may be used by offices during system review, pre-submission audits, and monitoring visits.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

CHAPTER FIVE
SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

This chapter presents the summary of findings, conclusions, and recommendations of the study on the automated analysis of data aggregation and reporting of Client Satisfaction Measurement (CSM) in DepEd Surigao del Sur. This chapter synthesizes the key results derived from both quantitative and qualitative analyses, highlighting the effectiveness of the proposed system in improving efficiency, accuracy, and compliance with standards. It also provides conclusions drawn from the findings and offers practical recommendations aimed at enhancing the implementation of automated client satisfaction systems and guiding future research and development in this field.

- *Summary of Findings*

The findings revealed that the technological solutions utilized in the study effectively streamlined the processes of data collection, storage, and processing. Based on the thematic analysis of key informant responses, several recurring themes emerged, including the adoption of web-based systems, centralized data management, and automated survey tools, all of which facilitated faster data consolidation and reduced reliance on manual processes.

Subsequently, the results showed that the algorithms used for real-time reporting and analysis performed at a consistently high level across all indicators, with security (4.77) and maintainability (4.75**)** obtaining the highest mean scores, followed by efficiency (4.60)**,** usability (4.35)**,** reliability (4.33)**,** and functionality (4.29) as the lowest. These findings suggest that while the system performed well overall in generating automated reports, certain aspects such as functionality may still require further refinement.

Moreover, the findings from the thematic analysis indicated that the use of automated data entry and validation mechanisms improved the accuracy and reliability of reports. Key themes identified include error reduction, validation controls, and system-generated outputs**,** which contributed to more consistent and dependable data reporting.

In addition, the results revealed that the BERT-based sentiment analysis model demonstrated a high level of effectiveness in classifying client feedback into appropriate sentiment categories. The performance of the classification task yielded high values in Accuracy, Precision, Recall, and F1-Score with overall Accuracy of 78.57%, indicating that the model can support automated classification, however, the performance level suggests that further optimization may be necessary, particularly in handling more nuanced or less represented sentiment classes

Finally, the findings based on thematic analysis showed that the system ensured compliance with ARTA standards through key themes such as transparency, timeliness, accessibility, and standardization of reports. These themes reflect how the system supported the required processes for client satisfaction reporting in alignment with government standards. Thus, the results indicate potential support for compliance, subject to further validation.

- *Conclusions*

The findings of the study lead to the conclusion that the integration of web-based and automated technological solutions plays a critical role in enhancing the efficiency of data aggregation and reporting systems. The emergence of themes such as centralized data management and automated survey tools indicates that transitioning from manual to digital processes results in a more streamlined and responsive system, enabling faster and more organized handling of client satisfaction data.

Furthermore, the consistently high performance of the algorithms across all system quality indicators suggests that the implemented system is robust, reliable, and capable of sustaining real-time data processing demands. The higher ratings in security and maintainability imply that the system is not only dependable in protecting data but also adaptable for long-term use, while the relatively lower rating in functionality highlights areas that may still require refinement to optimize system features.

In addition, the thematic findings suggest that the use of automated data entry and validation mechanisms is associated with improved consistency in reporting. Themes such as error reduction and validation controls indicate that automation may help minimize inconsistencies in data handling. However, these observations are drawn from participant perceptions rather than controlled comparative measures.

With regard to sentiment analysis, the BERT-based model demonstrated acceptable performance in classifying client feedback into sentiment categories. The obtained accuracy of 78.57%, along with the corresponding precision, recall, and F1-score values, suggests that the model can support automated classification of textual feedback. Nevertheless, the level of performance indicates that further refinement may be needed, particularly in handling complex or less represented sentiment classes.

Finally, the identified themes related to transparency, timeliness, accessibility, and standardization indicate that the system aligns with ARTA standards for client satisfaction reporting. This suggests that the automated system not only improves internal processes but also supports compliance with government requirements, contributing to more accountable and efficient public service delivery.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

- *Recommendations*

Based on the conclusions drawn and identified limitations of the study, several technical and operational recommendations are proposed.

For the DepEd Surigao del Sur Division, it is recommended that the implementation of the web-based Client Satisfaction Measurement (CSM) system be carried out in a phased manner across selected offices before full-scale deployment. Pilot testing in multiple functional units may allow further validation of system performance under varying transaction volumes and service contexts. Continuous monitoring of system outputs, including report accuracy, response time, and user engagement, is also recommended to assess long-term operational performance.

In relation to the sentiment analysis component, it is recommended that the dataset be expanded to include a larger volume of client feedback, particularly those reflecting neutral and negative sentiments. The inclusion of more localized language data, such as Filipino and regional dialect expressions, may improve the model's ability to capture context-specific meanings. Periodic retraining and fine-tuning of the BERT model using updated datasets are also advised to enhance classification performance and address observed limitations in certain sentiment categories.

For system development and maintenance, it is recommended that IT practitioners strengthen data security and system reliability mechanisms. This may include the implementation of role-based access control, encrypted data storage, and regular system backup protocols to prevent data loss. In addition, improvements in system functionality and user interface design should be prioritized, particularly in areas identified with relatively lower user ratings, to enhance usability and overall user experience.

For administrators and decision-makers, it is recommended that the outputs of the system, including sentiment analysis results and automated reports, be used as supplementary decision-support tools rather than sole bases for policy actions. Further validation of trends through additional data sources or manual review processes may help ensure more accurate interpretation of client feedback.

For future research, it is recommended to conduct further evaluation of the system using larger and more diverse samples, as well as to explore comparative analysis with other machine learning models to determine optimal approaches for sentiment classification. Longitudinal studies may also be undertaken to examine the sustained effectiveness, adaptability, and scalability of the system over time in different institutional settings.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

REFERENCES

[1]. Al-Hawari, M., & Ward, T. (2020). The effect of automated service quality on customer satisfaction and financial performance: An artificial neural network analysis. Journal of Computer Information Systems, 60(4), 389-397.

[2]. Aliazas, J. V., Pasia, A., & Madrideo, J. (2024, September). Teaching Effectiveness: A Design Feedback Process of One State University in the Philippines. ResearchGate.

[3]. Alqurafi, A., & Alsanoosy, T. (2024). Measuring customers' satisfaction using sentiment Analysis: Model and tool. *Journal of Computer Science*, *20*(4), 419–430. https://doi.org/10.3844/jcssp.2024.419.430

[4]. Alzaydi, Z. M., Al-Hajla, A., Nguyen, B., & Jayawardhena, C. (2018). A review of service quality and service delivery: Towards a customer co-production and customer-integration approach. Business Process Management Journal, 24(1), 295-328.

[5]. Arli, D., Van Esch, P., & Weaven, S. (2024). The impact of SERVQUAL on consumers' satisfaction, loyalty, and intention to use online food delivery services. *Journal of Promotion Management*, *30*(7), 1159–1188. https://doi.org/10.1080/10496491.2024.2372858

[6]. Bagozzi, R. P. (2007). The legacy of the technology acceptance model and a proposal for a paradigm shift. Journal of the Association for Information Systems, 8(4), 244-254.

[7]. Bashiri, H., & Naderi, H. (2024). Comprehensive review and comparative analysis of transformer models in sentiment analysis. *Knowledge and Information Systems*, *66*(12), 7305–7361. https://doi.org/10.1007/s10115-024-02214-3

[8]. Cervantes, J. S., Armingol, M. B., Abrasaldo, K. G., Amit, D. P. & Alejandrino, M. D. (2024). The Relationship of Utilization of Digital Payment and Customer Satisfaction among DepEd Employees in Digos City. Journal of Interdisciplinary Perspectives, 2(7), 385-391.

[9]. Chuttur, M. Y. (2009). Overview of the technology acceptance model: Origins, developments and future directions. Working Papers on Information Systems, 9(37), 9-37.

[10]. Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

[11]. Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

[12]. De Goma, J., Mercado, R. M., Saporna, T. J., & Vallo, S. P. (2024). Sentiment Analysis of Tweets Towards Vaccines via Bidirectional Encoder Representations from Transformers (BERT). *ICCAI*, 419–424. https://doi.org/10.1145/3669754.3669819

[13]. Doria, J. (2025). A Data-Driven Customer Satisfaction Survey Model Integrating Sentiment and Prescriptive Analytics in Higher Educations Institutions. *Southeast Asian Journal of Science and Technology*, *10*(1), 137-145.

[14]. Etikan, I. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1. https://doi.org/10.11648/j.ajtas.20160501.11

[15]. Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, *35*(3), 379–383. https://doi.org/10.3758/BF03195514

[16]. Flood, R. L. (2010). The relationship of 'systems thinking' to action research. Systemic Practice and Action Research, 23(4), 269-284.

[17]. Gabia, R. M., Urbiztondo, F. O. (2025). Streamlining Public Service In Higher Education: A Study On The Anti-Red Tape Act Implementation In Caraga State University. In *IJARIIE* (pp. 3046–3047) [Journal-article]. https://ijariie.com/AdminUploadPdf/Streamlining_Public_Service_In_Higher_Education__A_Study_On_The_Anti_Red_Tape_Act_Implementation_In_Caraga_State_University_ijariie26842.pdf?srsltid=AfmBOooPc3tCD47irvbuLBinX6PKIQQw2fbqinSxCuGopaiH3xvBIYOS

[18]. Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

[19]. Hien, B. N., Tuyen, N. T. K., Lan, N. T., Ngan, N. T. K., & Thanh, N. N. (2024). The Impact of Digital Government Initiatives on Public Value Creation: Evidence from Ho Chi Minh City –Vietnam. Revista de Gestão Social e Ambiental, 18(2), e04892–e04892.

[20]. Indar, I. N., Razi, D. F., Atmansyah, L., Yani, A. A., & Ismail, N. (2024). Digital Transformation of Public Services: An Analysis of the Effectiveness of Online Services at the Local Government Level. Journal of Digital Sociohumanities, 1(2), 92–101.

[21]. Jackson, M. C. (2003). Systems thinking: Creative holism for managers. Wiley.

[22]. Laureta, M., Feria, W. A., Limbag, P. C., & Montecillo, B. (2025). Profanity and sentiment detection in Filipino social media comments using Transformer-Based NLP models. *Isabela State University Linker Journal of Education Social Sciences and Allied Health*, 102. https://doi.org/10.65141/jessah.v2i1.n8

[23]. Li, Z., Yang, C., & Huang, C. (2023). A Comparative Sentiment Analysis of Airline Customer Reviews Using Bidirectional Encoder Representations from Transformers (BERT) and Its Variants. *Mathematics*, *12*(1), 53. https://doi.org/10.3390/math12010053

[24]. Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

[25]. Manderscheid, E., & Lee, M. (2024). Predicting Customer Satisfaction by Replicating the Survey Response Distribution. https://doi.org/10.48550/arxiv.2411.12539
[26]. Muñoz, A. V. (2024). *AI-Powered Governance: Navigating the Future of Public Policy with Intelligent Systems*. https://doi.org/10.13140/rg.2.2.18678.41281
[27]. Mozumder, M. a. S., Nguyen, T. N., Devi, S., Arif, M., Ahmed, M. P., Ahmed, E., Bhuiyan, M., Rahman, M. H., Mamun, A. A., & Uddin, A. (2024). Enhancing customer satisfaction analysis using advanced machine learning techniques in fintech industry. *Journal of Computer Science and Technology Studies*, *6*(3), 35–41. https://doi.org/10.32996/jcsts.2024.6.3.4
[28]. Navo, N. (2024, September 26). Empowering Philippine MSMEs Through Digital Transformation. Institute for Small-Scale Industries.
[29]. Oh, H. (1999). Service quality, customer satisfaction, and customer value: A holistic perspective. International Journal of Hospitality Management, 18(1), 67-82.
[30]. Omas-As, R. M. A., & Encarnacion, R. E. (2024). Stakeholders' Satisfaction on Institutional Assessment: A Proposal for Unified Feedback Management System with Text Analytics and Sentiment Analysis. *International Journal of Advanced Research in Science Communication and Technology*, 133–144. https://doi.org/10.48175/ijarsct-18716
[31]. Parkavi, A., Shetty, T. B. N., Shibu, S. G., Ismail, R., & Muthirakalayil, R. A. (2023). Customer Feedback Analysis Based on Emotion Detection Using Machine Learning Techniques with Privacy Preservation. 1617–1624.
[32]. Patton, M. Q. (2015). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications.
[33]. Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147
[34]. Rahman, B., & Maryani. (2024). Optimizing customer satisfaction through sentiment analysis: a BERT-Based machine learning approach to extract insights. *IEEE Access*, *12*, 151476–151489. https://doi.org/10.1109/access.2024.3478835
[35]. Richey, R. C., & Klein, J. D. (2014). *Design and development research: Methods, strategies, and issues* (2nd ed.). Routledge.
[36]. Rivero, D. M., Suson, R., Arnejo, A., Atibing, N. M., Aro, J. L., Wenceslao, C., Burdeos, A., Yamagishi, K., & Ocampo, L. (2023). Service recovery and innovation on customer satisfaction amidst massive typhoon-induced disruptions: The mediating role of SERVQUAL. *International Journal of Disaster Risk Reduction*, *99*, 104130. https://doi.org/10.1016/j.ijdrr.2023.104130
[37]. Sadhu, A. K. R., Parfenov, M., Saripov, D., Muravev, M., & Sadhu, A. K. R. (2024). Enhancing Customer Service Automation and User Satisfaction: An Exploration of AI-powered Chatbot Implementation within Customer Relationship Management Systems. Journal of Computational Intelligence and Robotics, 4(1), 103–123.
[38]. Singhal, P., Chaudhary, B., & Singh, V. (2023). Improving feedback analysis: Deep learning approach to college customer satisfaction assessments. Multidisciplinary Science Journal.
[39]. Sy, C. Y., Canon, M. J. P., Maceda, L. L., Flores, N. M., Palaoag, T. D., & Abisado, M. B. (2024). Leveraging transformer-based BERTopic model on stakeholder insights towards Philippine UAQTE. *International Journal of Engineering Trends and Technology*, *72*(3), 277–287. https://doi.org/10.14445/22315381/ijett-v72i3p125
[40]. Tongco, M. D. C. (2007). Purposive sampling as a tool for informant selection. *Ethnobotany Research and Applications*, *5*, 147–158. https://doi.org/10.17348/era.5.0.147-158
[41]. Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors*, *34*(4), 457–468. https://doi.org/10.1177/001872089203400407
[42]. Von Bertalanffy, L. (1972). The history and status of general systems theory. Academy of Management Journal, 15(4), 407-426.
[43]. Yüksel, A., & Yüksel, F. (2001). The expectancy-disconfirmation paradigm: A critique. Journal of Hospitality & Tourism Research, 25(2), 107-131.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDICES

APPENDIX A:
SURVEY QUESTIONNAIRE

Dear Respondents,

Please complete this survey questionnaire form. Rest assured that the information will be kept confidential. Thank you.

MARVIN G. MINGUILLAN

Researcher

I. PERSONAL DETAILS

Name: _____

II. EDUCATIONAL ATTAINMENT (Please check the one that applies)

_____ Elementary Graduate

_____ High School Graduate

_____ Vocational Graduate

_____ College Level

_____ College Graduate

_____ Masteral

_____ Doctorate

III. OCCUPATION AND SKILLS (Please check one that applies)

_____ House owner

_____ Business owner

_____ Other. Please Specify _____

_____ Government Employee

_____ IT Expert

Do you know how to use a mobile phone? _____ Yes _____ No.

Evaluate the developed ***Hose Leak Detection System using Sensor Technology*** using the ISO 25010 software standard.

5 Very Great Extent
4 Great Extent
3 Moderate Extent
2 Little Extent
1 Very Little Extent

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

A. Functionality	5	4	3	2	1	Comments
1. The required task is performed by the system satisfactorily						
2. The output generated by the system is correct.						
3. The system helps the user overcome any problem encountered.						
4. The system regulates user access to data.						
B. Reliability	**5**	**4**	**3**	**2**	**1**	**Comments**
1. Most errors have been corrected over time.						
2. The system is capable of maintaining its performance despite the presence of errors during operation.						
3. The system restores the operation of lost data after failure.						
4. The system provides real-time monitoring to ensure availability.						
C. Usability	**5**	**4**	**3**	**2**	**1**	**Comments**
1. The system is intuitive and receptive to learning						
2. The information generated by the system is easy to understand and act upon.						
3. The documentation is very informative.						
4. The system straightforwardly performs tasks.						
D. Efficiency	**5**	**4**	**3**	**2**	**1**	**Comments**
1. The system responds immediately to sensor inputs.						
2. The system maximizes the efficient utilization of available resources (i.e., people, data, machines, and materials).						
3. The system performs efficiently on intended tasks with ease and scalability.						
4. The system meets established efficiency benchmarks for industry standards						
E. Maintainability	**5**	**4**	**3**	**2**	**1**	**Comments**
1. Deficiencies can be easily diagnosed and modified from the system.						
2. The system can be easily customized and modified.						
3. The system continues functioning if changes are made.						
4. The system can easily be tested.						
5. Deficiencies can be easily diagnosed and modified from the system.						
6. The system can be easily customized and modified.						
7. The system continues functioning if changes are made.						
8. The system can easily be tested.						

Source: ISO 25010:2011 Software Quality Standard

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX B:
SEMI-STRUCTURED INTERVIEW GUIDE

KEY INFORMANTS: DEPARTMENT HEADS (END USERS OF CSM RESULTS)

Introductory Script (To Be Read by Interviewer)

Good day. This interview is part of a study about improving the Customer Satisfaction Measurement (CSM) system.
We would like to ask your opinion on how the proposed system, especially the online kiosk and automatic recording of responses, can improve the accuracy and reliability of CSM reports.
We will also ask how the system can help ensure compliance with ARTA standards for client satisfaction reporting.
There are no right or wrong answers. Your honest opinion is important.

PART I SYSTEM ENSURANCE COMPLIANCE WITH ARTA STANDARDS FOR CLIENT SATISFACTION REPORTING

(Reference: Anti-Red Tape Act of 2007 and its amendments under Ease of Doing Business and Efficient Government Service Delivery Act of 2018)

Main Questions

1. How critical it is for your department to meet ARTA requirements when reporting satisfaction?
2. What challenges does your department face in meeting ARTA standards with the current system?
3. How can an automated CSM system help ensure that reports follow ARTA standards?
4. Do you think automatic report generation will help in timely submission of required reports? Why?
5. What features should the system have to make sure your department

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX C:
SENTIMENT ANALYSIS VALIDATION TOOL (FOR VALIDATORS)

System/Module: Client Satisfaction Measurement (CSM) – Sentiment Analysis & Automated Reporting

Validator Name: _____

Position: _____

Agency/ Office/Unit: _____

Date: _____

Purpose: This tool is used to validate the accuracy, reliability, and usefulness of the system's sentiment analysis results (Positive/Neutral/Negative) and how these results support CSM reporting for decision-making.

PART I. SENTIMENT LABELING VALIDATION (Ground-Truth Check)

Instruction: Below is a sample set of actual CSM open-ended comments. For each item, (a) assign your "human label" (Positive/Neutral/Negative) based on your judgment, and (b) check whether the system label matches your judgment. Indicate the remarks if the label is true positive, true negative, false positive, and false negative.

1.1.1 **True Positive (TP) -** A comment is labeled **Positive** by the human validator, and the system also correctly classifies it as **Positive**. This means the model correctly identifies a positive sentiment.

1.1.2 **True Negative (TN)** - A comment is labeled **Negative** by the human validator, and the system also correctly classifies it as **Negative**. This means the model correctly identifies a negative sentiment.

1.1.3 **False Positive (FP) -** A comment is labeled **Negative (or Neutral)** by the human validator, but the system incorrectly classifies it as **Positive**. This means the model mistakenly predicts a positive sentiment when it is not positive.

1.1.4 **False Negative (FN)** - A comment is labeled **Positive** by the human validator, but the system incorrectly classifies it as **Negative (or Neutral)**. This means the model fails to detect a positive sentiment.

No.	CSM Comment (verbatim)	Human Label (Positive / Neutral / Negative)	System Label (Positive / Neutral / Negative)	Match? (Yes/No)	Remarks (true positive, true negative, false positive, false negative
1	Basta deped Surigao ♡				
2	the security personnel should improve their attitude				
3	Security personnel is approachable				
4	none				
5	Dugay kayo ug wait				
6	Dako ang tabang ninyo sa ako				
7	nknek				
8	Online processing must be available				
9	help				
10	Dili ka gwapa				
11	Nered onloineer transaction				
12	keep ukp the good wdsork				
13	ok				
14	Accommodating accommodating				
15	kapoy				

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

16	Swabe ra kaayo ang transaction				
17	Appreciated kaayo ang effort				
18	Complicated ang forms				
19	libog				
20	Naa puy guide. Helpful				
21	Need more information kiosk				
22	no comment				
23	Deserve nyo ng award sa service				
24	n/a				
25	Kulang ang staff				
26	satisfy				
27	Air conditioned pa. Comfy jud				
28	Dapat may aircon				
29	GOODS KAU				
30	Medjo init sa waiting area				

Validator Signature: _____

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX D:
LETTER TO CONDUCT A STUDY

Permission to Conduct a Study

DR. LORENZO O. MACASOCOL, CESO V
Schools Division Superintendent
Department of Education
Division of Surigao del Sur
Tandag City, Surigao del Sur

Dear Dr. Lorenzo

Greetings of peace and goodwill.

I am currently pursuing a Master's Degree in Computer Science and conducting a study entitled *"Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement Using BERT Model and Sentiment Analysis Algorithm."*

In this regard, I respectfully request permission to conduct the study in the Schools Division of Surigao del Sur. The study will involve selected teaching and non-teaching personnel who have utilized the Client Satisfaction Measurement (CSM) system, as well as key informants from various departments. All data gathered will be used solely for academic purposes and handled with strict confidentiality.

May I humbly request your approval for the conduct of this study. Your support will greatly contribute to its successful completion and to the improvement of client satisfaction reporting in the division.

Thank you very much for your consideration.

Respectfully yours,

MARVIN G. MINGUILLAN
Researcher

Approved by:

LORENZO O. MACASOCOL PhD, CESO V
Schools Division Superintendent

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX E:
CLIENT SATISFACTION SURVEY APPLICATION

➢ *User Side*

1. The user will login to ehris.surigoadelsur.com, the click the button “click here”.

Fig 11 Welcome Page of the System

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

2. The user will start answering the survey form. By clicking the “Start Survey” then check the check box button agreeing to the data privacy notice and click the “I accept and agree" button.

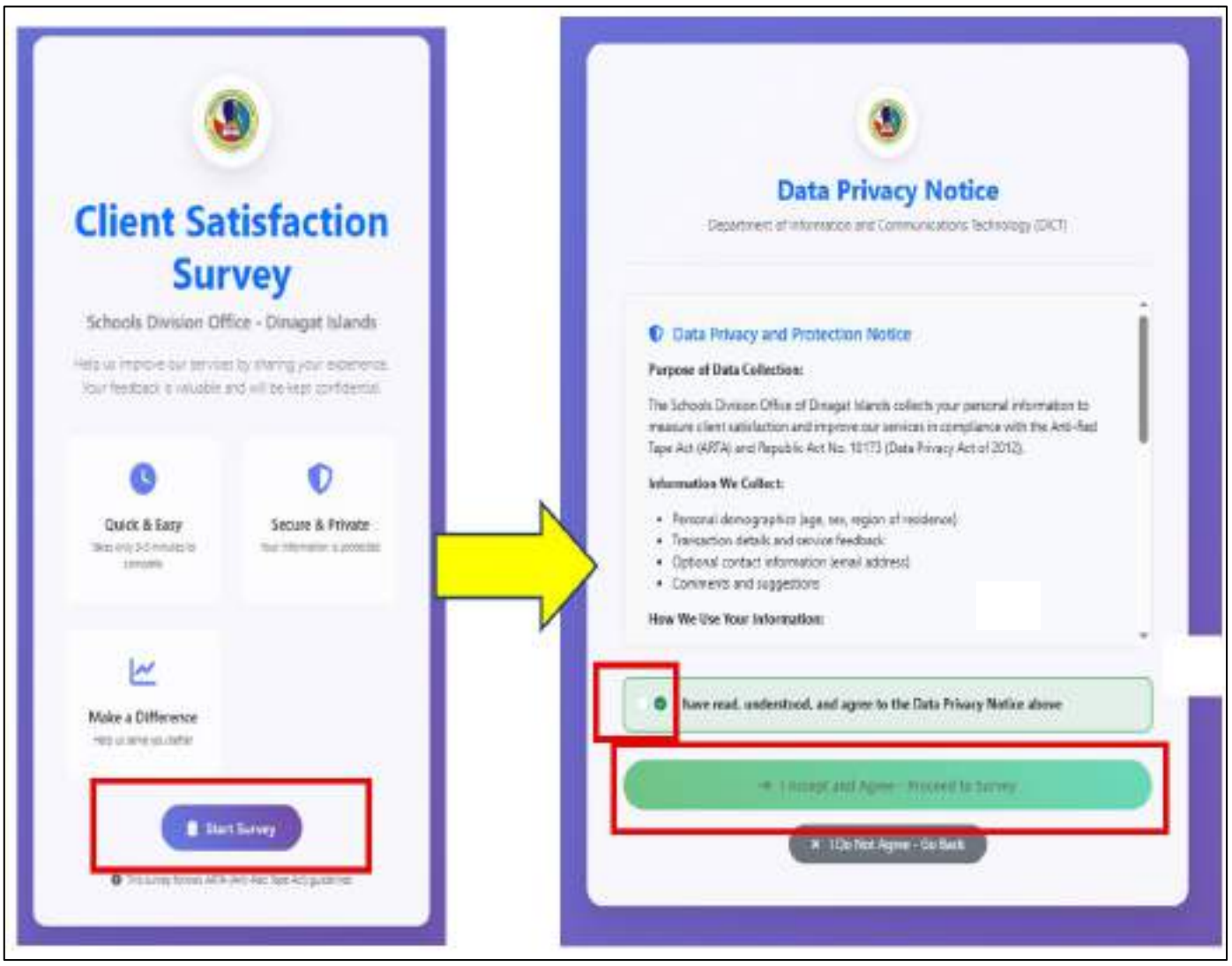


Fig 12 Welcome Page and Data Privacy Disclaimer

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

3. The user will now start answering the Client Satisfaction Survey.

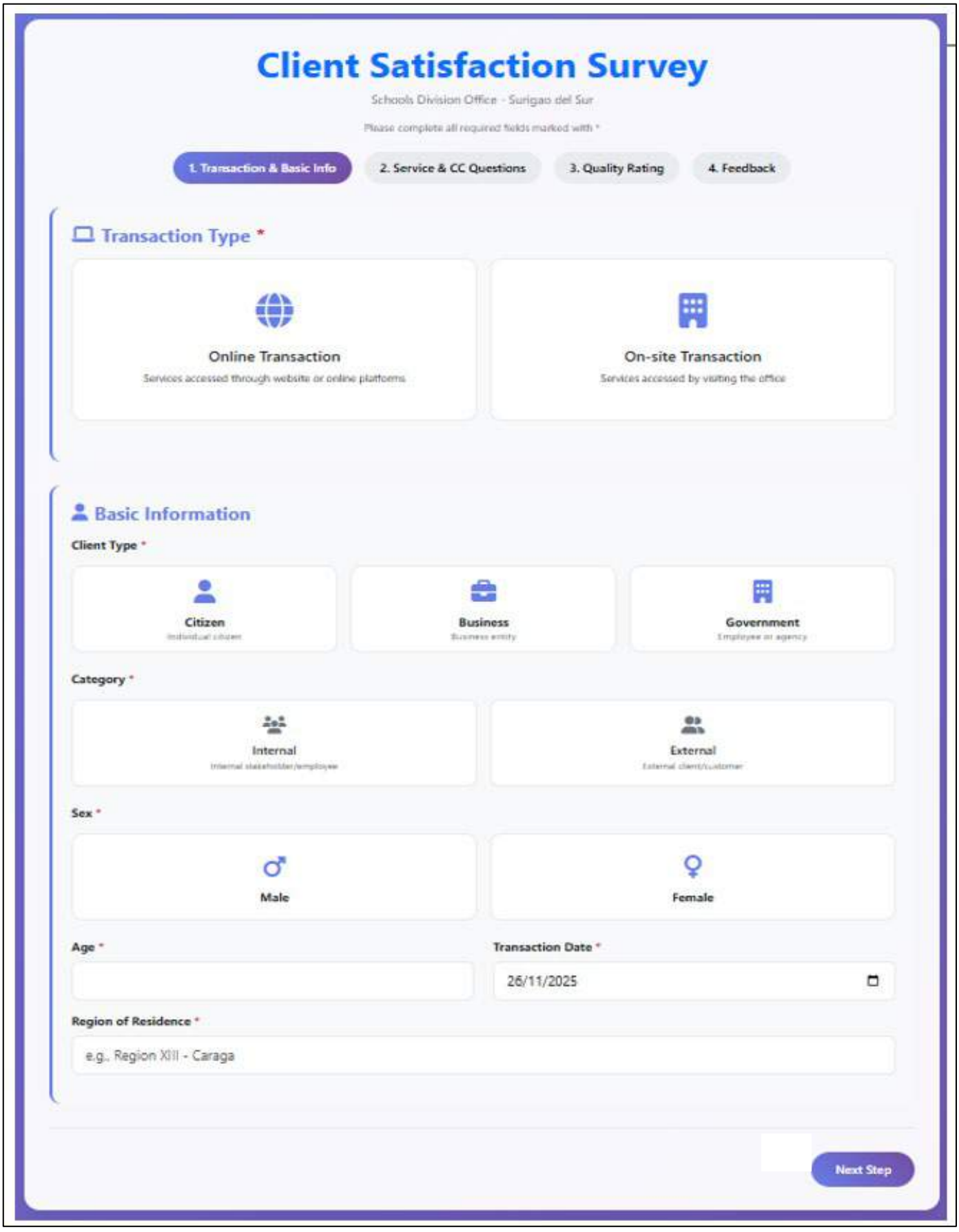


Fig 13 Client Satisfaction Basic Info

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Transaction Type (Required)*

This section asks how you availed the service.

a. Online Transaction

Select this if you accessed the service through:

- Website
- Online platform
- Email
- Messenger or any online communication

b. On-site Transaction

➢ *Basic Information Section*

This section collects basic details to categorize your survey submission. Every item marked with * is required.

I. Client Type (Required)
Who you are as a client.

- Citizen – Individual client
- Business – Private company or business entity
- Government – Employee or representative of another government agency

II. Category (Required).
What type of stakeholder you are.

- Internal – DepEd employee or internal stakeholder
- External – Students, parents, businesses, or general public clients

III. Sex (Required)
Choose your biological sex.

- Male – Click the Male icon
- Female – Click the Female icon

IV. Age (Required)
Enter your age in numbers using the text field.

V. Transaction Date (Required)
This is the date you availed the service.

How to select:

- Click the date field
- Choose the correct date from the calendar popup
- Date will automatically appear (e.g., 26/11/2025)

VI. Region of Residence (Required)
Enter the region where you live (e.g., Region XIII – Caraga).

Proceeding to the Next Step. After filling in all required fields: Click “Next Step” (bottom-right button). This will take you to Step 2: Service & Citizen’s Charter Questions.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

4. The user will answer the Citizen's Charter Questions and select what department and service they avail.

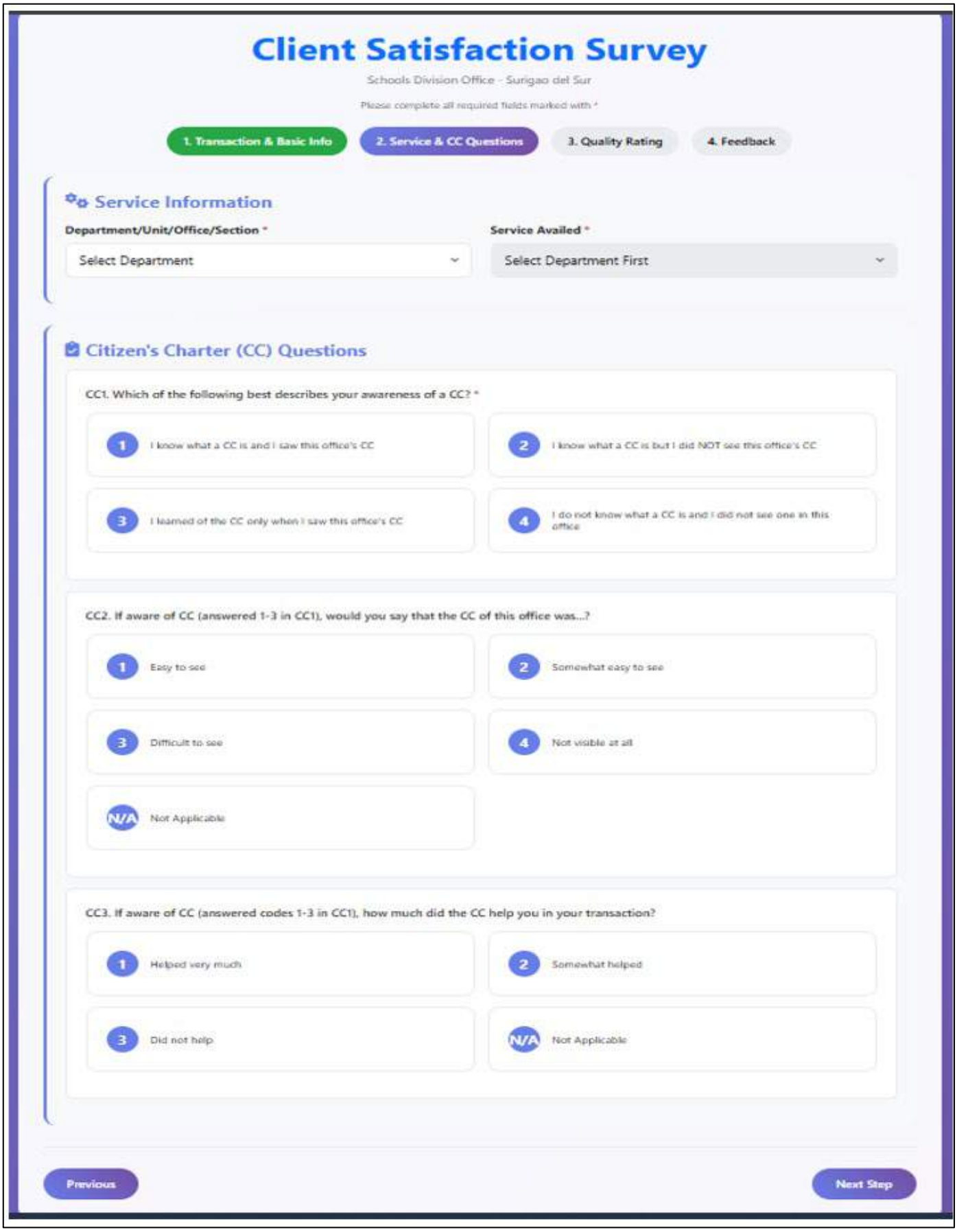


Fig 14 Citizen' Charter Questions

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Client Satisfaction Survey – Service & Citizen's Charter (Cc) Section*

This section gathers details about the client's transaction and their awareness of the office's Citizen's Charter (CC).

The user must first select the **Department/Unit/Office/Section** they transacted with and choose the **Service Availed** from the available options.

After selecting these, the user answers brief multiple-choice questions about:

- **CC Awareness** – Whether they know what a Citizen's Charter is and if they saw it in the office.
- **CC Visibility** – How easy or difficult it was to see the Citizen's Charter.
- **CC Helpfulness** – How useful the Citizen's Charter was in completing their transaction.

This information helps the office assess how well the Citizen's Charter is displayed, understood, and utilized by clients.

5. The next step, they will be answering the Service Quality Dimensions. The user will rate now the service quality offered by the office.

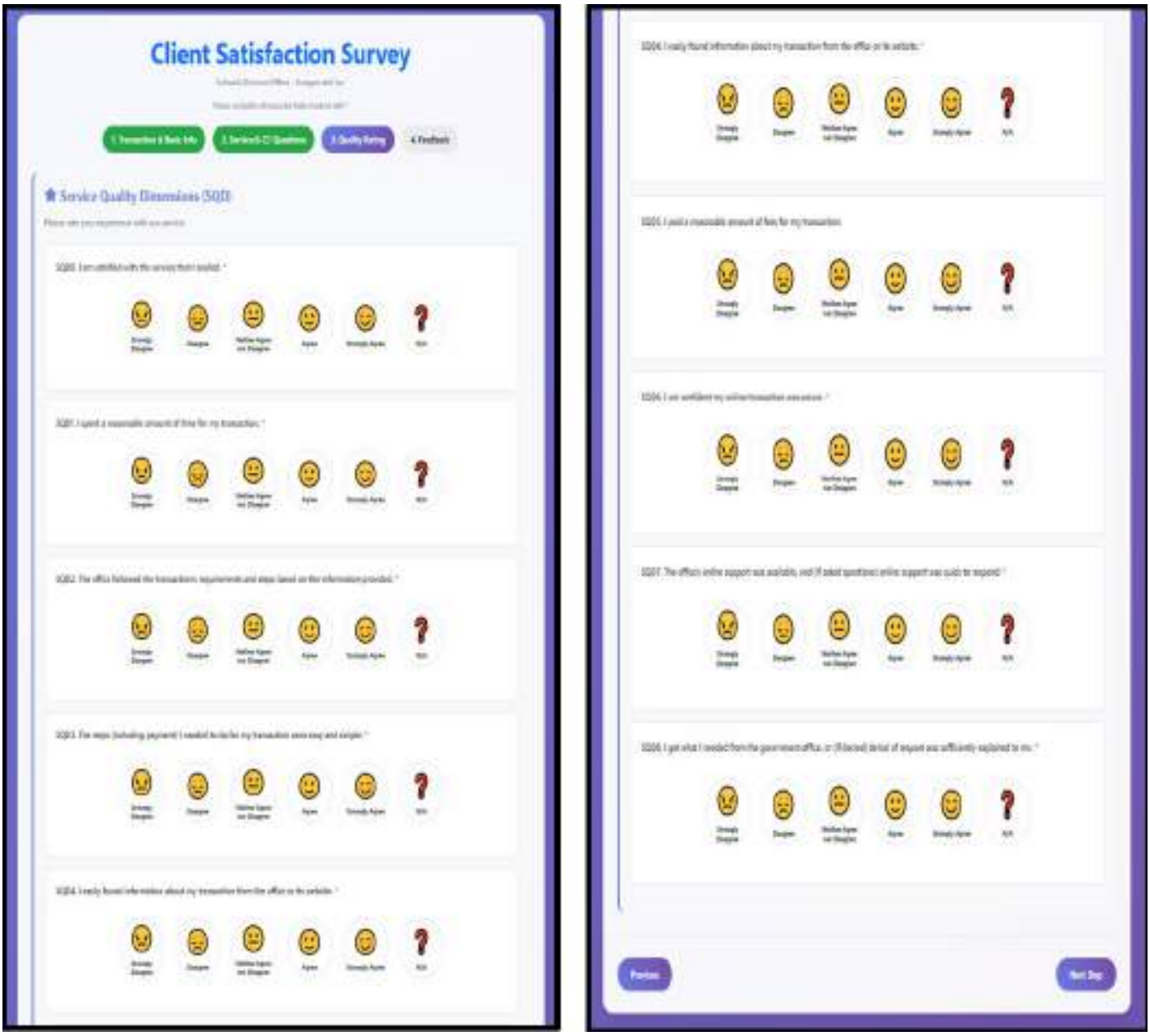


Fig 15 Service Quality Dimensions

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Service Quality Dimensions (Sqd)*

This section measures the client's overall experience with the service they received. Clients rate each statement using a satisfaction scale (from *Strongly Disagree* to *Strongly Agree*). The questions cover key areas such as:

- **Service Satisfaction** – How satisfied they were with the service.
- **Timeliness** – Whether the time spent for the transaction was reasonable.
- **Process Compliance** – If the office followed proper steps and requirements.
- **Simplicity of Steps** – How easy and simple the transaction steps were.
- **Information Accessibility** – How easily information was found in the office or online.
- **Fees** – Whether the fees were reasonable.
- **Online Security** – Confidence in the security of online transactions.
- **Online Support** – Availability and responsiveness of online assistance.
- **Results/Outcome** – Whether they received what they needed or were given a clear explanation if denied.

6. Lastly, the user will reach the final part of the form, which is the Suggestions section. The user is required to provide their suggestions so that the system can generate a sentiment analysis, determining whether the feedback is positive or negative through AI interpretation

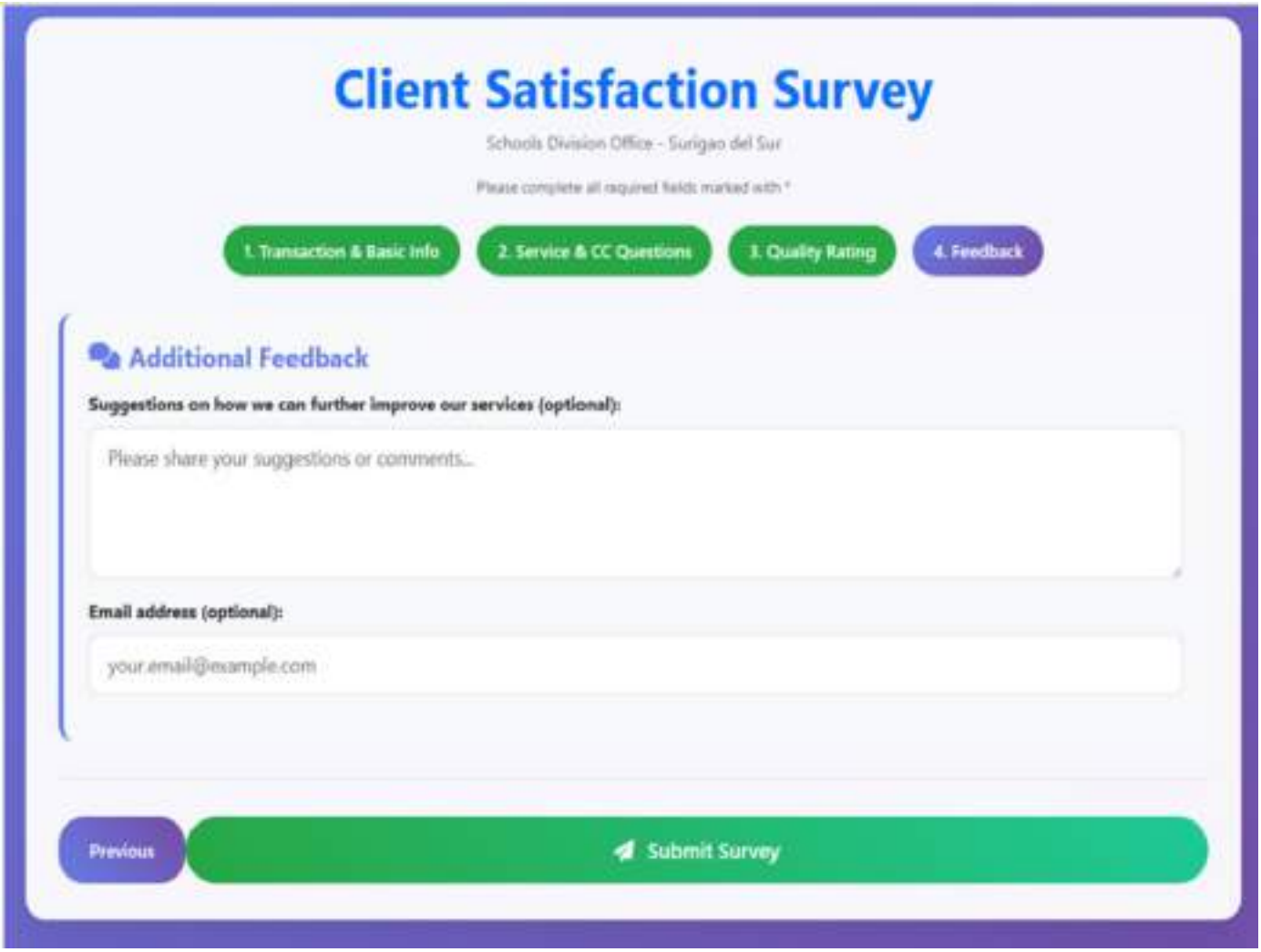
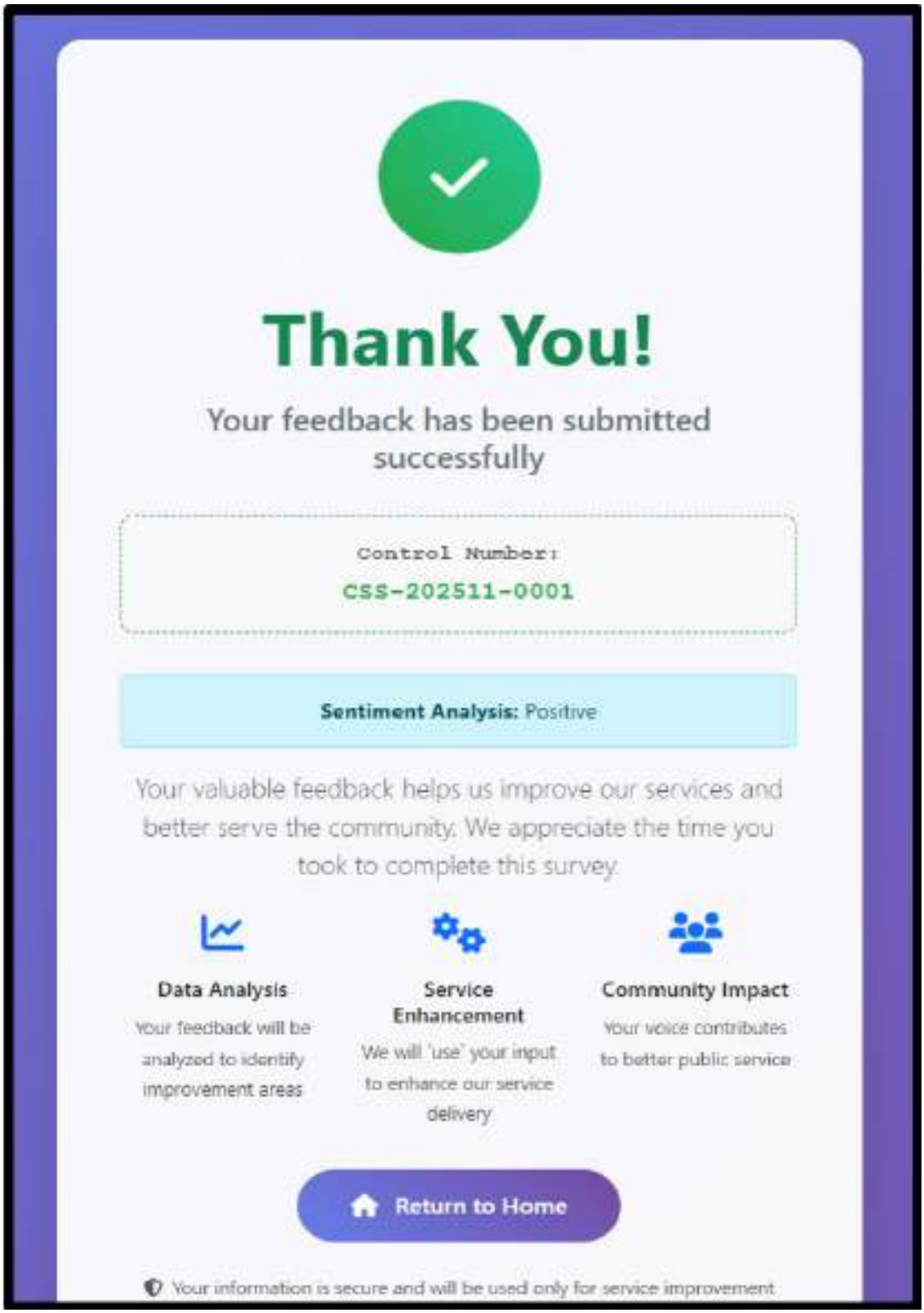


Fig 16 Feedback

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

7. After completing the survey form, a control number and the sentiment analysis result will be displayed on the screen. The data will then be saved and processed on the admin side. The users' responses will play a significant role in analyzing and improving the office's services.



ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Admin Side*

On the admin side of the system, there is a dashboard with filters that allow the admin to view specific data based on department, services, dates, and transaction types. The dashboard also displays overall satisfaction ratings, percentages, and detailed descriptions of the ratings. Additionally, the system provides insights into client satisfaction trends, highlights areas where services are performing well, and identifies aspects that may require improvement. This allows the administration to make data-driven decisions to enhance service quality and respond effectively to user feedback.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

In the dashboard also, the admin has the ability to generate a PDF report that follows the required format according to ARTA guidelines. This report consolidates all survey responses and sentiment analysis results, providing a clear overview of user feedback. The discussion and insights in the report are automatically generated by AI, summarizing trends, highlighting areas of high satisfaction, and identifying aspects that need improvement. This allows the administration to quickly understand overall service performance, track progress over time, and make informed decisions to enhance client satisfaction. Additionally, the PDF report can be easily shared with stakeholders, ensuring transparency and accountability in service delivery.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

CLIENT SATISFACTION MEASUREMENT REPORT

ARTA Memorandum Circular 2022-05

Period: October 01, 2025 to October 30, 2025
Generated: November 26, 2025 03:59 PM

I. EXECUTIVE SUMMARY

Total Survey Responses:	236
Survey Period:	October 01, 2025 to October 30, 2025
Overall Satisfaction Score (SQD0):	4.77/5 (95.4%)
Performance Rating:	Outstanding
ARTA Compliance Status:	COMPLIANT

Citizen's Charter Answers

CC1. Which of the following describes your awareness of the CC?	Responses	Percentage
1. I know what a CC is and I saw this office's CC.	64	27.1%
2. I know what a CC is but I did not see this office's CC.	60	25.4%
3. I learned of the CC only when I saw this office's CC.	57	24.2%
4. I do not know what a CC is and I did not see this office's CC.	55	23.3%

CC2. If aware of CC, would you say that the CC of this office was...?	Responses	Percentage
1. Easy to see	49	20.8%
2. Somewhat easy to see	40	16.9%
3. Difficult to see	46	19.5%
4. Not visible at all	49	20.8%

CC3. If aware of CC, how much did the CC help you in your transaction?	Responses	Percentage
1. Helped very much	62	26.3%
2. Somewhat helped	54	22.9%
3. Did not help	55	23.3%

II. SERVICES SURVEYED

The services surveyed were the following:

External Services		
Service Name	Responses	Total Transactions
Processing of Retirement and Separation	10	10
Release and Receiving of Records / Documents	4	4
Issuance of Certificates (COE, Service Records, Leave Credits Cert)	10	10
Certified True Copy (CTC)/Photocopy of documents	11	11
ORS Obligation Request and Status	20	20

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Dimension	Description	Score	%
SQD1	Responsiveness	4.34/5	86.8%
SQD2	Reliability	4.33/5	86.6%
SQD3	Access and Facilities	4.33/5	86.6%
SQD4	Communication	4.28/5	85.6%
SQD5	Costs	5.89/5	117.8%
SQD6	Integrity	1.62/5	32.4%
SQD7	Assurance	1.63/5	32.6%
SQD8	Outcome	1.61/5	32.2%

[SQD1 (Responsiveness) Discussion]

With a score of 4.34/5 (86.8%), processing time satisfaction is performing well above satisfactory levels. This indicates strong performance in this dimension and should be maintained as a standard of excellence.

[SQD2 (Reliability) Discussion]

With a score of 4.33/5 (86.6%), document requirements clarity is performing well above satisfactory levels. This indicates strong performance in this dimension and should be maintained as a standard of excellence.

[SQD3 (Access and Facilities) Discussion]

With a score of 4.33/5 (86.6%), procedural steps understanding is performing well above satisfactory levels. This indicates strong performance in this dimension and should be maintained as a standard of excellence.

[SQD4 (Communication) Discussion]

With a score of 4.28/5 (85.6%), information accessibility is performing well above satisfactory levels. This indicates strong performance in this dimension and should be maintained as a standard of excellence.

[SQD5 (Costs) Discussion]

With a score of 5.89/5 (117.8%), fee structure transparency is performing well above satisfactory levels. This indicates strong performance in this dimension and should be maintained as a standard of excellence.

[SQD6 (Integrity) Discussion]

At 1.62/5 (32.4%), office environment quality requires immediate attention and improvement. This below-satisfactory performance indicates significant client dissatisfaction that must be addressed through targeted corrective actions.

[SQD7 (Assurance) Discussion]

At 1.63/5 (32.6%), staff courtesy and professionalism requires immediate attention and improvement. This below-satisfactory performance indicates significant client dissatisfaction that must be addressed through targeted corrective actions.

[SQD8 (Outcome) Discussion]

At 1.61/5 (32.2%), overall government office experience requires immediate attention and improvement. This below-satisfactory performance indicates significant client dissatisfaction that must be addressed through targeted corrective actions.

VI. RECOMMENDATIONS

The organization demonstrates exemplary service delivery. To sustain this excellence, continue implementing best practices, share successful strategies across departments, recognize and reward high-performing staff, and maintain vigilant monitoring of client satisfaction metrics. Priority areas requiring attention include: Overall Government Office (score: 1.61/5), Office Environment (score: 1.62/5), Staff Courtesy (score: 1.63/5).

VII. ARTA COMPLIANCE ASSESSMENT

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX F:
CONSOLIDATED DATA OF CONFUSION MATRIX

Consolidated the 3 IT experts' human labels, applied majority voting per comment, compared the result with the system label, and then computed the confusion matrix, Accuracy, Precision, Recall, and F1-Score for the BERT-based sentiment classification.

Majority-Vote Final Labels

No.	CSM Comment	Expert 1	Expert 2	Expert 3	Final Human Label	System Label
1	Basta deped Surigao ♡	Positive	Positive	Positive	**Positive**	Positive
2	the security personnel should improve their attitude	Positive	Positive	Negative	**Positive**	Negative
3	Security personnel is approachable	Positive	Positive	Positive	**Positive**	Positive
4	none	Neutral	Neutral	Neutral	**Neutral**	Neutral
5	Dugay kayo ug wait	Negative	Negative	Negative	**Negative**	Negative
6	Dako ang tabang ninyo sa ako	Positive	Positive	Positive	**Positive**	Positive
7	nknek	Neutral	Neutral	Neutral	**Neutral**	Positive
8	Online processing must be available	Positive	Negative	Neutral	**No majority / tied**	Neutral
9	help	Neutral	Neutral	Neutral	**Neutral**	Positive
10	Dili ka gwapa	Neutral	Negative	Negative	**Negative**	Negative
11	Nered onloineer transaction	Neutral	Neutral	Neutral	**Neutral**	Neutral
12	keep ukp the good wdsork	Positive	Positive	Positive	**Positive**	Positive
13	ok	Positive	Positive	Positive	**Positive**	Neutral
14	Accommodating accommodating	Positive	Positive	Positive	**Positive**	Positive
15	kapoy	Neutral	Negative	Negative	**Negative**	Neutral
16	Swabe ra kaayo ang transaction	Positive	Positive	Positive	**Positive**	Positive
17	Appreciated kaayo ang effort	Positive	Positive	Positive	**Positive**	Positive
18	Complicated ang forms	Positive	Negative	Negative	**Negative**	Neutral
19	libog	Negative	Negative	Negative	**Negative**	Neutral
20	Naa puy guide. Helpful	Positive	Positive	Positive	**Positive**	Positive
21	Need more information kiosk	Positive	Neutral	Neutral	**Neutral**	Positive
22	no comment	Neutral	Neutral	Neutral	**Neutral**	Neutral
23	Deserve nyo ng award sa service	Positive	Positive	Positive	**Positive**	Positive
24	n/a	Neutral	Neutral	Neutral	**Neutral**	Neutral
25	Kulang ang staff	Positive	Negative	Neutral	**No majority / tied**	Neutral
26	satisfy	Positive	Positive	Positive	**Positive**	Positive
27	Air conditioned pa. Comfy jud	Positive	Positive	Positive	**Positive**	Positive
28	Dapat may aircon	Neutral	Negative	Neutral	**Neutral**	Neutral
29	GOODS KAU	Positive	Positive	Positive	**Positive**	Positive
30	Medjo init sa waiting area	Neutral	Negative	Negative	**Negative**	Negative

- *Note:*

Yellow rows (Items 8 and 25) are excluded from metric computation due to tied human ratings. Red cells in System Label column indicate mismatch with the Final Human Label.

➢ *Note on tied items*

Two comments had no majority human label among the 3 experts:

No. 8: Positive / Negative / Neutral
No. 25: Positive / Negative / Neutral

For strict evaluation, these two are best treated as excluded from metric computation because there is no stable ground truth from the validators. That leaves 28 valid items for performance measurement.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

➢ *Confusion matrix based on majority-vote labels*

Using the 28 items with a majority human label, the corrected confusion matrix is:

Actual \ Predicted	Positive	Neutral	Negative	Total
Positive	12	1	1	14
Neutral	3	5	0	8
Negative	0	3	3	6
Total	15	9	4	28

- *Note:*

Diagonal cells (green) indicate correctly classified instances.

➢ *Computation of Accuracy, Precision, Recall, and F1-Score*

To measure the performance of the BERT Model, the following formulas were used:

Where: *TP = True Positives, FP = False Positives, FN = False Negatives, TN = True Negatives*

- *Overall Accuracy*

The overall accuracy is computed as the total number of correct predictions divided by the total number of valid items:

$$\text{Accuracy} = (\text{Sum of diagonal values}) / \text{Total items}$$

$$\text{Accuracy} = (TP_{pos} + TP_{neu} + TP_{neg}) / n$$

$$\text{Accuracy} = (12 + 5 + 3) / 28$$

$$\text{Accuracy} = 20 / 28 = 0.7143$$

The overall accuracy of the BERT model is 71.43%.

- *Per-class Precision, Recall, and F1-Score*

Per-class metrics are computed using the one-vs-rest approach, where each class is treated as the positive class against all remaining classes. The TP, FP, FN, and TN values for each class are derived from the confusion matrix.

✓ *Positive Class*

- TP = 12 (correctly predicted as Positive)
- FP = 3 (predicted Positive but actually Neutral or Negative)
- FN = 2 (actually Positive but predicted Neutral or Negative)
- TN = 11 (correctly predicted as non-Positive) *[TN = 28 − TP − FP − FN = 28 − 12 − 3 − 2 = 11]*

$$\text{Accuracy} = (TP + TN) / \text{Total} = (12 + 11) / 28 = 23 / 28 = 0.8214 \rightarrow 82.14\%$$

$$\text{Precision} = TP / (TP + FP) = 12 / (12 + 3) = 12 / 15 = 0.8000 \rightarrow 80.00\%$$

$$\text{Recall} = TP / (TP + FN) = 12 / (12 + 2) = 12 / 14 = 0.8571 \rightarrow 85.71\%$$

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$\text{F1-Score} = 2 \times (0.8000 \times 0.8571) / (0.8000 + 0.8571)$$

$$\text{F1-Score} = 2 \times 0.6857 / 1.6571 = 1.3714 / 1.6571 = 0.8276 \rightarrow 82.76\%$$

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

✓ *Neutral Class*

- TP = 5 (correctly predicted as Neutral)
- FP = 4 (predicted Neutral but actually Positive or Negative)
- FN = 3 (actually Neutral but predicted Positive or Negative)
- TN = 16 (correctly predicted as non-Neutral) *[TN = 28 − 5 − 4 − 3 = 16]*

$$\text{Accuracy} = (TP + TN) / \text{Total} = (5 + 16) / 28 = 21 / 28 = 0.7500 \rightarrow 75.00\%$$

$$\text{Precision} = TP / (TP + FP) = 5 / (5 + 4) = 5 / 9 = 0.5556 \rightarrow 55.56\%$$

$$\text{Recall} = TP / (TP + FN) = 5 / (5 + 3) = 5 / 8 = 0.6250 \rightarrow 62.50\%$$

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$\text{F1-Score} = 2 \times (0.5556 \times 0.6250) / (0.5556 + 0.6250)$$
$$\text{F1-Score} = 2 \times 0.3472 / 1.1806 = 0.6944 / 1.1806 = 0.5882 \rightarrow 58.82\%$$

✓ *3) Negative Class*

- TP = 3 (correctly predicted as Negative)
- FP = 1 (predicted Negative but actually Positive or Neutral)
- FN = 3 (actually Negative but predicted Positive or Neutral)
- TN = 21 (correctly predicted as non-Negative) *[TN = 28 − 3 − 1 − 3 = 21]*

$$\text{Accuracy} = (TP + TN) / \text{Total} = (3 + 21) / 28 = 24 / 28 = 0.8571 \rightarrow 85.71\%$$

$$\text{Precision} = TP / (TP + FP) = 3 / (3 + 1) = 3 / 4 = 0.7500 \rightarrow 75.00\%$$

$$\text{Recall} = TP / (TP + FN) = 3 / (3 + 3) = 3 / 6 = 0.5000 \rightarrow 50.00\%$$

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$\text{F1-Score} = 2 \times (0.7500 \times 0.5000) / (0.7500 + 0.5000)$$

$$\text{F1-Score} = 2 \times 0.3750 / 1.2500 = 0.7500 / 1.2500 = 0.6000 \rightarrow 60.00\%$$

Summary of Performance Metrics

Class	TP	FP	FN	TN	Accuracy	Precision	Recall	F1-Score
Positive	12	3	2	11	82.14%	80.00%	85.71%	82.76%
Neutral	5	4	3	16	75.00%	55.56%	62.50%	58.82%
Negative	3	1	3	21	85.71%	75.00%	50.00%	60.00%
Overall (Macro-avg)	**20**	**—**	**—**	**—**	**71.43%**	**70.19%**	**66.07%**	**67.19%**

- *Note:*

Metrics computed using the one-vs-rest approach. TP = True Positive, FP = False Positive, FN = False Negative, TN = True Negative. Accuracy per class = (TP+TN)/Total. Overall row reports macro-averages across all three classes.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX G:
ARTA COMPLIANCE VERIFICATION CHECKLIST

ARTA COMPLIANCE VERIFICATION CHECKLIST

Automated Client Satisfaction Measurement (CSM) System

Dapfd Division of Surigao del Sur

Instructions: Use this checklist during system review, monitoring visits, or pre-submission audits to verify that the automated CSM system complies with ARTA standards under Republic Act No. 11032. Check (✓) the appropriate column for each indicator.

No.	Compliance Indicator	✓ Complied	✗ Not Complied	N/A	Remarks
A. Survey Administration and Coverage					
1	The CSM survey is administered to all clients upon completion of each frontline service transaction.	✓	✗		
2	The system is accessible across all service offices of the agency through the online kiosk, web-based platform, or QR-code access.	✓	✗		
3	The survey remains available throughout the entire required reporting period without service interruption.	✓	✗		
4	Clients are not required to create an account or log in before submitting feedback.	✓	✗		
5	The system minimizes duplicate submissions per transaction through automated controls.	✓	✗		
B. Survey Instrument and ARTA Indicators					
6	The survey uses ARTA-prescribed CSM indicators (e.g., responsiveness, quality, timeliness, efficiency, and overall satisfaction).	✓	✗		
7	All survey questions are standardized and uniform across all offices and service types.	✓	✗		
8	The survey includes a required section for open-ended client comments or suggestions.	✓	✗		
9	All required fields are enforced by the system — no submission is accepted with incomplete responses.	✓	✗		
10	The rating scale used in the survey conforms to the scale prescribed by ARTA guidelines.	✓	✗		
C. Data Processing and Consolidation					
11	All submitted responses are automatically recorded in a centralized database in real time upon submission.	✓	✗		
12	Responses from all offices and service types are consolidated at the division or agency level without manual re-encoding.	✓	✗		
13	The system automatically computes satisfaction scores and ratings from submitted responses.	✓	✗		
14	The system can filter and disaggregate data by office, service type, and reporting period as required by ARTA.	✓	✗		
15	Submitted data cannot be altered or modified by unauthorized personnel after final recording.	✓	✗		
D. Report Generation and ARTA Format					
16	The system automatically generates CSM reports following the format and structure prescribed by ARTA.	✓	✗		
17	Generated reports include all required content: office-level results, service-type breakdown, overall rating, and period covered.	✓	✗		
18	Reports are available in downloadable format (PDF or equivalent) suitable for regulatory submission.	✓	✗		
19	The system produces a consolidated summary report covering all offices without requiring manual data gathering from individual units.	✓	✗		
20	Reports include a summary or classification of open-ended client feedback using the automated sentiment classification model.	✓	✗		
E. Timeliness of Submission					
21	CSM reports are generated and made available within the reporting period prescribed by ARTA (quarterly or semester basis).	✓	✗		
22	Authorized personnel can access and download the final report immediately after the reporting period closes, without additional preparation.	✓	✗		
23	Report preparation does not require additional manual consolidation or reformatting prior to submission to ARTA.	✓	✗		
24	The system notifies responsible personnel (e.g., weekly office ratings) to support continuous monitoring and deadline awareness.	✓	✗		
F. Transparency, Accessibility, and Data Security					
25	A real-time monitoring dashboard is available to authorized administrators for tracking office-level satisfaction ratings.	✓	✗		
26	CSM results are made accessible to authorized stakeholders and the public in accordance with ARTA transparency requirements.	✓	✗		
27	The system maintains an audit trail or activity log of all submitted responses and generated reports.	✓	✗		
28	Data privacy and confidentiality of individual responses are protected in accordance with Republic Act No. 10173 (Data Privacy Act of 2012).	✓	✗		
29	Role-based access control ensures only authorized personnel can view, export, or manage sensitive data.	✓	✗		
30	The system includes accessibility features that support differently-abled clients (e.g., visual or hearing impairments), consistent with ARTA's inclusivity mandate.	✓	✗		

Overall Compliance Status:

☑ Fully Complied ☐ Partially Complied ☐ Not Complied

Remarks / Observations:

The automated CSM system is highly functional, user-friendly, and compliant with data privacy and accessibility standards.

Verified by: _____ Date: _____

Position: _____ Office: _____

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX H:
BERT SENTIMENT ANALYSIS TEST RUN REPORT

Automated Client Satisfaction Measurement (CSM) System

DepEd Division of Surigao del Sur

Data Collection Period: February 2026 to March 2026

1. Dataset Framing and Methodology

This section presents the dataset preparation and descriptive analysis for the BERT-based sentiment analysis experiment using feedback collected through the Client Satisfaction Measurement (CSM) module. The analysis is anchored on actual survey remarks drawn from the production database and restricted to entries with non-empty textual feedback.

From an available pool of 1,319 feedback-bearing records dated February 2026 to March 2026, a seeded random sample of 350 responses was selected using a fixed random seed (20260523) to ensure reproducibility. This sample served as the working corpus for the sentiment classification workflow.

The sampled corpus is dominated by positive sentiment (68%), followed by neutral sentiment (30.9%), while negative sentiment remains limited at 1.1%. The resulting class structure indicates a favorable respondent outlook but also reveals a strong class imbalance that must be considered during model training and evaluation.

Summary Statistics

Indicator	Value
Eligible Feedback Rows	1,319
Random Sample Size	350
Average Satisfaction (SQD0)	4.72
Class Balance Risk	High (Minority class count: 4)

2. Statistical Summary of the Sampled Corpus

Dataset	Records	Positive	Neutral	Negative	Avg. SQD0
Overall sampled corpus	350	238 (68%)	108 (30.9%)	4 (1.1%)	4.72
Training subset (30%)	105	73 (69.5%)	31 (29.5%)	1 (1%)	4.79
Validation subset (50%)	175	113 (64.6%)	60 (34.3%)	2 (1.1%)	4.63
Testing subset (20%)	70	52 (74.3%)	17 (24.3%)	1 (1.4%)	4.84

Partition	Assigned Share	Actual Records	Purpose in Experiment
Training	30%	105	Used to fit the BERT classifier parameters on the sampled feedback corpus.
Validation	50%	175	Used to monitor convergence, inspect class behavior, and guide model tuning decisions.
Testing	20%	70	Reserved for final performance checking on held-out examples.

3. Interpretation and Implications

The statistical summary shows that the sampled remarks remain consistently positive across all three partitions. Validation was intentionally assigned the largest share of the corpus (50%), reflecting the experimental design and enabling closer monitoring of model behavior during development.

The negative class is severely underrepresented (4 instances out of 350), which increases the risk of optimistic accuracy estimates and weak recall for adverse feedback. Class-wise precision, recall, and F1-Score are therefore necessary for a complete evaluation of classification performance.

Class balance diagnostics: majority class size = 238; minority class size = 4; imbalance ratio = 59.5; risk level = High. For a production-grade BERT experiment, additional negative feedback samples or class-balancing strategies such as weighted loss or stratified annotation should be considered before drawing comparative conclusions.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

4. Illustrative Feedback Instances from the Sample

Record ID	Date	Sentiment	Feedback
3242	2026-03-25	Neutral	Good
2910	2026-03-10	Positive	accommodating
2628	2026-02-26	Positive	NONE
2345	2026-02-16	Positive	NICE
2984	2026-03-12	Positive	None

5. Illustrative BERT workflow code

A. Dataset Preparation and Split

```
import pandas as pd
from sklearn.model_selection import train_test_split

# df columns expected from the CSM export:
# suggestions -> feedback text
# sentiment -> positive / neutral / negative

df = pd.read_csv("csm_feedback_feb_mar_2026.csv")
df = df[df["suggestions"].notna()].copy()
df["suggestions"] = df["suggestions"].astype(str).str.strip()
df = df[df["suggestions"] != ""]

# Keep the same report sample size
df = df.sample(n=200, random_state=20260523).reset_index(drop=True)

label_map = {"negative": 0, "neutral": 1, "positive": 2}
df["label"] = df["sentiment"].str.lower().map(label_map)

train_df = df.iloc[:100].copy()
val_df = df.iloc[100:150].copy()
test_df = df.iloc[150:].copy()

print("Train:", len(train_df))
print("Validation:", len(val_df))
print("Test:", len(test_df))
```

B. Tokenization and Dataset Construction

```
from datasets import Dataset
from transformers import AutoTokenizer

tokenizer = AutoTokenizer.from_pretrained("bert-base-uncased")

def tokenize(batch):
    return tokenizer(
        batch["suggestions"],
        truncation=True,
        padding="max_length",
        max_length=128,
    )

train_dataset = Dataset.from_pandas(train_df[["suggestions", "label"]]).map(tokenize, batched=True)
val_dataset = Dataset.from_pandas(val_df[["suggestions", "label"]]).map(tokenize, batched=True)
test_dataset = Dataset.from_pandas(test_df[["suggestions", "label"]]).map(tokenize, batched=True)

train_dataset = train_dataset.rename_column("label", "labels")
val_dataset = val_dataset.rename_column("label", "labels")
test_dataset = test_dataset.rename_column("label", "labels")

train_dataset.set_format(type="torch", columns=["input_ids", "attention_mask", "labels"])
val_dataset.set_format(type="torch", columns=["input_ids", "attention_mask", "labels"])
test_dataset.set_format(type="torch", columns=["input_ids", "attention_mask", "labels"])
```

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

C. Training and Validation Code

```
import numpy as np
from sklearn.metrics import accuracy_score, precision_recall_fscore_support
from transformers import AutoModelForSequenceClassification, TrainingArguments, Trainer

model = AutoModelForSequenceClassification.from_pretrained(
    "bert-base-uncased",
    num_labels=3,
)

def compute_metrics(eval_pred):
    logits, labels = eval_pred
    predictions = np.argmax(logits, axis=1)
    precision, recall, f1, _ = precision_recall_fscore_support(
        labels,
        predictions,
        average="macro",
        zero_division=0,
    )
    accuracy = accuracy_score(labels, predictions)
    return {
        "accuracy": accuracy,
        "precision_macro": precision,
        "recall_macro": recall,
        "f1_macro": f1,
    }

training_args = TrainingArguments(
    output_dir="./bert-csm-results",
    eval_strategy="epoch",
    save_strategy="epoch",
    logging_strategy="epoch",
    learning_rate=2e-5,
    per_device_train_batch_size=8,
    per_device_eval_batch_size=8,
    num_train_epochs=5,
    weight_decay=0.01,
    load_best_model_at_end=True,
    metric_for_best_model="f1_macro",
    greater_is_better=True,
)

trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=train_dataset,
    eval_dataset=val_dataset,
    tokenizer=tokenizer,
    compute_metrics=compute_metrics,
)

trainer.train()
```

D. Testing and Final Evaluation Code

```
test_metrics = trainer.evaluate(test_dataset)
print("Held-out test metrics:", test_metrics)

predictions = trainer.predict(test_dataset)
predicted_labels = predictions.predictions.argmax(axis=1)

results = test_df.copy()
results["predicted_label"] = predicted_labels
inverse_label_map = {0: "negative", 1: "neutral", 2: "positive"}
results["predicted_sentiment"] = results["predicted_label"].map(inverse_label_map)

print(results[["suggestions", "sentiment", "predicted_sentiment"]].head(10))
results.to_csv("bert_csm_test_predictions.csv", index=False)
```

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX I:
QUALITATIVE RESEARCH INSTRUMENT VALIDATION SHEET AND LETTER

May 14, 2026

UCILLE P. GALVEZ
Senior Education Program Specialist / Research Coordinator
Schools Division of Surigao del Sur
Brgy. Mabua, Tandag City

Dear Ma'am Galvez:

Greetings of peace and excellence!

The undersigned researcher is currently conducting a research study entitled "Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao Del Sur" in partial fulfillment of the requirements for Master of Science in Computer Science.

In line with this, may we respectfully request your expertise as a Research Coordinator and academic professional to validate the qualitative research instrument, particularly the Interview Guide, which will be utilized in gathering the necessary data for the study. Your comments, suggestions, and recommendations regarding the clarity, relevance, coherence, and appropriateness of the interview questions will be highly appreciated and will contribute significantly to the improvement of the research instrument.

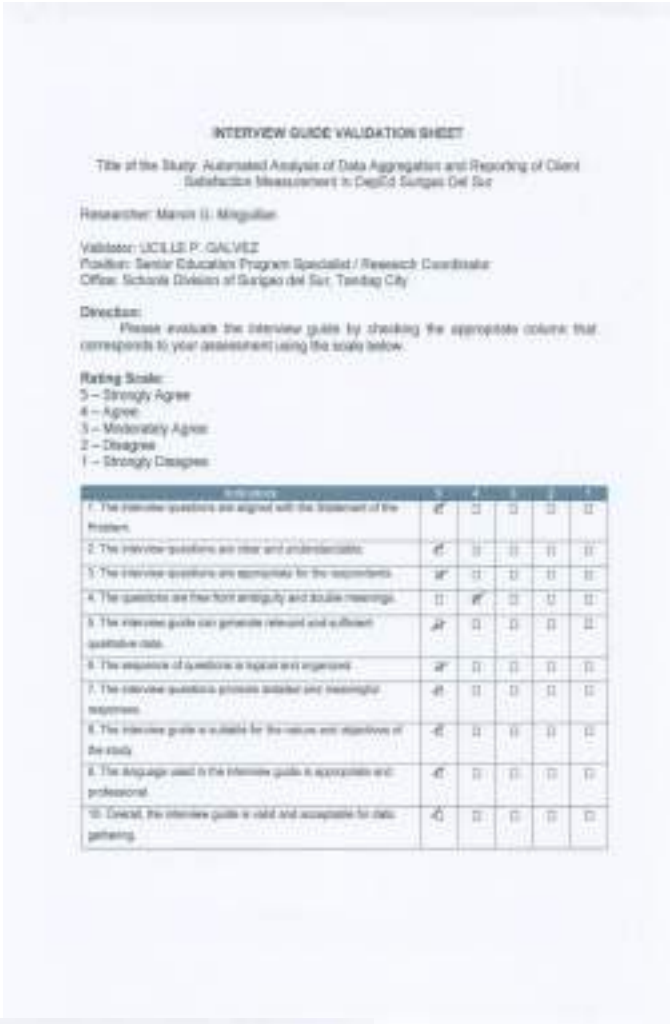
Attached herewith is the Validation Sheet for your evaluation and assessment. Rest assured that all your inputs and recommendations will be treated with utmost confidentiality and will solely be used for academic purposes.

We sincerely hope for your favorable response and valuable assistance in this endeavor.

Thank you very much and more power!

Respectfully yours,

MARVIN G. MINGUILLAN
Researcher



INTERVIEW GUIDE VALIDATION SHEET

Title of the Study: Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao Del Sur

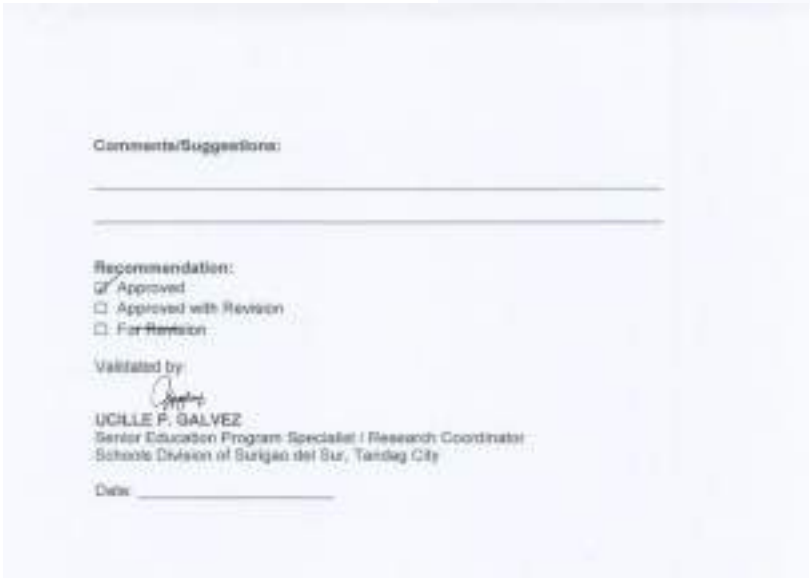
Researcher: Marvin G. Minguillan

Validator: UCILLE P. GALVEZ
Position: Senior Education Program Specialist / Research Coordinator
Office: Schools Division of Surigao del Sur, Tandag City

Direction:
Please evaluate the interview guide by checking the appropriate column that corresponds to your assessment using the scale below.

Rating Scale:
5 – Strongly Agree
4 – Agree
3 – Moderately Agree
2 – Disagree
1 – Strongly Disagree

Statement	5	4	3	2	1
1. The interview questions are aligned with the Statement of the Problem.	☑	☐	☐	☐	☐
2. The interview questions are clear and understandable.	☑	☐	☐	☐	☐
3. The interview questions are appropriate for the respondents.	☑	☐	☐	☐	☐
4. The questions are free from ambiguity and double meanings.	☐	☑	☐	☐	☐
5. The interview guide can generate relevant and sufficient qualitative data.	☑	☐	☐	☐	☐
6. The sequence of questions is logical and organized.	☑	☐	☐	☐	☐
7. The interview questions promote detailed and meaningful responses.	☑	☐	☐	☐	☐
8. The interview guide is suitable for the nature and objectives of the study.	☑	☐	☐	☐	☐
9. The language used in the interview guide is appropriate and professional.	☑	☐	☐	☐	☐
10. Overall, the interview guide is valid and acceptable for data gathering.	☑	☐	☐	☐	☐



Comments/Suggestions:

_____

_____

Recommendation:
☑ Approved
☐ Approved with Revision
☐ For Revision

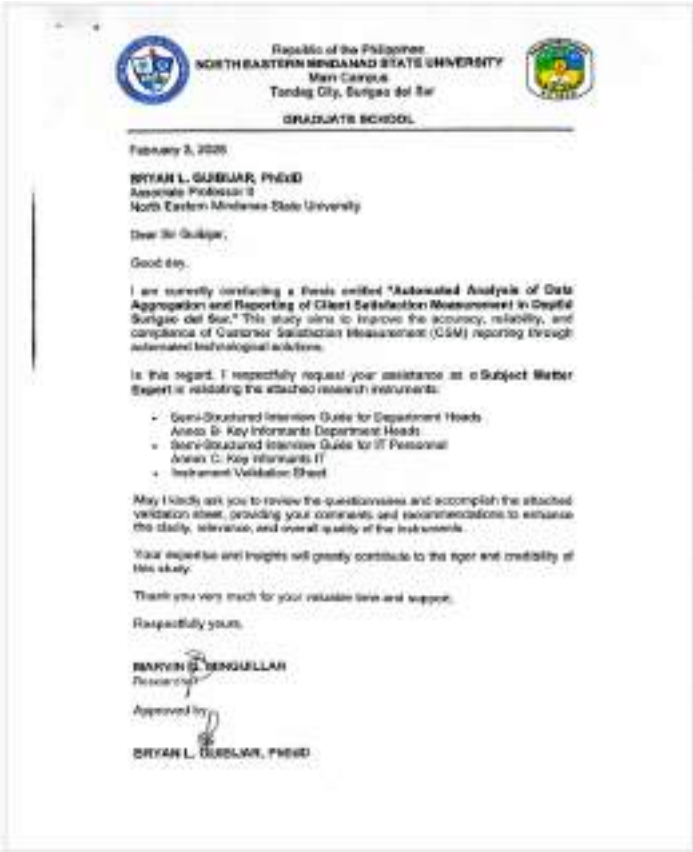
Validated by:

UCILLE P. GALVEZ
Senior Education Program Specialist / Research Coordinator
Schools Division of Surigao del Sur, Tandag City

Date: _____

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX J:
LETTER TO VALIDATORS AND VALIDATION SHEETS



Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

February 3, 2026

BRYAN L. GUBALANI, PhDEd
Associate Professor II
North Eastern Mindanao State University

Dear Sir Gubalan,

Good day.

I am currently conducting a thesis entitled "**Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in Depted Surigao del Sur.**" This study aims to improve the accuracy, reliability, and compliance of Customer Satisfaction Measurement (CSM) reporting through automated technological solutions.

In this regard, I respectfully request your assistance as a **Subject Matter Expert** in validating the attached research instruments:

- Semi-Structured Interview Guide for Department Heads
  Annex B- Key Informants Department Heads
- Semi-Structured Interview Guide for IT Personnel
  Annex C- Key Informants IT
- Instrument Validation Sheet

May I kindly ask you to review the questionnaires and accomplish the attached validation sheet, providing your comments and recommendations to enhance the clarity, relevance, and overall quality of the instruments.

Your expertise and insights will greatly contribute to the rigor and credibility of this study.

Thank you very much for your valuable time and support.

Respectfully yours,

MARVIN G. MINGUILLAN
Researcher

Approved by:

BRYAN L. GUBALANI, PhDEd

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

VALIDATION SHEET
For Qualitative Research Instrument (Semi-Structured Interview Guide)

Thesis Title:

Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in Depted Surigao del Sur

Researcher: Marvin G. Minguillan

Validator's Name: BRYAN L. GUBALANI
Designation/Position: ASSOCIATE PROFESSOR II
Area of Expertise: COMPUTER SCIENCE
Years of Experience:
Date: FEB 13, 2026

I. PURPOSE OF VALIDATION

This validation form is intended to evaluate the content validity, clarity, relevance, and appropriateness of the Semi-Structured Interview Guides designed for Department Heads and IT Key Informants. Your expert judgment will help ensure that the instrument is aligned with the Statement of the Problem and is suitable for qualitative data collection.

II. INSTRUCTIONS TO THE VALIDATOR

Please review the attached Semi-Structured Interview Guide carefully. Rate each criterion using the scale below and provide comments or recommendations for improvement.

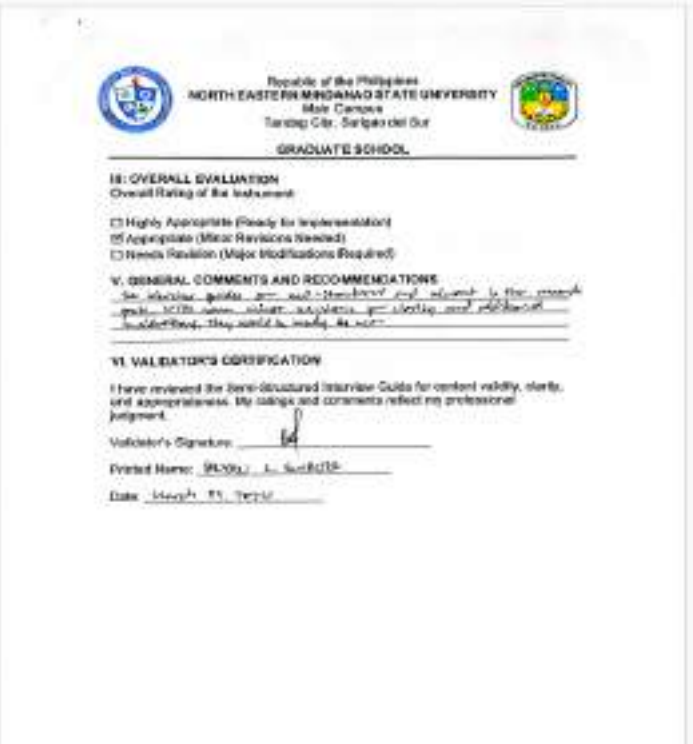
Rating Scale

Rating	Description
4	Highly Appropriate
3	Appropriate
2	Needs Minor Revision
1	Needs Major Revision

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

Criteria	4	3	2	1	Comments/ Recommendations
A. CONTENT VALIDITY					
1. The questions are aligned with the Statement of the Problem.	✓				
2. The questions adequately cover accuracy and reliability of CSM reports.	✓				
3. The questions address compliance with ARTA standards.	✓				
4. The questions are relevant to the roles of the key informants.	✓				
5. The instrument captures sufficient data to answer the research objectives.	✓				
B. CLARITY AND STRUCTURE					
6. The questions are clear and easy to understand.		✓			
7. The language used is appropriate for the intended key informants.		✓			
8. The sequence of questions is logical and well-organized.		✓			
9. The follow-up prompts support deeper responses.		✓			
10. The introductory script is clear and ethically appropriate.		✓			
C. TECHNICAL AND RESEARCH QUALITY					
11. The instrument is suitable for qualitative thematic analysis.	✓				
12. The questions are open-ended and non-leading.	✓				
13. The instrument avoids ambiguity or double-barreled questions.	✓				
14. The number of questions is appropriate for semi-structured interviews.	✓				
15. The instrument can generate meaningful insights for system improvement.	✓				



Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

III. OVERALL EVALUATION
Overall Rating of the Instrument

☐ Highly Appropriate (Ready for Implementation)
☑ Appropriate (Minor Revisions Needed)
☐ Needs Revision (Major Modifications Required)

V. GENERAL COMMENTS AND RECOMMENDATIONS

the interview guides are well-structured and relevant to the research goals. With some minor adjustments for clarity and refinement, they would be ready to use.

VI. VALIDATOR'S CERTIFICATION

I have reviewed the Semi-Structured Interview Guide for content validity, clarity, and appropriateness. My ratings and comments reflect my professional judgment.

Validator's Signature:
Printed Name: BRYAN L. GUBALANI
Date: March 11, 2026

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

February 3, 2026

MELBRICK A. EVALLAR, MSCS
Associate Professor II
North Eastern Mindanao State University

Dear Sir Evallar,

Good day

I am currently conducting a thesis entitled **"Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao del Sur."** This study aims to improve the accuracy, reliability, and compliance of Customer Satisfaction Measurement (CSM) reporting through automated technological solutions.

In this regard, I respectfully request your assistance as a **Subject Matter Expert** in validating the attached research instruments:

- Semi-Structured Interview Guide for Department Heads
  Annex B- Key Informants Department Heads
- Semi-Structured Interview Guide for IT Personnel
  Annex C- Key Informants IT
- Instrument Validation Sheet

May I kindly ask you to review the questionnaires and accomplish the attached validation sheet, providing your comments and recommendations to enhance the clarity, relevance, and overall quality of the instruments.

Your expertise and insights will greatly contribute to the rigor and credibility of this study.

Thank you very much for your valuable time and support.

Respectfully yours,

MARVIN G. MINGUILLAN
Researcher

Approved by:

MELBRICK A. EVALLAR, MSCS

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

VALIDATION SHEET
For Qualitative Research Instrument (Semi-Structured Interview Guide)

Thesis Title:

Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao del Sur

Researcher: Marvin G. Minguillan

Validator's Name: [illegible]
Designation/Position: [illegible]
Area of Expertise: [illegible]
Years of Experience: [illegible]
Date: 01-03-2026

I. PURPOSE OF VALIDATION

This validation form is intended to evaluate the content validity, clarity, relevance, and appropriateness of the Semi-Structured Interview Guides designed for Department Heads and IT Key Informants. Your expert judgment will help ensure that the instrument is aligned with the Statement of the Problem and is suitable for qualitative data collection.

II. INSTRUCTIONS TO THE VALIDATOR

Please review the attached Semi-Structured Interview Guide carefully. Rate each criterion using the scale below and provide comments or recommendations for improvement.

Rating Scale

Rating	Description
4	Highly Appropriate
3	Appropriate
2	Needs Minor Revision
1	Needs Major Revision

[illegible]

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

Criteria	4	3	2	1	Comments/ Recommendations
A. CONTENT VALIDITY					
1. The questions are aligned with the Statement of the Problem.		✓			
2. The questions adequately cover accuracy and reliability of CSM reports.		✓			
3. The questions address compliance with ARTA standards.	✓				
4. The questions are relevant to the roles of the key informants.	✓				
5. The instrument captures sufficient data to answer the research objectives.	✓				
B. CLARITY AND STRUCTURE					
6. The questions are clear and easy to understand.	✓				
7. The language used is appropriate for the intended key informants.	✓				
8. The sequence of questions is logical and well-organized.	✓				
9. The follow-up prompts support deeper responses.		✓			
10. The introductory script is clear and ethically appropriate.	✓				
C. TECHNICAL AND RESEARCH QUALITY					
11. The instrument is suitable for qualitative thematic analysis.	✓				
12. The questions are open-ended and non-leading.	✓				
13. The instrument avoids ambiguity or double-barreled questions.	✓				
14. The number of questions is appropriate for semi-structured interviews.	✓				
15. The instrument can generate meaningful insights for system improvement.	✓				

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

III. OVERALL EVALUATION
Overall Rating of the Instrument:

☐ Highly Appropriate (Ready for Implementation)
☐ Appropriate (Minor Revisions Needed)
☐ Needs Revision (Major Modifications Required)

V. GENERAL COMMENTS AND RECOMMENDATIONS

VI. VALIDATOR'S CERTIFICATION

I have reviewed the Semi-Structured Interview Guide for content validity, clarity, and appropriateness. My ratings and comments reflect my professional judgment.

Validator's Signature: _____
Printed Name: [illegible]
Date: _____

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

February 3, 2026

MICHAEL L. ESTAL, MSCS, EDPSE
Instructor III
North Eastern Mindanao State University

Dear Sir Estal,

Good day.

I am currently conducting a thesis entitled **"Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao del Sur."** This study aims to improve the accuracy, reliability, and compliance of Customer Satisfaction Measurement (CSM) reporting through automated technological solutions.

In this regard, I respectfully request your assistance as a **Subject Matter Expert** in validating the attached research instruments:

- Semi-Structured Interview Guide for Department Heads
  Annex B- Key Informants Department Heads
- Semi-Structured Interview Guide for IT Personnel
  Annex C- Key Informants IT
- Instrument Validation Sheet

May I kindly ask you to review the questionnaires and accomplish the attached validation sheet, providing your comments and recommendations to enhance the clarity, relevance, and overall quality of the instruments.

Your expertise and insights will greatly contribute to the rigor and credibility of this study.

Thank you very much for your valuable time and support.

Respectfully yours,

MARVIN G. MINGUILLAN
Researcher

Approved by:

MICHAEL L. ESTAL, MSCS, EDPSE
Thesis Adviser

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

VALIDATION SHEET
For Qualitative Research Instrument (Semi-Structured Interview Guide)

Thesis Title:

Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao del Sur

Researcher: Marvin G. Minguillan

Validator's Name: MICHAEL L. ESTAL
Designation/Position: Instructor III
Area of Expertise: Computer Science
Years of Experience: 7
Date: March 5, 2026

I. PURPOSE OF VALIDATION

This validation form is intended to evaluate the content validity, clarity, relevance, and appropriateness of the Semi-Structured Interview Guides designed for Department Heads and IT Key Informants. Your expert judgment will help ensure that the instrument is aligned with the Statement of the Problem and is suitable for qualitative data collection.

II. INSTRUCTIONS TO THE VALIDATOR

Please review the attached Semi-Structured Interview Guide carefully. Rate each criterion using the scale below and provide comments or recommendations for improvement.

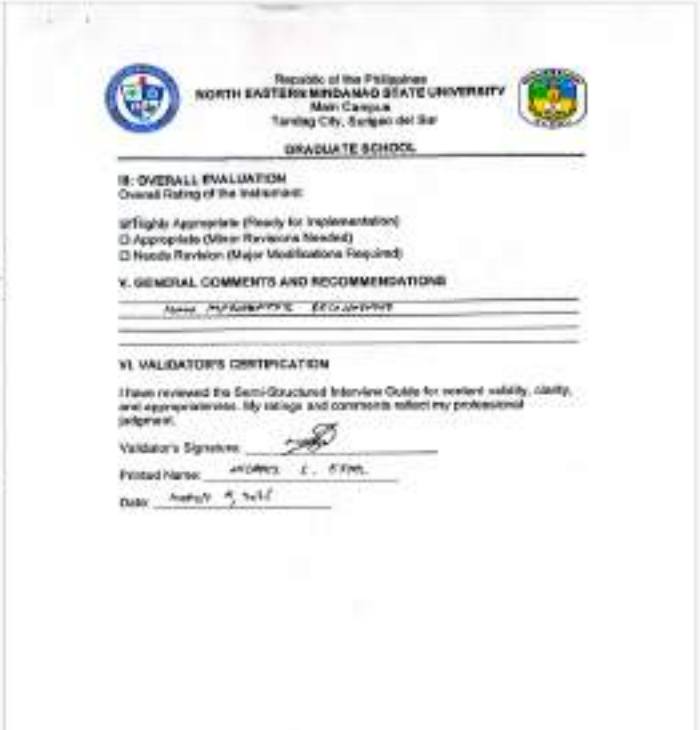
Rating Scale

Rating	Description
4	Highly Appropriate
3	Appropriate
2	Needs Minor Revision
1	Needs Major Revision

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

Criteria	4	3	2	1	Comments/ Recommendations
A. CONTENT VALIDITY					
1. The questions are aligned with the Statement of the Problem.	✓				
2. The questions adequately cover accuracy and reliability of CSM reports.	✓				confirm against the Statement about data validation report inconsistencies
3. The questions address compliance with ARTA standards.	✓				evaluate against and ARTA Citizen's Charter
4. The questions are relevant to the roles of the key informants.	✓				
5. The instrument captures sufficient data to answer the research objectives.	✓				
B. CLARITY AND STRUCTURE					
6. The questions are clear and easy to understand.	✓				
7. The language used is appropriate for the intended key informants.	✓				
8. The sequence of questions is logical and well-organized.	✓				
9. The follow-up prompts support deeper responses.	✓				"Last add one more question (optional)"
10. The introductory script is clear and ethically appropriate.	✓				
C. TECHNICAL AND RESEARCH QUALITY					
11. The instrument is suitable for qualitative thematic analysis.	✓				
12. The questions are open-ended and non-leading.	✓				
13. The instrument avoids ambiguity or double-barreled questions.		✓			Add the phrasing with separate concepts
14. The number of questions is appropriate for semi-structured interviews.	✓				(optional)
15. The instrument can generate meaningful insights for system improvement.	✓				



Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

III. OVERALL EVALUATION
Overall Rating of the Instrument:

☑ Highly Appropriate (Ready for Implementation)
☐ Appropriate (Minor Revisions Needed)
☐ Needs Revision (Major Modifications Required)

V. GENERAL COMMENTS AND RECOMMENDATIONS

None. Appropriate. Excellent.

VI. VALIDATOR'S CERTIFICATION

I have reviewed the Semi-Structured Interview Guide for content validity, clarity, and appropriateness. My ratings and comments reflect my professional judgment.

Validator's Signature: _____
Printed Name: MICHAEL L. ESTAL
Date: March 5, 2026

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

February 3, 2026

CHRISTOPHER JABAGAT
IT Officer – I
SDO Bislig City

Dear Sir Jabagat,

Good day.

I am currently conducting a graduate research study entitled **"Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao del Sur."** As part of this research, I have developed a **Sentiment Analysis Validation Tool** which includes the Ground-Truth Check for sentiment labeling, construction of the Confusion Matrix, and computation of Accuracy, Precision, Recall, and F1-Score for evaluating the performance of a BERT-based sentiment classification model.

In this connection, I respectfully seek your expertise as a Subject Matter Expert to review and validate the attached instrument. Your evaluation will help ensure that the tool is clear, relevant, statistically sound, and appropriate for validating customer feedback classification results. Specifically, your assessment will focus on the clarity of instructions, correctness of the definitions of True Positive, True Negative, False Positive, and False Negative, accuracy of the formulas used, and overall technical soundness of the computation procedures.

Attached to this letter are the following documents for your review:

1. Sentiment Analysis Validation Tool
2. Subject Matter Expert Validation Sheet

Your comments, recommendations, and validation decision will greatly contribute to strengthening the methodological rigor and reliability of this study.

Thank you very much for your valuable time, expertise, and support.

Respectfully yours,

MARVIN S. MANGULLAH
Researcher

Approved by:

CHRISTOPHER JABAGAT

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

SUBJECT MATTER EXPERT VALIDATION SHEET

Study Title:

AUTOMATED ANALYSIS OF DATA AGGREGATION AND REPORTING OF CLIENT SATISFACTION MEASUREMENT IN DEPED SURIGAO DEL SUR

Instrument Being Validated:

Sentiment Analysis Validation Tool (Ground Truth Check, Confusion Matrix, and Metrics Computation)

Validator's Information

Name: CHRISTOPHER T. JABAGAT
Area of Expertise: [illegible]
Position/Designation: IT OFFICER I
Institution/Agency: DEPED TANDAG CITY DIVISION
Years of Experience: 10 years
Email (Optional):
Date:

Rating Scale

4 – Very Appropriate
3 – Appropriate
2 – Needs Minor Revision
1 – Needs Major Revision

A. Clarity and Structure of the Instrument

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The instructions are clear and easy to understand.		✓		
The sentiment definitions (TP, TN, FP, FN) are clearly explained.		✓		
The confusion matrix section is logically organized.		✓		
The computation formulas are clearly presented.		✓		
The tool follows a logical sequence (labeling → matrix → metrics).		✓		

B. Relevance

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The tool properly validates sentiment classification using BERT.		✓		
The tool allows accurate encoding of confusion matrix values.		✓		
The formulas for Accuracy, Precision, Recall, and F1-Score are correct.		✓		
The worked sample computation is mathematically correct.		✓		
The tool is appropriate for validating CSM open-ended responses.	✓			

C. Technical and Statistical Soundness

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The definition of TP, TN, FP, FN is statistically accurate.		✓		
The one-vs-rest computation for three-class sentiment is appropriate.		✓		
The formulas for macro and weighted averages are appropriate.		✓		
The instrument can generate reliable evaluation metrics.		✓		

Overall Evaluation

☐ Highly Valid
☐ Valid
☐ Valid with Minor Revisions
☐ Needs Major Revision

Comments and Recommendations

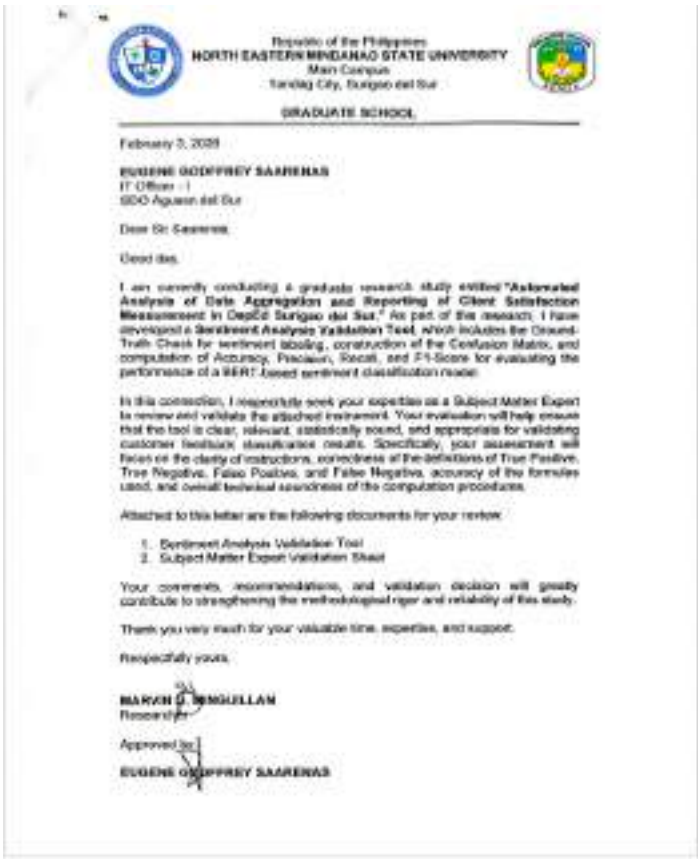
Suitable method, Dependable data source

Validation Decision

☑ The instrument is approved for pilot testing.
☐ The instrument is approved subject to minor revisions.
☐ The instrument requires major revision before use.

Validator Signature:
Date:

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615



Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

February 3, 2026

EUGENE GODFFREY SAARENAS
IT Officer - I
SDO Agusan del Sur

Dear Sir Saarenas,

Good day.

I am currently conducting a graduate research study entitled **"Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao del Sur."** As part of the research, I have developed a **Sentiment Analysis Validation Tool**, which includes the Ground-Truth Check for sentiment labeling, construction of the Confusion Matrix, and computation of Accuracy, Precision, Recall, and F1-Score for evaluating the performance of a BERT-based sentiment classification model.

In this connection, I respectfully seek your expertise as a Subject Matter Expert to review and validate the attached instrument. Your evaluation will help ensure that the tool is clear, relevant, statistically sound, and appropriate for validating customer feedback classification results. Specifically, your assessment will focus on the clarity of instructions, correctness of the definitions of True Positive, True Negative, False Positive, and False Negative, accuracy of the formulas used, and overall technical soundness of the computation procedures.

Attached to this letter are the following documents for your review:

1. Sentiment Analysis Validation Tool
2. Subject Matter Expert Validation Sheet

Your comments, recommendations, and validation decision will greatly contribute to strengthening the methodological rigor and reliability of this study.

Thank you very much for your valuable time, expertise, and support.

Respectfully yours,

MARVIN J. BINGULLAN
Researcher

Approved by:

EUGENE GODFFREY SAARENAS

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

SUBJECT MATTER EXPERT VALIDATION SHEET

Study Title:

AUTOMATED ANALYSIS OF DATA AGGREGATION AND REPORTING OF CLIENT SATISFACTION MEASUREMENT IN DEPED SURIGAO DEL SUR

Instrument Being Validated:

Sentiment Analysis Validation Tool (Ground Truth Check, Confusion Matrix, and Metrics Computation)

Validator's Information

Name: Eugene Godfrey P. Saarenas
Area of Expertise: Information Technology
Position/Designation: IT Officer I
Validation Agency: DepEd
Years of Experience: 14 years
Email (Optional): eugene.saarenas@deped.gov.ph
Date: 3/3/26

Rating Scale

4 – Very Appropriate
3 – Appropriate
2 – Needs Minor Revision
1 – Needs Major Revision

A. Clarity and Structure of the Instrument

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The instructions are clear and easy to understand.	✓			
The sentiment definitions (TP, TN, FP, FN) are clearly explained.	✓			
The confusion matrix section is logically organized.	✓			
The computation formulas are clearly presented.	✓			
The tool follows a logical sequence (labeling → matrix → metrics).	✓			

B. Relevance

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The tool properly validates sentiment classification using BERT.	✓			
The tool allows accurate encoding of confusion matrix values.	✓			
The formulas for Accuracy, Precision, Recall, and F1-Score are correct.	✓			
The worked example computation is mathematically correct.	✓			
The tool is appropriate for validating CSM open-ended responses.	✓			

C. Technical and Statistical Soundness

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The definition of TP, TN, FP, FN is statistically accurate.	✓			
The one-vs-rest computation for three-class sentiment is appropriate.	✓			
The formulas for macro and weighted averages are appropriate.	✓			
The instrument can generate reliable evaluation metrics.	✓			

Overall Evaluation

☑ Highly Valid
☐ Valid
☐ Valid with Minor Revisions
☐ Needs Major Revision

Comments and Recommendations

The instrument properly validates the accuracy of the sentiment analysis tool.

Validation Decision

☐ The instrument is approved for pilot testing.
☐ The instrument is approved subject to minor revisions.
☐ The instrument requires major revision before use.

Validator Signature: _____
Date: _____

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

February 3, 2026

KRISTIAN DIOR LAYNO
IT Officer I
SDO Cabadbaran City

Dear Sir Layno,

Good day.

I am currently conducting a graduate research study entitled "**Automated Analysis of Data Aggregation and Reporting of Client Satisfaction Measurement in DepEd Surigao del Sur.**" As part of this research, I have developed a **Sentiment Analysis Validation Tool**, which includes the Ground-Truth Check for sentiment labeling, construction of the Confusion Matrix, and computation of Accuracy, Precision, Recall, and F1-Score for evaluating the performance of a BERT-based sentiment classification model.

In this connection, I respectfully seek your expertise as a Subject Matter Expert to review and validate the attached instrument. Your evaluation will help ensure that the tool is clear, relevant, statistically sound, and appropriate for validating customer feedback classification results. Specifically, your assessment will focus on the clarity of instructions, correctness of the definitions of True Positive, True Negative, False Positive, and False Negative, accuracy of the formulas used, and overall technical soundness of the computation procedures.

Attached to this letter are the following documents for your review:

1. Sentiment Analysis Validation Tool
2. Subject Matter Expert Validation Sheet

Your comments, recommendations, and validation decision will greatly contribute to strengthening the methodological rigor and reliability of this study.

Thank you very much for your valuable time, expertise, and support.

Respectfully yours,

MARVIN C. MINGUILLAN
Researcher

Approved by:

KRISTIAN DIOR LAYNO

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

SUBJECT MATTER EXPERT VALIDATION SHEET

Study Title

AUTOMATED ANALYSIS OF DATA AGGREGATION AND REPORTING OF CLIENT SATISFACTION MEASUREMENT IN DEPED SURIGAO DEL SUR

Instrument Being Validated

Sentiment Analysis Validation Tool (Ground Truth Check, Confusion Matrix, and Metrics Computation)

Validator's Information

Name: KRISTIAN DIOR A. LAYNO
Area of Expertise: IT Practitioner
Position/Designation: IT Officer I
Institution/Agency: DepEd – Cabadbaran City
Years of Experience: 10 years
Email (Optional): kristiandior.layno@deped.gov.ph
Date:

Rating Scale

4 – Very Appropriate
3 – Appropriate
2 – Needs Minor Revision
1 – Needs Major Revision

A. Clarity and Structure of the Instrument

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The instructions are clear and easy to understand.	✓			
The sentiment definitions (TP, TN, FP, FN) are clearly explained.	✓			

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

The confusion matrix section is logically organized.	✓			
The computation formulas are clearly presented.	✓			
The tool follows a logical sequence (labeling → matrix → metrics).	✓			

B. Relevance

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The tool properly validates sentiment classification using BERT.	✓			
The tool allows accurate encoding of confusion matrix values.	✓			
The formulas for Accuracy, Precision, Recall, and F1-Score are correct.	✓			
The worked example computation is mathematically correct.	✓			
The tool is appropriate for validating CSM open-ended responses.	✓			

Republic of the Philippines
NORTH EASTERN MINDANAO STATE UNIVERSITY
Main Campus
Tandag City, Surigao del Sur

GRADUATE SCHOOL

C. Technical and Statistical Soundness

Indicator	Very Appropriate (4)	Appropriate (3)	Needs Minor Revision (2)	Needs Major Revision (1)
The definition of TP, TN, FP, FN is statistically accurate.	✓			
The one-vs-rest computation for three-class sentiment is appropriate.	✓			
The formulas for macro and weighted averages are appropriate.	✓			
The instrument can generate reliable evaluation metrics.	✓			

Overall Evaluation

☑ Highly Valid
☐ Valid
☐ Valid with Minor Revisions
☐ Needs Major Revision

Comments and Recommendations

Content Verification

Validation Decision

☑ The instrument is approved for pilot testing.
☐ The instrument is approved subject to minor revisions.
☐ The instrument requires major revision before use.

Validator Signature: _____
Date: _____

VALIDATION SHEET

Name of Evaluator: Pierce Andre Abarro
Degree: BSIT
Position: Information Systems Analyst I
Number of Years in Teaching/Service: 2 years

To the Evaluator: Please check the appropriate box for your rating:

Point Equivalent:
5 -Excellent
4 -Very Good
3 -Good
2 -Fair
1 -Poor

1.) Clarify of Direction and Items	5	4	3	2	1
The vocabulary level, language structure and conceptual of the questions suit the level of respondents. The test direction and items are written in a clear and understandable manner.	✓				
2.) Presentation/Organization of Items					
The items are organized in logical manner.		✓			
3.) Suitability of Items					
The items appropriately represent the substance of the research. The questions are designed to determine the conditions, knowledge, skills and attitudes that are supposed to be measured.		✓			
4. Adequateness of Items per Category					
The items represent the coverage of the research adequately. The number of questions per category is representative enough of all questions needed for the research.		✓			
5.) Attainment of the Purpose					
The instrument as a whole fulfills the objective for which it was constructed.	✓				
6.) Objectivity					
In each item question, there is only one specific answer to measure one behavior and no aspect of the questionnaire suggests bias on the part of the research.	✓				
7.) Scale and Evaluation in Rating System					
The scale adopted is appropriate for the items.	✓				

Remarks:

PIERCE ANDRE ABARRO
Signature over Printed Name

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX K:
SAMPLE ANSWERED SURVEY QUESTIONNAIRES

Republic of the Philippines
NORTHEASTERN MINDANAO STATE UNIVERSITY
Main Campus
Rosario, Tandag City, Surigao del Sur

GRADUATE SCHOOL

APPENDIX A
Survey Questionnaire

Dear Respondents,
Please complete this survey questionnaire form. Rest assured that the information will be kept confidential. Thank you.

MARVING MINGUILLAN
Researcher

I. PERSONAL DETAILS
Name: [illegible]

II. EDUCATIONAL ATTAINMENT (Please check the one that applies)
_____ Elementary Graduate
_____ High School Graduate
_____ Vocational Graduate
___/___ College Level
___/___ College Graduate
_____ Masteral
_____ Doctorate

III. OCCUPATION AND SKILLS (Please check one that applies)
_____ House owner
_____ Business owner
_____ Other. Please Specify
___/___ Government Employee
_____ IT Expert

Do you know how to use a mobile phone? ___/___ Yes _____ No

Evaluate the developed *Wine Leak Detection System using Sensor Technology* using the ISO 25010 software standard.

5 Very Great Extent
4 Great Extent
3 Moderate Extent
2 Little Extent
1 Very Little Extent

121

Republic of the Philippines
NORTHEASTERN MINDANAO STATE UNIVERSITY
Main Campus
Rosario, Tandag City, Surigao del Sur

GRADUATE SCHOOL

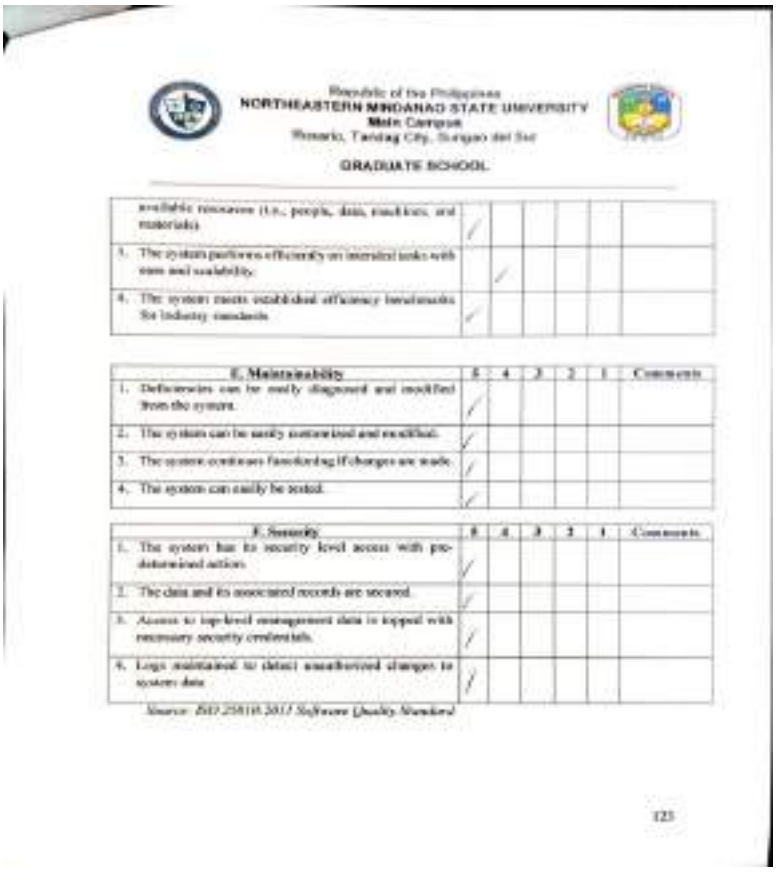
A. Functionality	5	4	3	2	1	Comments
1. The required task is performed by the system satisfactorily	/					
2. The output generated by the system is correct			/			
3. The system helps the user overcome any problem encountered	/					
4. The system regulates user access to data	/					

B. Reliability	5	4	3	2	1	Comments
1. Most errors have been corrected over time.		/				
2. The system is capable of maintaining its performance despite the presence of errors during operation.	/					
3. The system restores the operation of lost data after failure.	/					
4. The system provides real-time monitoring to ensure availability.	/					

C. Usability	5	4	3	2	1	Comments
1. The system is intuitive and receptive in learning	/					
2. The information generated by the system is easy to understand and act upon.	/					
3. The documentation is very informative.		/				
4. The system straightforwardly performs tasks	/					

D. Efficiency	5	4	3	2	1	Comments
1. The system responds immediately to sensor inputs.	/					
2. The system maximizes the efficient utilization of						

122



Republic of the Philippines
NORTHEASTERN MINDANAO STATE UNIVERSITY
Main Campus
Rosario, Tandag City, Surigao del Sur

GRADUATE SCHOOL

available resources (i.e., people, data, machines, and materials)	/					
3. The system performs efficiently on intended tasks with ease and usability.		/				
4. The system meets established efficiency benchmarks for industry standards	/					

E. Maintainability	5	4	3	2	1	Comments
1. Deficiencies can be easily diagnosed and modified from the system.	/					
2. The system can be easily customized and modified.	/					
3. The system continues functioning if changes are made.	/					
4. The system can easily be tested.	/					

F. Security	5	4	3	2	1	Comments
1. The system has its security level access with pre-determined action.	/					
2. The data and its associated records are secured.	/					
3. Access to top-level management data is topped with necessary security credentials.	/					
4. Logs maintained to detect unauthorized changes to system data	/					

Source: ISO 25010:2011 Software Quality Standard

123

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

Republic of the Philippines
NORTHEASTERN MINDANAO STATE UNIVERSITY
Main Campus
Rosario, Tandag City, Surigao del Sur

GRADUATE SCHOOL

APPENDIX A
Survey Questionnaire

Dear Respondents,

Please complete this survey questionnaire form. Rest assured that the information will be kept confidential. Thank you.

MARVIN G. MINGUILLAN
Researcher

I. PERSONAL DETAILS
Name: Anthony Llenos

II. EDUCATIONAL ATTAINMENT (Please check the one that applies)
_____ Elementary Graduate
_____ High School Graduate
_____ Vocational Graduate
_____ College Level
__/__ College Graduate
_____ Masteral
_____ Doctorate

III. OCCUPATION AND SKILLS (Please check one that applies)
_____ Home owner
_____ Business owner
_____ Other. Please Specify
__/__ Government Employee
_____ IT Expert

Do you know how to use a mobile phone? __/__ Yes _____ No

Evaluate the developed *Mean Look Detection System using Sensor Technology* using the ISO 25010 software standard.

5 Very Great Extent
4 Great Extent
3 Moderate Extent
2 Little Extent
1 Very Little Extent

131

Republic of the Philippines
NORTHEASTERN MINDANAO STATE UNIVERSITY
Main Campus
Rosario, Tandag City, Surigao del Sur

GRADUATE SCHOOL

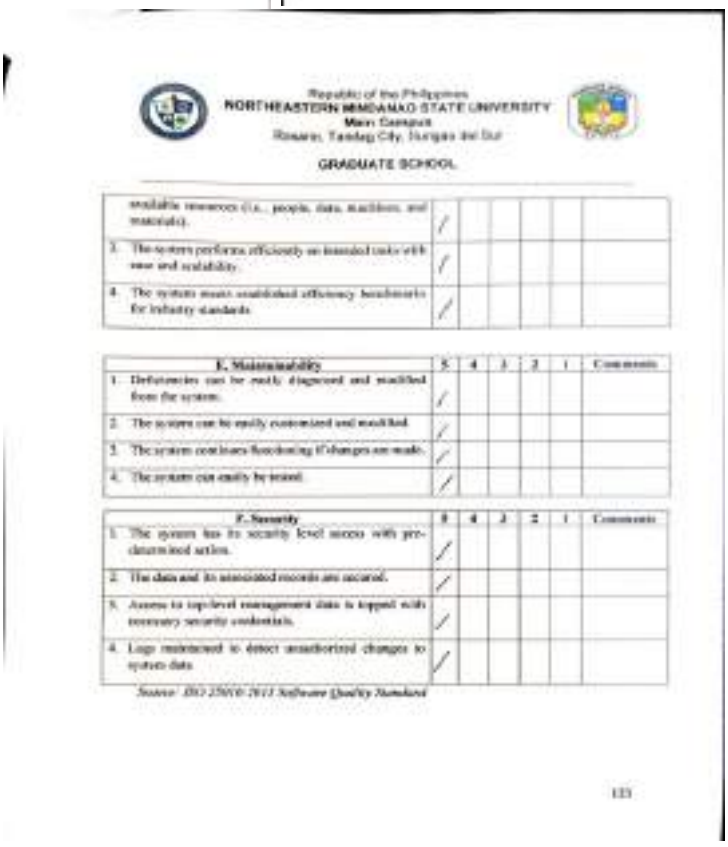
A. Functionality	5	4	3	2	1	Comments
1. The required task is performed by the system satisfactorily.	/					
2. The output generated by the system is correct.	/					
3. The system helps the user overcome any problem encountered.		/				
4. The system regulates user access to data.	/					

B. Reliability	5	4	3	2	1	Comments
1. Most errors have been corrected over time.	/					
2. The system is capable of maintaining its performance despite the presence of errors during operation.	/					
3. The system restores the operation of lost data after failure.	/					
4. The system provides real-time monitoring to ensure availability.	/					

C. Usability	5	4	3	2	1	Comments
1. The system is intuitive and responsive to learning.	/					
2. The information generated by the system is easy to understand and act upon.	/					
3. The documentation is very informative.	/					
4. The system straightforwardly performs tasks.	/					

D. Efficiency	5	4	3	2	1	Comments
1. The system responds immediately to sensor inputs.	/					
2. The system maximizes the efficient utilization of						

122



Republic of the Philippines
NORTHEASTERN MINDANAO STATE UNIVERSITY
Main Campus
Rosario, Tandag City, Surigao del Sur

GRADUATE SCHOOL

available resources (i.e., people, data, machines, and materials).	/					
3. The system performs efficiently as intended tasks with ease and reliability.	/					
4. The system meets established efficiency benchmarks for industry standards.	/					

E. Maintainability	5	4	3	2	1	Comments
1. Deficiencies can be easily diagnosed and modified from the system.	/					
2. The system can be easily customized and modified.	/					
3. The system continues functioning if changes are made.	/					
4. The system can easily be tested.	/					

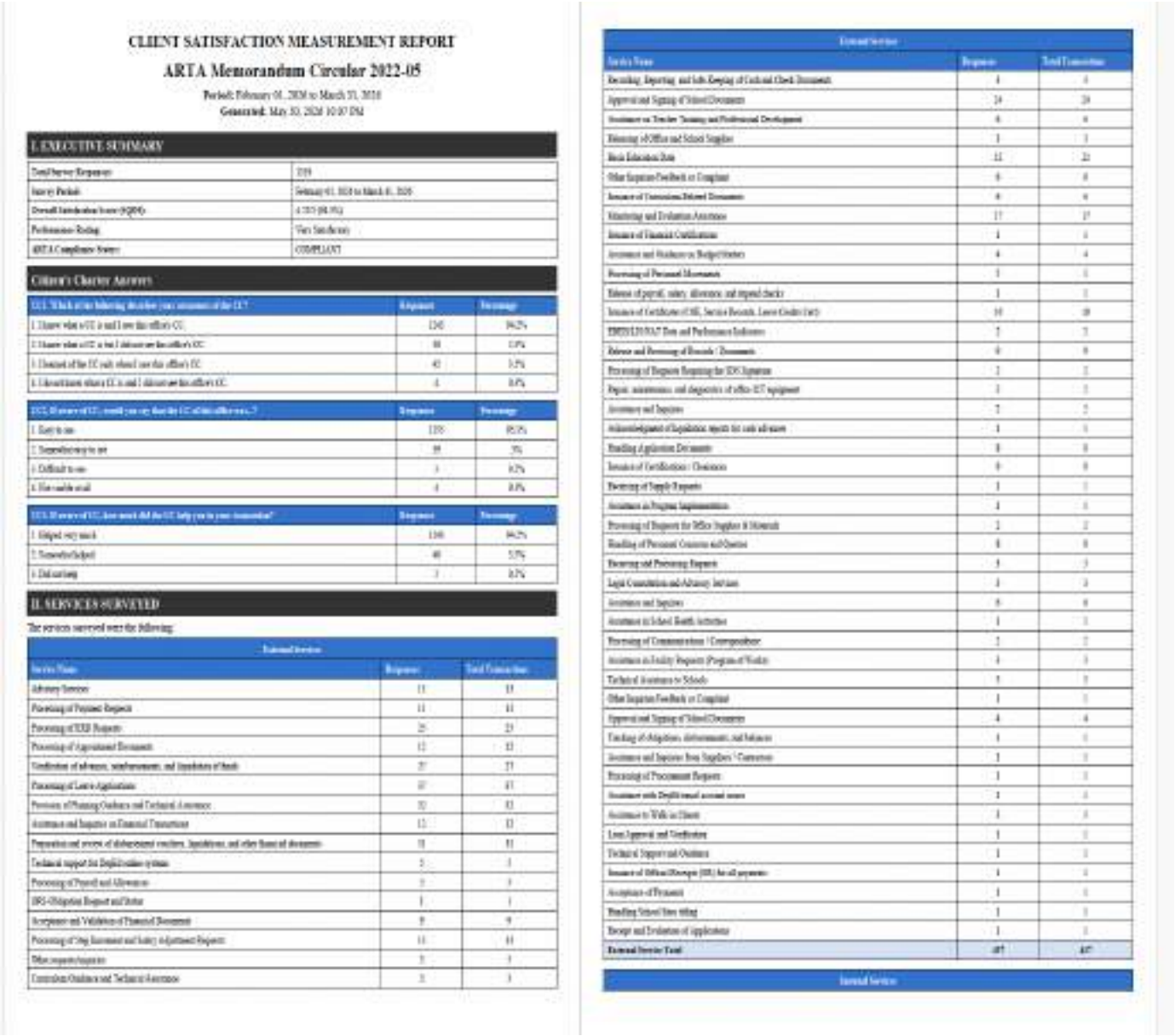
F. Security	5	4	3	2	1	Comments
1. The system has its security level access with pre-determined action.	/					
2. The data and its associated records are secured.	/					
3. Access to top-level management data is topped with necessary security credentials.	/					
4. Logs maintained to detect unauthorized changes to system data.	/					

Source: ISO 25010:2011 Software Quality Standard

123

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615

APPENDIX L:
SAMPLE PDF ARTA REPORT RESULT



CLIENT SATISFACTION MEASUREMENT REPORT

ARTA Memorandum Circular 2022-05

Period: February 01, 2026 to March 31, 2026
Generated: May 30, 2026 10:07 PM

I. EXECUTIVE SUMMARY

Total Survey Responses	239
Survey Period	February 01, 2026 to March 31, 2026
Overall Satisfaction Score (SQD0)	4.333 (86.9%)
Performance Rating	Very Satisfactory
ARTA Compliance Status	COMPLIANT

Citizen's Charter Answers

CC1: Which of the following describes your awareness of the CC?	Response	Percentage
1. I know what a CC is and I saw this office's CC.	226	94.2%
2. I know what a CC is but I did not see this office's CC.	10	4.0%
3. I learned of the CC only when I saw this office's CC.	03	1.2%
4. I do not know what a CC is and I did not see this office's CC.	1	0.4%

CC2: If aware of CC, would you say that the CC of this office was...?	Response	Percentage
1. Easy to see	229	95.8%
2. Somewhat easy to see	09	3%
3. Difficult to see	1	0.2%
4. Not visible at all	1	0.4%

CC3: If aware of CC, how much did the CC help you in your transaction?	Response	Percentage
1. Helped very much	230	96.2%
2. Somewhat helpful	08	3.3%
3. Did not help	1	0.2%

II. SERVICES SURVEYED

The services surveyed were the following:

External Service		
Service Name	Response	Total Transaction
Advisory Services	11	11
Processing of Payment Requests	11	11
Processing of TESF Requests	25	25
Processing of Appointment Documents	12	12
Coordination of advocacies, reimbursements, and liquidations of funds	27	27
Processing of Leave Applications	17	17
Provision of Planning Guidance and Technical Assistance	32	32
Assistance and Inquiries on Financial Transactions	12	12
Preparation and review of disbursement vouchers, liquidations, and other financial documents	11	11
Technical Support for DepEd online systems	3	3
Processing of Travel and Allowances	3	3
DRS-Obligation Request and Status	1	1
Acceptance and Validation of Financial Documents	9	9
Processing of Step Increment and Salary Adjustment Requests	11	11
Other requests/inquiries	3	3
Curriculum Guidance and Technical Assistance	3	3

External Service		
Service Name	Response	Total Transaction
Recording, Reporting, and Safe Keeping of Cash and Check Documents	1	1
Approval and Signing of School Documents	24	24
Assistance on Teacher Training and Professional Development	6	6
Ordering of Office and School Supplies	1	1
Basic Education Data	11	21
Other Inquiries Feedback or Complaint	0	0
Issuance of Commission Related Documents	6	6
Monitoring and Evaluation Assistance	17	17
Issuance of Financial Certifications	1	1
Assistance and Guidance on Budget Matters	4	4
Processing of Personnel Movements	1	1
Release of payroll, salary, allowance, and stipend checks	1	1
Issuance of Certificates of GSIS, Service Records, Leave Credits Data	10	10
[illegible] Data and Performance Indicators	7	7
Release and Processing of Records / Documents	0	0
Processing of Requests Requiring the DSO Signature	2	2
Repair, maintenance, and diagnostics of office ICT equipment	1	1
Assistance and Inquiries	7	7
Acknowledgement of Legislative requests for cash advances	1	1
Handling of application Documents	0	0
Issuance of Certifications / Clearances	0	0
Processing of Supply Requests	1	1
Assistance on Program Implementation	1	1
Processing of Requests for Office Supplies & Materials	2	2
Handling of Personnel Concerns and Queries	0	0
Processing and Processing Requests	3	3
Legal Consultations and Advocacy Services	3	3
Assistance and Inquiries	0	0
Assistance in School Health Activities	1	1
Processing of Communications / Correspondence	2	2
Assistance on Facility Requests (Programs of Works)	3	3
Technical Assistance to Schools	3	3
Other Inquiries Feedback or Complaint	1	1
Approval and Signing of School Documents	4	4
Tracking of Allegations, disbursements, and balances	1	1
Assistance and Inquiries from Suppliers / Contractors	2	2
Processing of Procurement Requests	1	1
Assistance with DepEd email account issues	1	1
Assistance to Walk-in Clients	1	1
Lease Approval and Verification	1	1
Technical Support and Guidance	1	1
Issuance of Official Receipts (ORs) for all payments	1	1
Acceptance of Documents	1	1
Handling School Data Gathering	1	1
Receipt and Evaluation of Applications	1	1
External Service Total	187	187

Internal Service

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26jun615