ISSN No:-2456-2165

Integrated Machine Learning Methodology to Enhance Automation in IT Service Desk

Abhishek Choudhury
Software Development Engineer (R&D) Cisco Systems Inc.
Bengaluru, India

Abstract:- In modern-day business, IT infrastructure is the pith of its base. The increasing ramification and high-octane infrastructure and the rising of cloud-based solutions bring challenges concerning time complexity to predict, diagnose and resolution. Generally, infrastructure lies between the data centre and some of the core products, which the business serves. Thus, the data centre is mostly affected by operating system related issues, physical disk crashes or issues in daemon services. The operating system has undergone a rise in the number of issues in recent years. While analysing and training employees to fix the issues can contribute to reducing the service downtime and control the business loses, however, anticipating the coming of new issues and holding the human resources for a long time is a big challenge for any organization. Therefore, industries are in a sense of urgency to find effective and viable solutions to suppress these problems. Incident management is comprised of layers of communication, mainly termed as follows: Tier-1, Tier-2, and Tier-3.

In this paper, we will discuss the possible approaches which can be adopted to minimize the workload on each layer. Compound with language tokenization, identifying the similarity in token using deep learning, and proactively perform tasks to fix the issues.

Keywords:- *Centralized Database, IT IS, Subject Matter Expert.*

I. INTRODUCTION

When an organization's business is expanded, then it brings more people associated with it; likewise, it demands technological innovations to tackle complex situations because clients never compromise on their service expectation. With the burgeoning pace of technology, service and support for technology are getting tortuous and often ineffective at times. The crux behind the failure is that organization often focuses on training their resources time- to-time with the expectation to lower the time complexity, however, this brute force approach backfires the company because IT incident management has many challenges on an employee perspective, such as: to provide support 24/7, Analyze and provide resolution within given time frame. Therefore, an employee's do not wish to commit more time in the organization. As we say-"Innovation occurs at the adverse conditions". Similarly, machine learning can be considered as a boon for ITIS industry. [1] In this paper, our contribution will be to provide solutions and techniques which can be adopted to bring the ITIS to its par. ITIS or ITIL follows a series of steps to resolve issues at their end. Today an undertaking application traditionally comprises of different interconnected utilitarian parts, for example, frontend web server, client validation server, application server, and back-end database server, and every segment may be sent in numerous cases for versatility under an expanded remaining burden. Flaw in any of the parts and their examples influence the application execution and cause issues to the end-clients. Typically, when end-user registers a complain it reaches to Tier-1. Tier-1(often called as service desks) is the first layer of abstraction which is responsible to manage many tasks such as- monitoring servers' health, provide configuration guidelines to the consumers, register a ticket for issues, etc. Now, if an issue is related to any fault occurred on servers then service desk i.e. raises a ticket and delegate to Tier-2, who is the subject matter expert. SME's execute set of instructions to solve the issue, if they succeed then they close the service request, else they delegate it to Tier-3. Further, we will discuss the data mining approach and deep learning to automate the entire process to a certain extent. [2]

II. BACKGROUND

The proposed solution for incident detection is different than the previous works through the data stored during monitoring and the architecture to identify problem determination platform.

A) Monitoring and Data Collection-

In recent years, many approaches have been used in accumulating data and integrating a software solution to gain access to relevant data. Popular monitoring tools have been developed, which runs autonomously into the various servers, generally, these tools run as a daemon service and constantly checks the health of the server. If any suspicion is noticed like hard disk failure, heap allocation, etc then it triggers a notice to the centralized database. Further, a notification will be sent to the service desk. Many service desk management is integrated with incident ticket creation as well, when the incident ticket is opened by the help desk, it immediately marks the incident with a incident ticket number in the centralized managed database.

ISSN No:-2456-2165

B) Issue Identification Strategy-

One the incident ticket is raised; it is the responsibility of the tier-2 SME's to fix the issue. SME's are trained to perform and execute specific set of instructions. If the issue doesn't get resolve they triage multiple teams and engineers to get expertise solution. However, in the past, there have been many service desk solution, which is capable of performing certain tasks for a specific kind of notification. For instance, if a specific server shows out of memory in heap area, then CMDB will have the set of commands to execute. This helped the industry to a great extent however in a scaled system this approach becomes equally twisted and challenging.

C) Guided Workflow-

Recently, another approach has been adopted by IT companies, mainly to solve service desk issues- Guided Workflow. Service desk management provides an interface on a mobile application or their website in the form of support. Service management database categorizes many problems. For instance, an internet service provider based organization put multiple issues in their support menu. e.g. Internet not working, not able to download bill etc. When the user selects, the internet is not working then the related solutions are provided such as Wireless not active, LAN is not active etc. Based on that a service incident is created and the related action is taken places. This entire process is automated. However, this approach has its limitation, it cannot be used for the data centre issues. Therefore, it is mostly used by organizations which are not focused on scalability. [3]

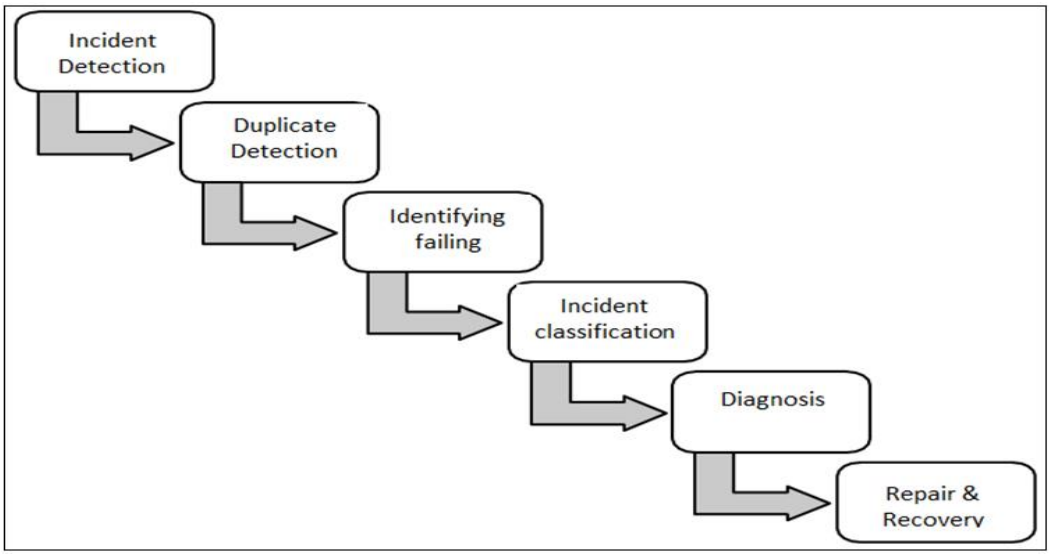

Fig 1:- Incident life cycle

III. PROPOSED SOLUTIONS

In this fragment, we will speak to an answer which computerizes the episode the executive's procedure. In this walk, we will make sense of the approaches to distinguish a comparable episode, which is additionally separated into parts which are as per the following: Text extraction, calculation and estimation, data Indexing, and deep learning. When the incident ticket is submitted to the service desk, the string tokenizer will look for keywords in the description. The tokenizer will also normalize the result set based on the keyword directory available in the CMDB. Further, the indexer will look for the similarity-based on past learning using indexes. The convey engineering computation part is responsible for abstracting the design affected by the occurrence. They all can be considered as the contribution for the Machine Learning.

A) String Tokenizer to Extract Keywords-

First sub-part of the catchphrase extractor, called ticket Parser, extricates the issue portrayal and passes it to the occurrence annotator. If tickets are managed through web pages then we can use HTTP parsing to fetch the incident description, and if it is standalone application then we can use XML parsing as well. Noun and noun phrases are the crux of the English language. In this way, when an episode shows up, we just concentrate things and things states as watchwords. Its portrayal is passed on to a grammatical feature tagger to distinguish thing phrases from the given episode depiction. We use these noun phrases as keywords. This component is responsible for searching the CMDB and returning the CIs matching with the input keywords. Note that CIs are put away as items and these articles must be gotten to utilizing techniques (APIs) gave already. These techniques incorporate including/refreshing an item, erasing an article and looking through utilizing SQL like inquiry. In a commonplace CMDB, there is no in-manufacture strategy to play out a regular catchphrase search. One guileless method

for looking over articles to recover the items and check whether any of its qualities coordinate with at least one of the info watchwords. This plan is delayed as we have to get to every one of the traits of the considerable number of CIs at runtime. To conquer this, we make files over the CMDB. Watchword file maps every catchphrase with a lot of aides. To actualize this record, we expand a transparently accessible system of report ordering. For list creation, we parse the CMDB database, offline, by accessing all the indexes of all the types. [4]

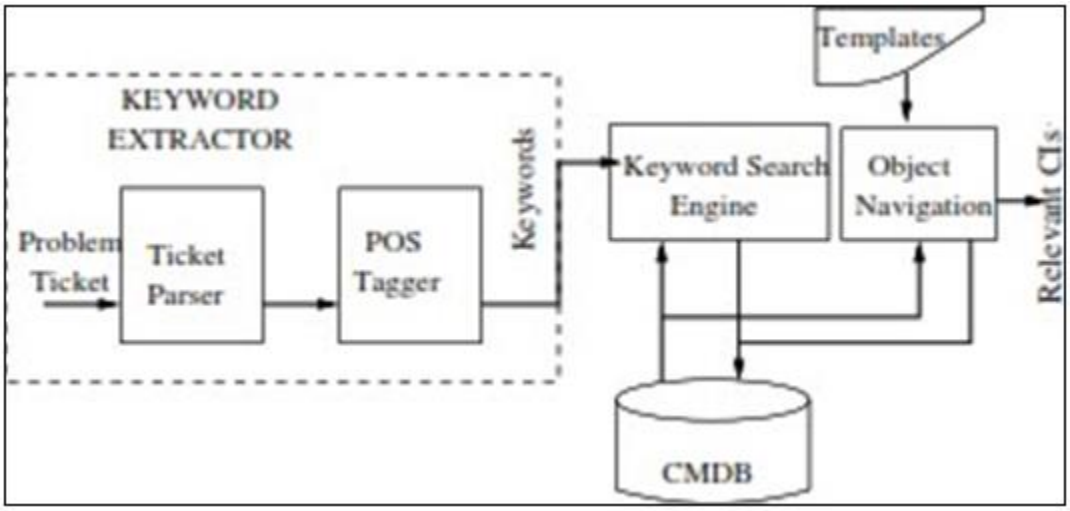

Fig 2:- Lifecycle of tokenizer to resolution

B) Indexing-

Linear search can be time consuming in CMDB. Thus the response time may become unacceptable. Therefore, to improve the time complexity CI indexing need to be imported. This is responsible to search and create indexes prior to the incident creation, to shorten the time. At first, the component rewords all the attributes of the indexes, and returns a normalized keyword set S={S1,S2,S3..}. In the wake of looking through all the lists in CMDB, the segment makes the list, which contains the organized data of record object, for watchwords that are remarkable in finding the occurrence, for example, application name, machine name, have IP and so forth. By and by, clients are various in learning foundation and composing propensities, which thusly may offer ascent to the improper depiction of the bombing module and misdirecting bolster engineers. The framework needs to locate all related list without manual tasks, to improve the automation procedure and beat the above issue. This paper utilizes a hunt strategy, which is appropriate for programmed progress, to locate the related record. Our framework limits the inquiry extension to diminish the execution necessity while holding minor loss of exactness. This pursuit strategy quantitatively estimated the connection between CIs by connection score. As appeared in Figure-3, Relative score is the condition of Relation score.

$$R = \sum_{i \in \Psi} \frac{tf(t_i) * \log\left(\frac{n}{n(t_i)}\right)}{\sum_{j \in \Omega} tf(t_j) * \log\left(\frac{n}{n(t_i)}\right)} \times W_N(i) \times \prod_{k \in X} \frac{W_R(T_{i \to k})}{\alpha}$$

Fig 3:- Relationship score

C) Estimation and calculation-

Now and again, we meet the conditions that parts, which are comparable in convey structures are considered not comparative as they allude to various lists. Be that as it may, the closeness in send engineering gives the plausibility of reusing the arrangement. Send engineering computation module is intended to improve this circumstance. This paper factors the aftereffects of past segments, including principal records. As indicated by the connections between files, we make a non-cyclic chart to outline the design of the episode. A case of occurrence engineering is appeared in Fig 2. From that point forward, the closeness between two diagram in various episode portrayals will be determined. In the figuring, our technique considers the hubs in the engineering as well as the edges and change between sending edges and in reverse edges. The cost capacity represents the cost when changing the parameters from one to the next. The worth is more like 1 when models are progressively comparable, and 0 signifies the two diagrams are completely unique.

D) Integrating Machine learning-

For text classification of a specified event di, using machine- learning, the aim is to find class cj such that assigned numerical score to < di, cj> D C is maximum where D is incident domain and C={c1, c2,.., cn} is the set of predefined incident classes. Figure 2 gives a case of such classes. A worth pij is doled out to

<di, cj> demonstrating likelihood of characterizing episode di under class cj. The capacity for allocating the likelihood to an occurrence class pair is tuned utilizing preparing. We use Naïve Bayesian strategy for episode

ISSN No:-2456-2165

association. According to this procedure a lot of words or arrangement of words are known as occurrence highlights. Every episode is signified as a sack of these highlights. During preparing, a tally nij for each pair of highlight fi and class cj is kept up. The Naïve Bayesian likelihood of finding a component fi in an occurrence of class cj is given by:

$$p(f_i \mid c_j) = \frac{n_{ij} + 1}{\sum_i n_{ij} + N}$$

Fig 4:- Bayesian probability

where N is the all out number of highlights. The estimation of p(fi|cj) can be determined at run-time dependent on nij. For an approaching episode, we make its printed depiction by connecting chosen occurrence fields and tally the occasions each element fi happens in the portrayal.

IV. CONCLUSIONS

Expanding insolubility and energy in IT framework and the rising Cloud administrations present difficulties to opportune incidents/problems conclusion and goals. In this paper, we displayed an issue assurance stage with multidimensional learning incorporation and enablement for productive episode/issue the executives; explicitly, we talked about three highlights: the mechanized occurrence/issue order, the coordination with arrangement database for robotized CI relationship of an occurrence or issue, and the joining with checking frameworks for the gathering of pertinent framework vitals. With the bits of knowledge from building up these highlights, we understood that viable occurrence/issue the executives isn't only an assistance activity issue, it is likewise an issue in each period of the administration life cycle. In unmistakable, occurrence the incident considerations and undertakings ought to be taken care and incorporate assistance plan and sending stages, for example plan and convey administration demand the management devices, policies, and formats for new administrations with the goal that the applicable administration instruments and activity information are appropriately arranged and empowered at administration activity. We proposed a proactive methodology by tending to and incorporating episode/issue the executives in the administration life cycle the executives. In this paper, we displayed a strategy to distinguish the bombing part by incorporating content determined in the issue ticket with organized information put away in the database. We exhibited our execution design and Java-based usage of the episode the board search apparatus. We are mechanizing the choice of episode utilizing content and rule- based arrangement systems which are not displayed in this paper. We intend to broaden the work for autonomic processing and issue the board. Numerous episode tickets have messy information as occurrence portrayal. Taking care of such episodes and better methods of record choice are portions of our future work.

REFERENCES

[1]. "Reducing IT support costs through automated electronic," [Online]. Available: http://www-935.ibm.com/services/us/its/pdf/g510-3499-00.pdf.
[2]. "ITIL v3 Service Operation: Information Technology Infrastructure," [Online]. Available: http://www.itlibrary.org/ index.php?page=ITIL_v3.
[3]. Bon, "The guide to IT Service Management," vol. 1, 2002.
[4]. H. Prasad, R. Gupta and M. Mohania, "Information Integration Techniques to Automate Incident," p. 980, 2008.
[5]. J. Roberts, "AI Driven Automation for Incident Management," [Online]. Available: https://medium.com/kmeanswhat/ai-automation-for-incident-management-c872ee10e833.