ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24FEB1647

Development of Faculty Qualification Analysis System using Naive Bayes Algorithm

Eliza B. Ayo[1]; Lani E. Sakay[2]
Centro Escolar University, Manila, Philippinnes

Anna Liza O. Villanueva[3]; Rosalina R. Pangilinan[4]
Perpetual Help College-Manila, Manila, Philippines

Celia L.Verano[5]
Manila High School, Manila, Philippines

Abstract:- Qualification prediction is a crucial process in determining whether an applicant is qualified for a particular position. However, traditional methods of evaluation often rely on the experience and intuition of the evaluator, which may not always be accurate. This study proposed the use of a supervised machine learning approach, specifically the Naïve Bayes algorithm, to predict faculty qualification based on a labeled dataset. The developed Faculty Qualification Analysis System for Perpetual Help College of Manila would allow users to input appropriate test data and generate results of qualified or not qualified. The system's effectiveness and acceptance had 4.3 and 4.4 ratings with verbal interpretation of very high and strongly acceptable. The results of this study demonstrated the potential of machine learning algorithms to improve the accuracy and efficiency of qualification prediction processes in educational institutions.

Keywords:- *Qualification Prediction, Machine Learning, Naïve Bayes Algorithm.*

I. INTRODUCTION

As educational institutions encounter heightened competition in attracting top-tier faculty members, Human Resources (HR) departments face the imperative of implementing efficient screening procedures to identify the most qualified candidates. Perpetual Help College of Manila is no exception. Its Human resource department must scrutinize applicants' credentials to ascertain compliance with the college's requirements before proceeding to demo teaching sessions. Nevertheless, the prevailing method of manually reviewing printed curriculum vitae can prove time-consuming and could introduce conflicts or errors during the assessment process. To overcome this challenge, this study is anchored on the development of a Faculty Qualification Screening system that uses machine learning and predictive analytics. This innovation can free HR personnel to do responsibilities and devote their time to other important tasks.

In the case of Perpetual Help College- Manila, before a faculty member conducts a demonstration of teaching skills, the Human Resource Department needs to assess whether the applicant's credentials align with the university's requirements. Collaboration between the Human Resource Department and the Dean is integral in the evaluation process to determine the eligibility of candidates for demo teaching. The current practice involves a manual review of printed curriculum vitae, which can be time-consuming as both parties need to thoroughly read through all the criteria. This manual approach occasionally leads to conflicts arising from differing interpretations of standards or variations in awareness of specific requirements. For some institutions, the search committee's faculty votes are more influential in tenure-track faculty selection decisions than the chair or dean's votes, with academic accomplishments, interview performance, and presentation skills being key determinants [22]. Others opted for AI adoption in their recruitment process because it leads to efficiency and qualitative gains for both clients and candidates, offering strategic insights into automation and AI implementation [23]. These technologies have the potential to save time and resources while improving the accuracy and efficiency of recruitment. In this study, the naive Bayes algorithm improves the traditional method by selecting a subset of attributes, enhancing classification accuracy while reducing computational overhead [1].

Similar to this is the use of Artificial intelligence in Indian software companies which positively impacts the recruitment process resulting in effective talent acquisition and sustainable development [24]. The study is expected to help organizations formulate recruitment strategies and policy interferences to align to develop its effective recruitment process to recruit qualified talent into the team to encounter competitive business and to develop a sustainable environment. Moreover, several educational institutions also have adopted predictive algorithm-based screening systems like, a college faculty recruitment system [25], an android application that helps in sorting out candidates and providing the best results in the recruitment cycle, and a comprehensive campus recruitment and placement system for optimizing the hiring process [26].

With the mentioned successful implementation, a shift in paradigm to ditch traditional methods of faculty recruitment process is necessary. Adopting a machine learning and predictive algorithm-based screening system can streamline processes, and save time, to ensure HR hires qualified candidates is necessary. With the help of machine learning using the naïve Bayes algorithm, automated faculty recruitment is possible. The developed systems allow greater faculty diversity to create a richer learning environment and ensure continuity to be able to recruit top talent which improves the overall quality of education provided at the institution.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24FEB1647

II. STATEMENT OF THE PROBLEM

- *This Study was Pursued to Answer the Following Problems:*

- *What are the challenges encountered in the existing screening process for faculty applicants to be qualified for teaching demonstrations?*
- *How is the Naive Bayes Algorithm used in the data sets to predict qualified faculty applicants?*
- *How does the developed prototype address the challenges found in the existing system?*
- *How do the I.T. experts assess the level of effectiveness* of the developed prototype in terms of ISO 25010 criteria

- ✓ *Functional – Correctness;*
- ✓ *Efficiency/Performance -Software Capacity;*
- ✓ *Usability – user error protection;*
- ✓ *Reliability – fault tolerance and recoverability;*
- ✓ *Compatibility;*
- ✓ *Security;*
- ✓ *Maintainability; and*
- ✓ *Portability?*

- *How do the users assess the acceptance of the developed prototype in terms of ISO 25010 criteria:*

- ✓ *Functional suitability – Completeness and Appropriateness;*
- ✓ *Time-behavior and resource utilization;*
- ✓ *Appropriateness, learnability, operability, and user interface aesthetics; and lastly*
- ✓ *Maturity and availability?*

III. THEORETICAL FRAMEWORK

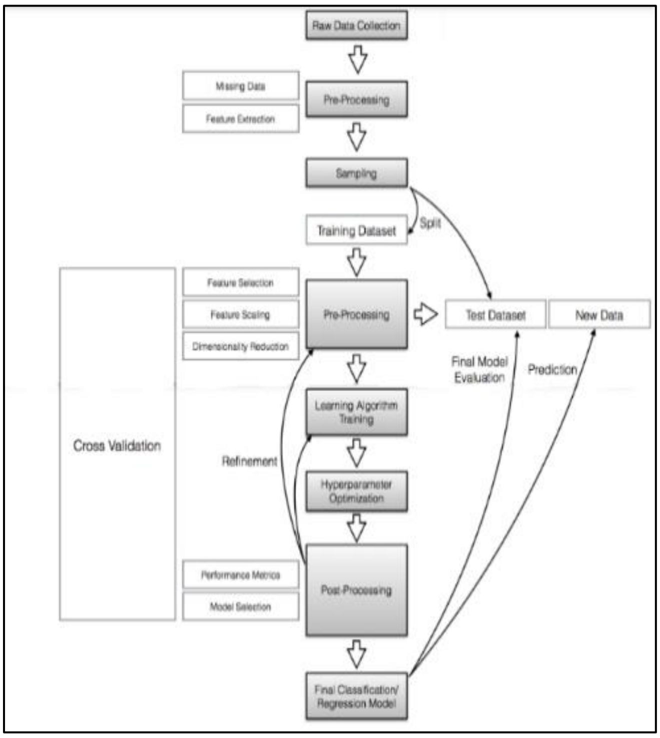


Fig 1 Supervised Machine Learning Process Model

This study employs the predictive modeling theory as illustrated in Figure 1, with a specific focus on supervised machine learning and the Naïve Bayes algorithm. The process begins with data collection to gather pertinent information for the subsequent modeling stages. The feature selection follows, involving the identification of crucial variables that impact the model's outcome. Subsequently, an appropriate model is selected to ensure accurate predictions based on the chosen features. The model is then trained using a designated training set, and its performance is evaluated through carefully selected test data and diverse metrics, parameters, feature selection, or preprocessing techniques to enhance its efficacy and optimization [27]. Once optimized, the model is ready for deployment. This comprehensive workflow diagram provides a systematic framework for constructing and implementing a supervised machine learning model for predictive purposes.

- *Assumptions of the Study*

The current process for determining qualified faculty applicants for teaching demonstrations has deficiencies and inefficiencies that can be addressed through the developed system.

- *Scope and Delimitation*

The study focuses on the development and evaluation of a faculty qualification analysis system utilizing the Naïve Bayes algorithm at the College of Computer Studies-Perpetual Help College of Manila. The key areas of emphasis include the screening process employed by the HR department for selecting faculty for teaching demonstrations, creating of training data set, designing the system's database and architecture, evaluating the system's effectiveness, and assessing user acceptance. Several limitations are inherent in this study. Firstly, the use of historical data from the Perpetual Help College of Manila and a Kaggle dataset may affect the results to other institutions or disciplines when implemented. The study does not evaluate hardware equipment for system implementation, and the absence of an exploration of ethical implications associated with machine learning in faculty selection, are further limitations. Despite these constraints, the study contributes valuable insights into the development and implementation of a faculty qualification screening system using the Naïve Bayes algorithm at Perpetual Help College of Manila, with a need for cautious interpretation due to the identified limitations.

- *Significance of the Study*

This study significantly impacts the faculty selection process within academic institutions, aimed to enhance efficiency, accuracy, and the selection of qualified candidates, ultimately fostering excellence in academia through the development of faculty qualification based on the Naïve Bayes Algorithm. This streamlined approach holds the potential to optimize resource allocation, make informed decisions about faculty appointments, and improve the overall effectiveness of the hiring process. Future researchers can draw upon the study's findings, methodologies, and challenges, using them as a stepping stone to contribute further to the field.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24FEB1647

IV. RESEARCH METHODOLOGY

The research design and methodology employed in the study used the descriptive developmental method. The descriptive method's primary focus is to describe, compare, analyze, and interpret existing data, aligning with the study's objective of identifying criteria for predicting the faculty qualification system and evaluating the software results. The development Faculty Qualification Analysis System using the Naïve Bayes Algorithm aimed to develop and implement an efficient system for analyzing the credentials of faculty applicants for demo teaching. The design and methodology involved several key phases, data collection and preparation, training data selection, algorithm development, system development tools, evaluation and testing, and project deliverables

The project commenced with data collection and preparation. The researcher gathered both primary and secondary data, showcasing the current business process of faculty screening of Perpetual Help College of Manila. A proposed system was designed based on the representation of the data flow of the current system. This design aimed at creating a more efficient and faster system, objectively reducing processes and generating results. To enable the system to predict qualified faculty for teaching demonstration, the researcher identified a suitable training data set. The Kaggle data set, considered a foundational resource in data science, was chosen to provide historical data for the system's machine learning training phase. The study extensively explored and employed the Naïve Bayes algorithm for predicting the probability of faculty qualification. The algorithm development phase involved designing the system, including the database, user interface, and logic design. The features of the Faculty Qualification Analysis System were carefully deliberated based on the expected project output. The development process adhered to a structured and systematic approach based on the Waterfall Model, a well-established model in the System Development Life Cycle (SDLC). Figure 2 provides a visual representation of the various steps undertaken in the development process.

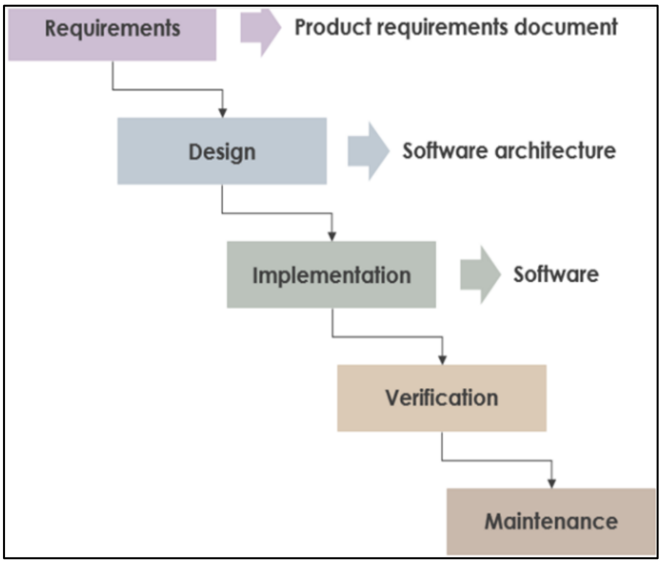


Fig 2 Waterfall Development Method

The project utilized specific software tools for its development, operating on the Windows 10 platform. The primary programming language employed was Python, known for its simplicity, readability, and support for modularity and code reuse. The Spyder IDE, included with Anaconda, served as the integrated development environment for editing, testing, and debugging the Python code. Additionally, CSV files were utilized for data storage.

The user acceptance and IT expert evaluation were integral components of the project design, following the ISO25010 standard. The evaluation considered factors such as the refined Naïve Bayes Algorithm, the system's functionality, and its efficiency in predicting faculty qualification. The project's deliverables encompassed various elements, including data collected from printed curriculum vitae and historical Kaggle datasets, a refined Naïve Bayes Algorithm, a fully functional Faculty Qualification Analysis System, and comprehensive reports on user acceptance and system effectiveness as evaluated by IT experts. These deliverables collectively reflected the successful implementation of the designed project.

➢ *Sources of Data*

The data for this study was sourced from multiple channels, employing a mix of primary and secondary data. The primary sources included interviews and focus group discussions, providing firsthand insights and opinions from relevant stakeholders and among four (4) IT experts, four (4) College Deans, and two (2) HRD staff members. This strategic sampling approach allowed for targeted insights from individuals directly involved in the faculty qualification screening process, ensuring a focused and relevant data collection process. The gathered data from these interactions underwent thorough analysis, leading to the formulation of the study's objectives. The historical data were extracted from the Kaggle dataset, a repository of information for machine learning training while the data from the survey conducted was used to gauge the system's effectiveness and acceptance. Collectively, these diverse sources of data contributed to a holistic and informed approach to conducting the study.

➢ *Statistical Treatment of Data*

To measure the user acceptance and system effectiveness of the Faculty Qualification Analysis System, a weighted mean approach was employed. The eight software characteristics and their sub-characteristics were given equal weights, resulting in the weighted mean being equivalent to the arithmetic mean or sample mean. This allowed for a simple assessment of both factors as seen in table 1.

Table 1 Measurement Ratings - Level of Effectiveness

Assigned Point	Numerical Ranges	Categorical Response	Verbal Interpretation
5	4.51 – 5.00	Strongly Agree	Very High
4	3.51 – 4.50	Agree	High
3	2.51 – 3.50	Neutral	Neutral
2	1.51 – 2.50	Disagree	Low
1	1.00 – 1.50	Strongly Disagree	Very Low

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24FEB1647

The acceptance level of the proposed system was evaluated using ISO/IEC 25010:2011, which utilizes a five-point scale to measure its level of acceptance. This evaluation helps determine how well the system meets the desired criteria, as presented in Table 2.

Table 2 Measurement Ratings - Level of Acceptance

Assigned Point	Numerical Ranges	Categorical Response	Verbal Interpretation
5	4.51 – 5.00	Strongly Agree	Strongly Acceptable
4	3.51 – 4.50	Agree	Acceptable
3	2.51 – 3.50	Neutral	Neutral
2	1.51 – 2.50	Disagree	Unacceptable
1	1.00 – 1.50	Strongly Disagree	Strongly Unacceptable

➢ *Instrumentation and Validation*

The researcher utilized the international standard ISO/IEC 2510:2011 to assess the acceptability of the software. This standard outlines eight key software characteristics, namely efficiency, compatibility, suitability, reliability, maintainability, security, usability, and portability. These characteristics served as the basis for evaluating the software's quality, ensuring that it met recognized industry standards.

➢ *System Architecture*

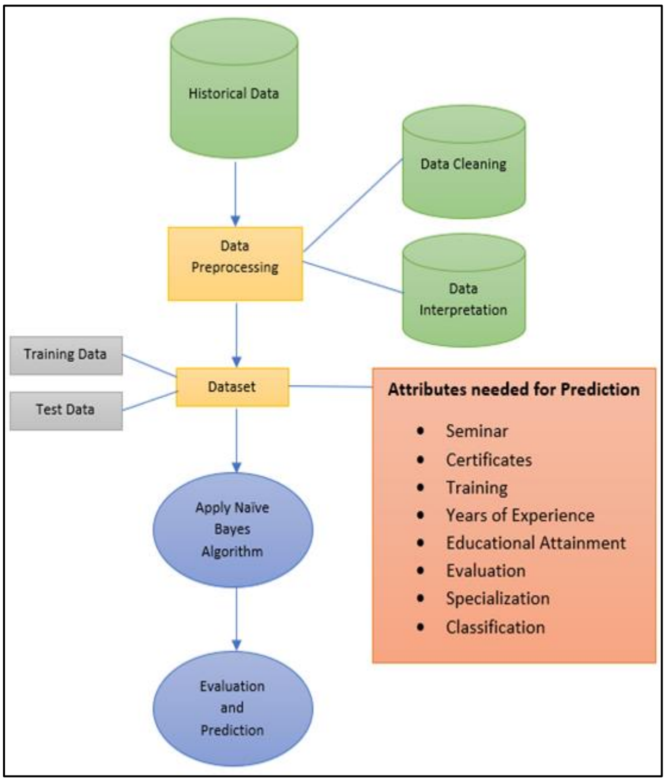


Fig 3 System Architecture

The system architecture employed in this study adheres to a systematic process as seen in Figure 3. The initial step involves acquiring historical data, which forms the foundational basis for the analysis. Following this, the data undergoes a meticulous preprocessing phase that includes both extraction and cleaning processes. Extraction involves the identification and selection of pertinent data for analysis while cleaning entails eliminating redundant or irrelevant information, rectifying errors, and converting the data into a required format. Subsequently, a comprehensive dataset is created, incorporating both training and test data, it served as a model for accurate prediction and analysis. The training dataset plays a crucial role in instructing the machine learning model because it will rely on identified patterns and relationships within the data to recognize trends. Following the model training, an evaluation phase follows, where the accuracy and effectiveness of the model are assessed by testing it on a separate portion of the dataset, specifically the curriculum vitae dataset. This validation step ensures that the system's predictions are aligned with the actual values stored in the database. Finally, the Naïve Bayes algorithm is applied to the prepared and validated dataset, enabling computation and analysis to determine the qualification status of faculty members. This comprehensive system architecture guarantees a methodical and robust process, encompassing data acquisition, preprocessing, dataset creation, and algorithmic utilization for accurate faculty qualification prediction.

➢ *Data Gathering Procedures*

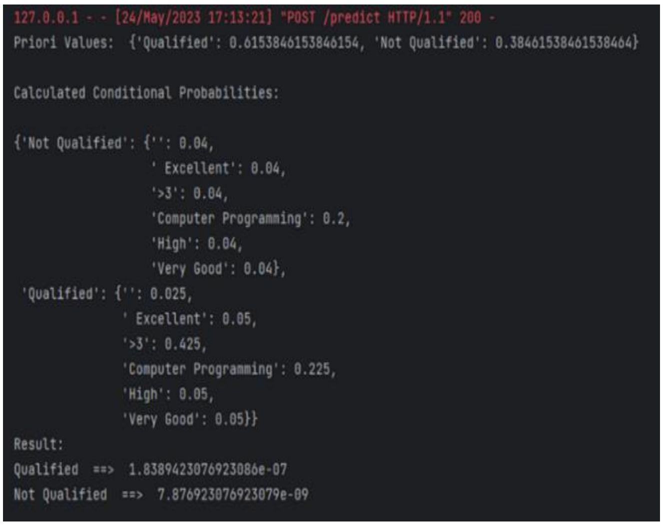
For this study, a comprehensive data-gathering procedure was employed to ensure the acquisition of relevant and valuable information. The process included the collection of data from diverse online sources, such as studies, research papers, documents, and articles, with a specific focus on the proposed Faculty Qualification Analysis System and the application of the Naïve Bayes algorithm, as outlined by Putra et al. in 2020. To assess the effectiveness and user acceptance of the system, a survey method was chosen as the primary data collection tool. The researcher meticulously designed survey questions aligning with the standards set by ISO/IEC 25010:2011, reflecting various software quality characteristics. Subsequently, the researcher personally administered the questionnaires to the identified participants. The survey responses were systematically collected and documented for further analysis. The collected data underwent rigorous statistical analysis to draw meaningful conclusions and insights. This analytical process allowed the researcher to evaluate the system's effectiveness and user acceptance, providing a robust foundation for formulating conclusions and offering recommendations for the study. The combination of online source analysis and survey data collection ensured a well-rounded approach to gathering information for the research study.

➢ *Testing Procedures*

The Black Box Testing for the Faculty Qualification Analysis System was employed. This aimed at evaluating the software's functionality from an end-user perspective without delving into the internal code of the system. The testing process adhered to the system's requirements and expected functionality, focusing on diverse aspects such as the user interface, APIs, database, security, client/server applications, and overall system functionality. Further testing measures were implemented, including Install/Uninstall testing. This comprehensive approach ensured a thorough evaluation of all system components to guarantee their excellent performance.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24FEB1647

* *Qualified Output – Evidence 1*

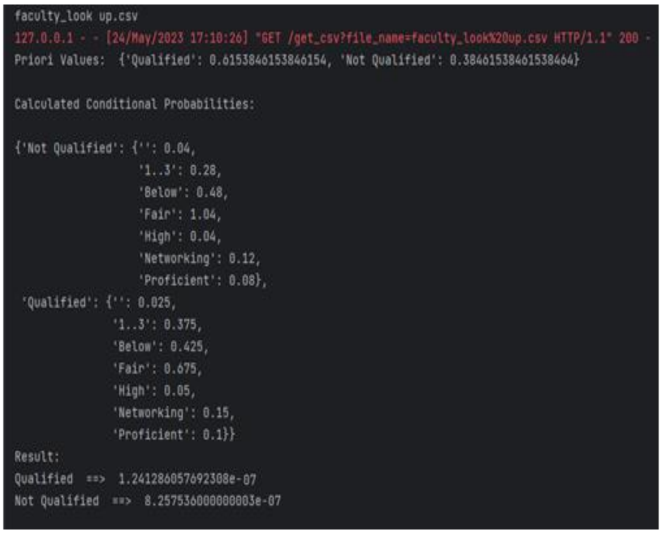


```
127.0.0.1 - - [24/May/2023 17:13:21] "POST /predict HTTP/1.1" 200 -
Priori Values:  {'Qualified': 0.6153846153846154, 'Not Qualified': 0.38461538461538464}

Calculated Conditional Probabilities:

{'Not Qualified': {'': 0.04,
                   ' Excellent': 0.04,
                   '>3': 0.04,
                   'Computer Programming': 0.2,
                   'High': 0.04,
                   'Very Good': 0.04},
 'Qualified': {'': 0.025,
               ' Excellent': 0.05,
               '>3': 0.425,
               'Computer Programming': 0.225,
               'High': 0.05,
               'Very Good': 0.05}}
Result:
Qualified  ==>  1.8389423076923086e-07
Not Qualified  ==>  7.876923076923079e-09
```

* *Not Qualified – Evidence 2*



```
faculty_look up.csv
127.0.0.1 - - [24/May/2023 17:10:26] "GET /get_csv?file_name=faculty_look%20up.csv HTTP/1.1" 200 -
Priori Values:  {'Qualified': 0.6153846153846154, 'Not Qualified': 0.38461538461538464}

Calculated Conditional Probabilities:

{'Not Qualified': {'': 0.04,
                   '1..3': 0.28,
                   'Below': 0.48,
                   'Fair': 1.04,
                   'High': 0.04,
                   'Networking': 0.12,
                   'Proficient': 0.08},
 'Qualified': {'': 0.025,
               '1..3': 0.375,
               'Below': 0.425,
               'Fair': 0.675,
               'High': 0.05,
               'Networking': 0.15,
               'Proficient': 0.1}}
Result:
Qualified  ==>  1.241286057692308e-07
Not Qualified  ==>  8.257536000000003e-07
```

V. ANALYSIS AND FINDINGS

A. The Current Screening Process of the Perpetual Help College of Manila for Faculty Applicants to be Qualified for Teaching Demonstration

The Perpetual Help College of Manila (PHCM) currently conducts teaching demonstrations after the College Dean and the HR Department review the curriculum vitae. These two stakeholders collaborate to assess if the applicant's credentials align with the college's requirements. Then the applicant will proceed to the next steps. However, the manual review of printed curriculum vitae has led to various challenges, including: (1) a delay of 3 days of forwarding CVs to the college dean, attributed to the HR secretary's busy schedule. (2) a 2-day duration for the dean to review CVs due to priority obligations; (3) subjective recommendations on qualification, influenced by biases and subjective judgments; (4) instances of CVs being lost due to mishandling; and (5) potential conflicts arising from considerations of standards, awareness of requirements, or inconsistencies in evaluation criteria. Figure – shows the data flow diagram of this process.

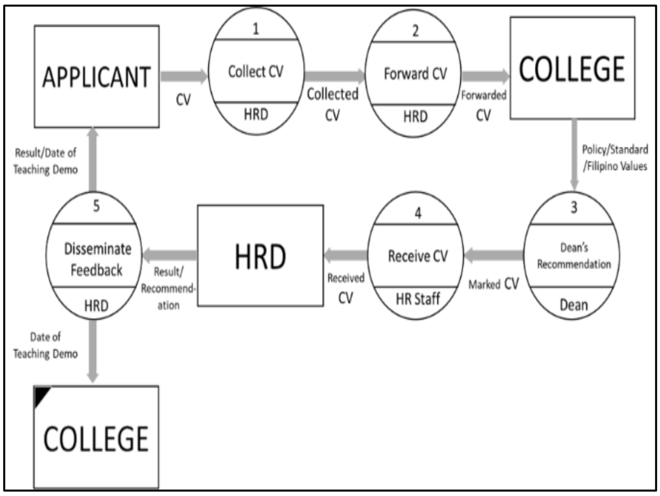


Fig 4 Data Flow Diagram

B. Training Data Set to be Used in Predicting Qualified-Faculty Applicants

The initial stage involves preprocessing the data, wherein relevant information is extracted and cleaned for subsequent analysis. Extraction entails identifying and selecting the necessary data, while cleaning involves eliminating irrelevant or duplicate data, rectifying errors, and formatting the data appropriately for analysis.

Following this, a dataset comprising both training and test data is generated for future use. This dataset serves as a model capable of accurately predicting and analyzing data. The training dataset is employed to teach the machine, relying on patterns and relationships within the data to identify trends. Once trained, the model's accuracy and effectiveness are assessed by testing it on a separate portion of the dataset, specifically the curriculum vitae. This dataset is utilized to validate the system's accuracy in producing predictions by comparing the results to actual values in the database.

Upon dataset creation, the Naïve Bayes algorithm can be applied to compute and analyze the data, determining the qualification of faculty members. The algorithm evaluates and predicts faculty qualification based on the patterns identified during training.

➢ *The Naive Bayes Algorithm*

The Naive Bayes algorithm is a widely employed method in machine learning, particularly for classification tasks. Grounded in Bayes' theorem, it computes the probability of a new data point belonging to a specific class based on its features. This algorithm calculates the conditional probability of each feature for the given class and integrates them to ascertain the overall probability, facilitating the prediction of a data point's class based on its features.

To delve into the mechanics of the Naive Bayes algorithm, the initial step involved transforming the provided dataset into a frequency table. This table is instrumental in determining the probabilities associated with the provided features. Subsequently, applying Bayes' theorem, we compute the posterior probability, enabling the estimation of the probability of a specific class given the observed features.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24FEB1647

For a more concrete illustration, let's consider an example from javatpoint.com. Suppose we have a training dataset comprising weather conditions and a corresponding target variable "Play," aiming to predict whether a player should play or not. The dataset is structured is presented in the table below.

Table 3 Example of Dataset

Outlook	Weather	Play
0	Rainy	Yes
1	Sunny	Yes
2	Overcast	Yes
3	Overcast	Yes
4	Sunny	No
5	Rainy	Yes
6	Sunny	Yes
7	Overcast	Yes
8	Rainy	No
9	Sunny	No
10	Sunny	Yes
11	Rainy	No
12	Overcast	Yes
13	Overcast	Yes

- *To Solve this Problem, Convert the Data set into Frequency as Shown after this Statement.*

Table 4 Frequency of Data

Weather	Yes	No
Overcast	5	0
Rainy	2	2
Sunny	3	2
Total	**10**	**5**

- *Then Create a Likelihood Table as Shown in Table 5.*

Table 5 Likelihood Table

Weather	No	Yes	
Overcast	0	5	5/14= 0.35
Rainy	2	2	4/14=0.29
Sunny	2	3	5/14=0.35
All	4/14=0.29	10/14=0.71	

Table 6 Prediction Outcome

P(Y|S)	P(N|S)
P(Y|S)=(P(S|Y)*P(Y)) / P(S)	P(N|S)=(P(N|Y)*P(N)) / P(S)
P (S|Y) = 3/10 = 0.30, P(S) = 5/14 = 0.36, P(Y)= 10/14 = 0.71	P (S|N) = 2/4 = 0.5, P(S) = 5/14 = 0.36, P(N)= 4/14 = 0.29
P(Y|S)=(0.30 * 0.71) / 0.36 = 0.60	P(N|S)=(0.50 * 0.29) / 0.36 = 0.40
Since P(Y|S) is higher than P(N|S), therefore Planer can play the game on a Sunny day.	

Lastly, we apply the Naive Bayes equation to compute the posterior probability for each class. The class with the highest posterior probability is considered the predicted outcome. This helps us determine the most likely class for the given data based on the calculated probabilities. The training data set is to be used in predicting qualified faculty applicants. To come up with a training data set the researcher went through the process of cleaning and interpretation of data. Table 7 is the result.

Table 7 Data for Filtering

Seminar	Certificates	Training	Years of Experience	Educational Attainment	Specialization	Classification
High	Very Good	High	>3	Excellent	Computer Programming	Qualified
Below	Very Good	Poor	1..3	Good	Capstone Project and Research	Not Qualified
Good	Fair	High	>3	Good	Data Structure and Algorithm	Qualified
Below	Very Good	Average	None	Good	Database Management System	Not Qualified
Good	Fair	High	None	Good	Human Computer Interaction	Not Qualified
Below	Fair	Poor	None	Good	Computer Security	Not Qualified
High	Good	Average	<1	Excellent	Networking	Qualified
Good	Very Good	Poor	1..3	Good	System Administration	Not Qualified
Good	Fair	Average	>3	Good	Web Development	Qualified
Good	Very Good	Poor	None	Good	Web Server Management	Not Qualified
Good	Fair	High	None	Good	Computer Programming	Not Qualified
Good	Very Good	Poor	None	Good	Computer Programming	Not Qualified
Good	Very Good	High	1..3	Excellent	Computer Programming	Qualified
High	Fair	Poor	1..3	Excellent	Computer Programming	Qualified
Good	Good	Average	>3	Good	Computer Programming	Qualified
Good	Good	Average	None	Proficient	Computer Programming	Qualified
High	Very Good	Average	None	Proficient	Capstone Project and Research	Qualified
High	Fair	High	None	Good	Data Structure and Algorithm	Qualified
High	Fair	Poor	>3	Excellent	Database Management System	Qualified
Good	Good	Average	>3	Excellent	Web Development	Qualified
Below	Very Good	Average	1..3	Good	Data Structure and Algorithm	Not Qualified
Below	Fair	Poor	1..3	Proficient	Database Management System	Qualified
High	Very Good	Average	None	Good	Web Development	Qualified
Below	Very Good	Poor	1..3	Good	Web Development	Qualified
Below	Very Good	Poor	<1	Excellent	Capstone Project and Research	Not Qualified
Good	Very Good	Average	>3	Good	Data Structure and Algorithm	Qualified
High	Fair	Poor	1..3	Proficient	Database Management System	Qualified
Good	Good	Average	1..3	Good	Web Development	Qualified
Below	Fair	Average	1..3	Proficient	Web Development	Qualified
Good	Very Good	Poor	1..3	Good	Web Development	Qualified
Good	Fair	Average	None	Proficient	Web Development	Qualified
Good	Good	Average	>3	Proficient	Networking	Qualified
Good	Good	Average	1..3	Good	Human Computer Interaction	Qualified
High	Fair	Poor	None	Proficient	Computer Security	Not Qualified
Good	Very Good	Average	1..3	Proficient	Capstone Project and Research	Not Qualified
Below	Very Good	Poor	1..3	Good	Data Structure and Algorithm	Not Qualified
Good	Fair	Poor	None	Proficient	Networking	Not Qualified
Good	Very Good	Average	None	Good	Computer Programming	Not Qualified
Below	Fair	Average	None	Good	System Administration	Not Qualified

C. Faculty Qualification Analysis System Designed for Perpetual Help College of Manila

The Faculty Qualification Analysis System was developed to determine the eligibility of an applicant for teaching demonstration. Employing a filtering technique commonly found in recommender systems, the system utilizes the Naïve Bayes Classifier for classifying large datasets, ensuring precise predictions for class variables with no parents due to its exact learning augmentation. The chosen algorithm, based on Bayes' theorem, is known for its simplicity and effectiveness. It demonstrates proficiency in handling high-dimensional data, a challenge for many other machine learning algorithms. Notably, the role of machine learning in providing recommendations, as seen in job recommendations, aligns with the system's purpose of predicting qualified faculty.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24FEB1647

Table 8 Sources of Data

Seminar	Certificates	Training	Years of Experience	Educational Attainment	Specialization	Classification
3	5	2	4 years	Doctorate	Computer Programming	Qualified
0	6	0	2 years	Bachelor	Capstone Project and Research	Not Qualified
1	0	2	5 years	Bachelor	Data Structure and Algorithm	Qualified
0	5	1	0	Bachelor	Database Management System	Not Qualified
1	0	2	0	Bachelor	Human Computer Interaction	Not Qualified
0	0	0	0	Bachelor	Computer Security	Not Qualified
4	1	1	6 months	Doctorate	Networking	Qualified
1	6	0	1 year and 4 months	Bachelor	System Administration	Not Qualified
1	0	1	4 years	Bachelor	Web Development	Qualified
1	4	0	0	Bachelor	Web Server Management	Not Qualified
1	0	2	0	Bachelor	Computer Programming	Not Qualified
1	5	0	0	Bachelor	Computer Programming	Not Qualified
1	6	3	2 years and 2 months	Doctorate	Computer Programming	Qualified
2	0	0	2 years and 4 months	Doctorate	Computer Programming	Qualified
1	1	1	5 years	Bachelor	Computer Programming	Qualified
1	1	1	0	Master	Computer Programming	Qualified
3	6	1	0	Master	Capstone Project and Research	Qualified
3	0	2	0	Bachelor	Data Structure and Algorithm	Qualified
5	0	0	4 years	Doctorate	Database Management System	Qualified
1	1	1	3 years and 4 months	Doctorate	Web Development	Qualified
0	5	1	2 years	Bachelor	Data Structure and Algorithm	Not Qualified
0	0	0	1 year and 6 months	Master	Database Management System	Qualified
5	6	1	0	Bachelor	Web Development	Qualified
0	9	0	2 years	Bachelor	Web Development	Qualified
0	6	0	6 months	Doctorate	Capstone Project and Research	Not Qualified
1	3	1	5 years	Bachelor	Data Structure and Algorithm	Qualified
5	0	0	2 years and 3 months	Master	Database Management System	Qualified
1	1	1	2 years	Bachelor	Web Development	Qualified
0	0	1	2 years	Master	Web Development	Qualified
1	4	0	1 year and 8 months	Bachelor	Web Development	Qualified
1	0	1	0	Master	Web Development	Qualified
1	1	1	4 years	Master	Networking	Qualified
1	1	1	2 years and 6 months	Bachelor	Human Computer Interaction	Qualified
5	0	0	0	Master	Computer Security	Not Qualified
1	5	1	2 years	Master	Capstone Project and Research	Not Qualified
0	5	0	1 year and 6 months	Bachelor	Data Structure and Algorithm	Not Qualified
1	0	0	0	Master	Networking	Not Qualified
1	6	1	0	Bachelor	Computer Programming	Not Qualified
0	0	1	0	Bachelor	System Administration	Not Qualified

As indicated in the project development phase, following the compilation of pertinent faculty attributes essential for the qualification screening, the next step involves obtaining historical data. This historical data plays a crucial role as it serves as the basis for creating a training dataset utilized by the Naïve Bayes algorithm. Table 8 showcases the data sourced by the researcher from www.kaggle.com, serving as the project's historical data and forming the foundation for constructing the training dataset. Furthermore, the researcher meticulously processed the selection of pertinent attributes from applicants' curriculum vitae, which were subsequently incorporated into the prediction model for faculty qualifications

Table 9 Attributes

Seminar	Certificates	Training	Years of Experience	Educational Attainment
High	Very Good	High	>3	Excellent
Good	Good	Average	1..3	Proficient
Below	Fair	Poor	<1	Good
			None	
>1	>1	>1	3 years and 6 months	Doctorate
1	1	1	2 years	Master
0	0	0	6 months	Bachelor
			0	

Upon evaluating the historical data, the researcher must undertake a crucial step: data conversion, depicted in Table 9. This process is essential to enable the Naive Bayes algorithm to effectively analyze and process the data, ensuring accurate predictions by appropriately transforming each attribute for the algorithm. After the data conversion process is successfully executed, the training data is now ready for the next stage, where it can be employed to predict the desired outcomes. With the completion of the data conversion process, the training data is primed for the subsequent step, involving the prediction of probabilities using the Naive Bayes algorithm.

Table 10 Loook up Table

Seminar	Qualified	Not Qualified
High	12/19	7/19
Good	16/29	13/29
Below	8/18	10/18

Certificates	Qualified	Not Qualified
Very Good	12/28	16/28
Good	10/13	3/13
Fair	14/25	11/25

Training	Qualified	Not Qualified
High	4/6	2/6
Average	22/35	13/35
Poor	10/25	15/25

Years of Experience	Qualified	Not Qualified
>3	13/17	4/17
1..3	11/20	9/20
<1	1/2	1/2
None	11/27	16/27

Educational Attainment	Qualified	Not Qualified
Excellent	6/7	1/7
Proficient	13/20	7/20
Good	17/39	22/39

Specialization	Qualified	Not Qualified
Capstone Project and Research	2/6	4/6
Computer Programming	7/12	5/12
Computer Security	4/7	3/7
Data Structure and Algorithm	5/9	4/9
Database Management System	3/4	1/4
Human Computer Interaction	2/4	2/4
Mobile Develoment	1/2	1/2
Networking	3/8	5/8
System Administration	1/5	4/5
Web Development	8/8	0/8
Web Server Management	0/1	1/1

Table 10 illustrates a lookup table that contains tabulated data showing the frequency counts of an attribute. By employing Bayes' theorem, the conditional probability (P(c|x)) can be calculated using the prior probability (P(c)) and the likelihood (P(x|c)). The following equation represents this conditional probability:

Likelihood

Class Prior Probability

$$P(c \mid x) = \frac{P(x \mid c)\,P(c)}{P(x)}$$

Posterior Probability

Predictor Prior Probability

$$P(c \mid \mathrm{X}) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Image Source: Author

Fig 4 Numerical Equation

In Table 12, shows the result of the computed values of the probabilities for sample Evidence 1 P(x). The P(x) are the attributes that serves as the predictors. Using the formula of Naïve Bayes theorem and the use of Look-up table in Figure 11 with the given attribute condition, the probability of qualified and not qualified is computed.

Table 11 Probability of Evidence 1

Probability of all Qualified Evidences: P(h|q), P(v|q),P(h|q),P(x|q),P(e|q), P(s|q)				
Attribute	P(a|q)	P(a)	P(q)	P(q|a)
Seminar (high)	0.3333	0.2879	0.5455	0.6316
Certificates (very good)	0.3333	0.4242	0.5455	0.4286
Training (high)	0.1111	0.0909	0.5455	0.6667
Experience (>3)	0.3611	0.2576	0.5455	0.7647
Educational Background (excellent)	0.1667	0.1061	0.5455	0.8571
Specialization (Computer Prog'g)	0.1944	0.1818	0.5455	0.5833
P(q|A) = P(q|h) x P(q|v) x P(q|h) x P(q|x) x P(q|e) x P(q|s) x P(q) =				0.0376
Average				65.53%

Probability of all Not Qualified Evidences: P(h|nq), P(v|nq),P(h|nq),P(x|nq),P(e|nq), P(s|nq)				
Attribute	P(a|nq)	P(a)	P(nq)	P(nq|a)
Seminar (high)	0.2333	0.2879	0.4545	0.3684
Certificates (very good)	0.5333	0.4242	0.4545	0.5714
Training (high)	0.0667	0.0909	0.4545	0.3333
Experience (>3)	0.1333	0.2576	0.4545	0.2353
Educational Background (excellent)	0.0333	0.1061	0.4545	0.1429
Specialization (Computer Prog'g)	0.1667	0.1818	0.4545	0.4167
P(nq|A) = P(nq|h) x P(nq|v) x P(nq|h) x P(nq|x) x P(nq|e) x P(nq|s) x P(nq) =				0.0005
Average				34.47%

Based on Table 11, P(q|A) the probability of qualified faculty with the given predictors is higher than the P(nq|A) probability of not qualified faculty. With the given faculty applicant attributes of high seminar, very good certificates, high training, more than three years of experience, excellent educational background and specialization in Computer Programming, the applicant is qualified with P(q|A) of 0.0376 and an equivalent average of 65.53% as shown in Table 11. Thus, the result was computed using Naïve Bayes Algorithm is an objective method that shows accuracy, precision, and truthfulness [21]. in determining qualified faculty applicant for teaching demonstration. Therefore, the applicant can proceed to teaching demonstration.

Table 12 Probability of Evidence 1 Result

Result 1			
Formula	Probability	Result	Average
P(q|A) = P(q|h) x P(q|v) x P(q|h) x P(q|x) x P(q|e) x P(q|s) x P(q)	Qualified	0.0376	65.53%
P(nq|A) = P(nq|h) x P(nq|v) x P(nq|h) x P(nq|x) x P(nq|e) x P(nq|s) x P(nq)	Not Qualified	0.0005	34.47%

➢ *To Provide Another Test on the not Qualified Result, Probability of Evidence 2 is Presented as Shown in Table 13.*

Table 13 Probability of Evidence 2

Probability of all Qualified Evidences: P(h|q), P(v|q),P(h|q),P(x|q),P(e|q), P(s|q)				
Attribute	P(a|q)	P(a)	P(q)	P(q|a)
Seminar (below)	0.2222	0.2727	0.5455	0.4444
Certificates (fair)	0.3889	0.3788	0.5455	0.5600
Training (high)	0.2778	0.3788	0.5455	0.4000
Experience (1..3)	0.3611	0.2576	0.5455	0.7647
Educational Background (proficient)	0.3611	0.3030	0.5455	0.6500
Specialization (Networking)	0.0000	0.0152	0.5455	0.0000
P(q|A) = P(q|h) x P(q|v) x P(q|h) x P(q|x) x P(q|e) x P(q|s) x P(q) =				0.0000
Average				46.99%

Probability of all Not Qualified Evidences: P(h|nq), P(v|nq),P(h|nq),P(x|nq),P(e|nq), P(s|nq)				
Attribute	P(a|nq)	P(a)	P(nq)	P(nq|a)
Seminar (below)	0.3333	0.2727	0.4545	0.5556
Certificates (fair)	0.3667	0.3788	0.4545	0.4400
Training (high)	0.5000	0.3788	0.4545	0.6000
Experience (1..3)	0.1333	0.2576	0.4545	0.2353
Educational Background (proficient)	0.2333	0.3030	0.4545	0.3500
Specialization (Networking)	0.0333	0.0152	0.4545	1.0000
P(nq|A) = P(nq|h) x P(nq|v) x P(nq|h) x P(nq|x) x P(nq|e) x P(nq|s) x P(nq) =				0.0055
Average				53.01%

Based on Table 12, P(*n*q|A) the probability of *not* qualified faculty with the given predictors is higher than the P(q|A) probability of qualified faculty. Therefore, the applicant is *not* qualified as shown in Table 11.

Table 14 Probability of Evidence 2 Result

Result 2			
Formula	Probability	Result	Percentage
P(q|A) = P(q|h) x P(q|v) x P(q|h) x P(q|x) x P(q|e) x P(q|s) x P(q)	Qualified	0.0000	46.99%
P(nq|A) = P(nq|h) x P(nq|v) x P(nq|h) x P(nq|x) x P(nq|e) x P(nq|s) x P(nq)	Not Qualified	0.0055	53.01%

The Level of Effectiveness of the Developed Faculty Qualification Analysis System by The IT Experts in Terms of (1) Functional – Correctness; (2) Efficiency/Performance -Software Capacity; (3) Usability – User Error Protection; (4) Reliability – Fault Tolerance and Recoverability; (5) Compatibility; (6) Security; (7) Maintainability; and (8) Portability.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24FEB1647

Table 15 Level of Effectiveness

Characteristics	Mean	Verbal interpretation
Functional suitability	4.0	High
Performance efficiency	4.5	High
Compatibility	4.5	High
Usability	4.0	High
Reliability	4.0	High
Security	4.5	High
Maintainability	4.5	High
Portability	4.5	High
Overall weighted mean	**4.3**	**High**

Table 15 provides a comprehensive evaluation of the system's characteristics, assigning mean scores on a scale from 1 to 5, where 5 represents the highest rating. The verbal interpretation accompanying each characteristic offers insights into the system's performance based on these mean scores. Notably, the system excels in various aspects, as indicated by high mean scores across functional suitability, performance efficiency, compatibility, usability, reliability, security, maintainability, and portability. The overall weighted mean of 4.3 underscores the system's excellence, positioning it as a high-performing entity. This collective assessment indicates that the system effectively meets its intended functions, operates efficiently, integrates seamlessly with other systems, provides a user-friendly experience, ensures reliability and security, facilitates maintainability, and demonstrates adaptability across different environments. In summary, the system receives a holistic "High" verbal interpretation, affirming its outstanding performance across the evaluated characteristics and meeting or surpassing expectations.

The Level of User Acceptance of the Developed Faculty Qualification Analysis System in Terms of the Following: (1) Functional Suitability – Completeness and Appropriateness; (2) Performance Efficiency - Time-Behavior and Resource Utilization; (3) Usability – Appropriateness, Learnability, Operability, and User Interface Aesthetics; and Lastly (4) Reliability – Maturity and Availability

- *The Result of the Faculty Qualification Analysis System user Level of Acceptance is Shown in Table 11.*

Table 16 User Level of Acceptance

Characteristics	Mean	Verbal interpretation
Functional suitability	4.0	Acceptable
Performance efficiency	4.5	Acceptable
Usability	4.5	Acceptable
Reliability	4.5	Acceptable
Overall weighted mean	**4.4**	**Acceptable**

Table 16 presents an evaluation of the system's characteristics, with mean scores assigned on a scale ranging from 1 to 5, where 5 signifies the highest rating. The accompanying verbal interpretation provides insights into the system's performance based on these mean scores. Across the assessed characteristics — functional suitability, performance efficiency, usability, and reliability — the system consistently achieves mean scores of 4.0 or higher, indicating an "Acceptable" level of performance. The overall weighted mean of 4.4 further reinforces this assessment, positioning the system as acceptable across the evaluated dimensions. In summary, the system is deemed acceptable in terms of its functional suitability, performance efficiency, usability, and reliability, collectively reflecting a solid performance that meets or exceeds acceptable standards.

VI. SUMMARY OF FINDINGS

After extensive research and software development, the developed systems have the potential to change the current approach which can lead to the implementation of the system. Specifically, the recommendation process at Perpetual Help College of Manila spans five working days, which is prone to errors and inefficiencies like mishandling, biases, subjective judgments, and conflicts due to evaluation inconsistencies.

Through the use of Naïve Bayes algorithm, and machine learning from the historical data the study computed the probability of a faculty applicant being qualified. The dataset is composed of attributes like seminars, certificates, training, years of experience, educational background, specialization, and classification. The study identified crucial predictors, including the number of seminars attended, certificates, training, years of related work experience, educational background, and specialization are all instrumental in determining an applicant's qualifications.

In terms of effectiveness, the system achieved an overall weighted mean of 4.3, accompanied by a "High" verbal interpretation, signifying recognition of the system's effectiveness as assessed IT experts. Furthermore, the user acceptance of the system garnered an overall weighted mean of 4.4, with an "Acceptable" verbal interpretation, indicating that the system fulfilled expected functionalities with an 88% satisfaction rate.

VII. CONCLUSIONS

Drawing from the findings of this study, the researcher has derived several significant conclusions about the development of the Faculty Qualification Analysis System(FQAS). The system achieved a high verbal interpretation rating, indicating its reliability in predicting faculty classification. The inclusion of seven specific attributes in the training dataset was found to be critical for accurately predicting the qualifications of faculty applicants for teaching demonstrations. Using collaborative data filtering combined with the Naïve Bayes algorithm shows great potential for predicting faculty classification, thereby streamlining the applicant screening process and reducing the need for manual screening. The system demonstrated promising results, ongoing evaluation and testing are necessary to ensure continued success and effectiveness. Overall, the FQAS holds great promise as a tool to enhance the accuracy and efficiency of qualification prediction processes in educational institutions.

RECOMMENDATIONS

Based on the derived conclusions, several recommendations are proposed for potential actions. These suggestions are grounded in the study's insights to have a potential impact to bring positive outcomes and address the identified areas for improvement.

- Drafting of policies to govern the use and implementation of the Faculty Qualification Analysis System in automating the screening process is recommended based on the result of the evaluation.
- Testing of alternative algorithms in predicting faculty classification to identify the most suitable algorithm for the Faculty Qualification Analysis System.
- Formulating a new dataset model in training the machine to enhance the system's prediction capabilities.
- Implementing a feedback mechanism within the system that would allow users to provide insights on performance and suggest areas for improvement in faculty qualifications, such as attending more seminars or acquiring additional certifications in specific areas. This feedback loop can contribute to continuous refinement and optimization of the system.

REFERENCES

[1]. Chen, S., Webb, G., Liu, L., & Ma, X. (2020). A novel selective naïve Bayes algorithm. *Knowl. Based Syst.*, 192,105361. https://doi.org/10.1016/j.knosys.2019.10 5361.

[2]. Pradeep Kumar Singh, Pijush Kanti Dutta Pramanik, Avick Kumar Dey,Prasenjit Choudhury (2021). Recommender Systems: An Overview, Research Trends and Future Directions. International Journal of Business and Systems Research 15(1): 14-52, DOI: 10.1504/IJBSR.2021.10033303

[3]. Mohanty, Sachi Nandan; Chatterjee, Jyotir Moy; Jain Sarika; Elngar, Ahmed A. ; Guta, Priya (2020). Recommender System with Machine Learning and Artificial Intelligence. Doi:10. 1002/9781119711582

[4]. Kaviani, A., & Dhotre, P. (2017). Naive Bayes Classifier for Text Classification with Python. International Journal of Computer Applications, 167(10), 37-40. DOI: 10.5120/ijca2017914383.

[5]. Autoridad Nacional del Servicio Civil. (2021). Recommendation system for employment in career-based social networking websites.

[6]. Tayade, S. B., Shinde, S. V., & Patel, R. B. (2021). A review of job recommender systems. International Journal of Computer Science and Mobile Computing, 10(3), 232-240.

[7]. Puspasari, A., Irwansyah, A., & Nurhadiyanto, P. (2021). Job vacancy recommendation system for undergraduate students using the K-Means clustering method. Journal of Physics: Conference Series, 1827(1), 012062.

[8]. Tentua, C. M., Kencana, R., & Arliansyah, A. (2020). Job recommendation system using multi-agent system approach. In 2020 3rd International Conference on Computer Science and Artificial Intelligence (ICOCSAI) (pp. 21-26). IEEE.

[9]. Mhamdi, T., Guermazi, H., & Gargouri, F. (2020). A recommender system for job offers based on clustering and user behavior. In 2020 International Conference on Advanced Systems and Electric Technologies (IC_ASET) (pp. 185-190). IEEE.

[10]. Loor, M., & De Tré, G. (2020). Interpretable predictions in naive Bayes classification using augmented appraisal degrees. Expert Systems with Applications, 155, 113404.

[11]. Appadoo, S., Chavda, R., & Maharaj, T. (2020). JobFit: A job recommendation system using machine learning techniques. In 2020 IEEE 6th International Conference on Big Data Analytics and Computational Intelligence (ICBDACI) (pp. 589-594). IEEE.

[12]. Putra, I. K. G., Sumiati, A., & Artawa, K. (2020). A Study on the Implementation of Teaching Materials in the Pandemic Era Using ADDIE and Kirkpatrick Evaluation Models. Journal of Physics: Conference Series, 1567(2), 022033. https://doi.org/10.1088/1742-6596/1567/2/022033

[13]. Nigam, S., Singh, S., Patel, R., & Gupta, A. (2019). A Machine Learning Approach for Job Recommendation using User Preferences. Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 1-6. doi: 10.1109/icccnt45670.2019.8945067.

[14]. Talukdar, D., Saha, S., & Hazarika, S. M. (2018). Analysis of Named Entity Recognition system using different feature sets. Journal of King Saud University-Computer and Information Sciences, 30(1), 1-10. doi: 10.1016/j.jksuci.2016.10.001

[15]. Sugahara, K., Hirasawa, K., & Yamanishi, K. (2018). Exact and Approximate Learning in Bayesian Networks: Comparing the Performance of Conditional Log Likelihood and Marginal Likelihood. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 15(1), 267-275. https://doi.org/10.1109/TCBB.2017.2650818

[16]. Wei, J. (2018). Naive Bayes Classifier. Towards Data Science. Retrieved from https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c

[17]. Urera, F. L. J. & Balahadia, F. F. (2019). ICTeachMUPO: An evaluation of information e-learning module system for faculty and students. International Journal of Computing Sciences Research, 3(1), 163-188. doi: 10.25147/ijcsr.2017.001.1.31.

[18]. Tentuaa, M. N., & Supraptob. (2020). Recommendation System Using Machine Learning: A Review Paper. International Journal of Innovation, Creativity and Change, 12(11), 40. Retrieved from www.ijicc.net.

[19]. Valenzuela, Alberto, Recruitment and Selection Process of Faculty in Higher Education Institutions in the Philippines (August 30, 2019). Available at SSRN: https://ssrn.com/abstract=3445566 or http://dx.doi.org/10.2139/ssrn.3445566

[20]. Dayal, A. (2015). Recruitment and selection process (PGDBA Thesis, HR). Symbiosis Center For Distance Learning, Pune. (Publication No. 78)

[21]. Kapur, R. (2018, March). Recruitment and Selection. Research Gate.

[22]. Tomlinson, G., & Freeman, S. (2017). Who really selected you? Insights into faculty selection processes in top-ranked higher education graduate programmes. *Journal of Further and Higher Education*, 42, 855 - 867. https://doi.org/10.1080/0309877X.2017.1323192

[23]. Upadhyay, A., & Khandelwal, K. (2018). Applying artificial intelligence: implications for recruitment. Strategic HR Review. https://doi.org/10.1108/SHR-07-2018-0051.

[24]. Nawaz, N. (2019). Artificial Intelligence Interchange Human Intervention in the Recruitment Process in Indian Software Industry. Labor: Personnel Economics eJournal. https://doi.org/10.2139/ssrn.3521912.

[25]. Kumar, C., Kumar, L., Muthyalappa, C., Shruthi, D., Shareena, S., & Sireesha, D. (2021). An Integrated College Faculty Recruitment System. *International Journal of Scientific Research in Science and Technology*. https://doi.org/10.32628/ijsrst2183140.

[26]. Singh, J., Salvi, R., Surve, J., & Sawant, A. (2023). A Comprehensive Campus Recruitment and Placement System for Optimizing the Hiring Process. *International Journal for Research in Applied Science and Engineering Technology*. https://doi.org/10.22214/ijraset.2023.50800.

[27]. Tatsat, H., Puri, S., & Lookabaugh, B. (n.d.). Machine learning and data Science blueprints for finance. O’Reilly Online Learning. https://www.oreilly.com/library/view/machine-learning-and/9781492073048/ch04.html