Improving Support Ticket Systems Using Machine Learning: A Literature Review

Simon Fuchs
Technical University of Munich
s.t.fuchs@tum.de

Clemens Drieschner
Technical University of Munich
clemens.drieschner@tum.de

Holger Wittges
Technical University of Munich
holger.wittges@tum.de

Abstract

Processing customer support requests via a support ticket system is a key-element for companies to provide support to their customers in an organized and professional way. However, distributing and processing such tickets is much work, increasing the cost for the support providing company and stretching the resolution time. The advancing potential of Machine Learning has led to the goal of automating those support ticket systems. Against this background, we conducted a Literature Review aiming at determining the present state-of-the-art technology in the field of automated support ticket systems. We provide an overview about present trends and topics discussed in this field. During the Literature Review, we found creating an automated incident management tool being the majority topic in the field followed by request escalation and customer sentiment prediction and identified Random Forrest and Support Vector Machine as best performing algorithms for classification in the field.

1. Introduction

Providing technical support for own IT products is an integral part of software developing or software providing companies [1, 2]. For this purpose, most companies provide their customers support ticket systems (STSs), in which users can create incident tickets describing their problem or request [1]. In most state-of-the-art STSs, at least some key decisions in distributing these support tickets to the responsible support assistant or support team are still made by support staff members [3]. Support ticket distribution has the potential to bind a lot of working time of technically skilled workers, wherefore big companies often use less-skilled or temporary workers for support ticket distribution or outsource the support to a third party entirely [4]. This process of manually distributing emerging support tickets by often less-skilled human workers is on the one hand ineffective and expensive [3]. On the other hand, it mostly increases the ticket resolution time and therefore lowers the satisfaction of the customer initially creating the ticket [1, 5, 6].

At the same time, the volume of support tickets in IT-companies created by customers has significantly grown due to the digitalization efforts currently made across all industries [7]. This means that IT-companies face an increasing pressure in automating their STSs to cope with the rising volume of tickets [7], to increase customer satisfaction [1, 8], to accelerate support management processes [8, 9], and to reduce costs [4, 10].

With Artificial Intelligence (AI) and Machine Learning (ML) algorithms becoming commonplace, the automating of STSs has become more interesting than ever before [11]. Technologies for automated ticket classification and automated ticket resolution using ML open the possibility of automating basic day-to-day IT tasks replacing the first level support staff members [4, 8].

This background in mind, we analyze current trends and topics in automation of STSs using ML by undertaking a narrative Literature Review [12]. We follow the principles of Watson and Webster [13] to identify the relevant literature and to analyze the present state of the art of automating STSs in the latest scientific literature. In the process, we defined the following research questions to guide our Literature Review: *What is the present state-of-the-art technology in automating STSs? Which ML algorithms have the highest accuracy in classifying support tickets?*

In answering these questions, we aim to provide an overview of the technical status quo of ML-driven automation of STSs. Further, we aim to identify research gaps in the current state of the art.

During our literature search, we recognized that only few Literature Reviews in this field of research have been published, not providing a general overview of the field. With this Literature Review we want to provide such a general overview over the present

URI: https://hdl.handle.net/10125/79570
978-0-9981331-5-7

technological state of the art and the current status quo of this particular field of research. Further, we want to provide a paper for newcomers in the field to start with.

2. Background

STSs mostly enable customers to create a support incident. The term "support incident" denotes the whole entity of one single support process. Such support incidents mostly comprise a support ticket, including a title, the plain text of the ticket and special information, often called meta-data, like the priority of the ticket, the category of the incident, a ticket-id, etc. Furthermore, such incidents can comprise files attached to a ticket or further customer data collected by the system [1, 6, 8]. If we speak of "incident management tools" in the course of this paper, we refer to a concrete IT artifact that is able to manage support incidents within a STS. In this context, managing an incident means: answering the incident, asking questions to the creator of it or distributing it to a responsible support agent.

The term "Machine Learning" has always denominated a very broad field of technical solutions aiming at making "intelligent" machines [14]. In the context of this Literature Review, the term "Machine Learning" is used for technologies comprising algorithms, mathematical models and approaches that enable an IT artifact to automatically classify support incidents based on previously provided training data.

At this point, we want to highlight the differences between chatbots and STSs. Chatbots are thought for 24-hour, real-time customer support, whereat mostly very simple questions are meant to be answered [15]. STSs are meant to be a communication tool between customers and technical agents to solve technical problems occurring at customer side [1, 5]. Mostly, a support agent has to actively do something to solve the customer request, for example install a program, modify a firewall, unlock ports, etc. or the question asked by the customer is difficult and one or more support agents have to think about or have to investigate in order to solve the request [1]. Simply said, a chatbot is meant to answer simple customer questions or for customer guidance [16], a STS is meant to help customers to create more difficult, technical requests [8]. For this reason, we think that beside the growing hype around chatbots the field of STSs will stay an independent and relevant field of research.

3. Literature Review Design

Our main objective was to investigate the status quo in the field of automating STSs using ML, to understand state-of-the-art technology and to identify research gaps in this field of research.

The Literature Review presented in this paper was performed during February 2021. We searched the databases Scopus, IEEE Xplore, Ebsco and Web of Science. We limited the search to papers in English language. It fast became obvious that precise search words had to be found to confine the field of ML-automated STSs from more general fields like support system improvement without ML or theoretical papers concerning topics like Natural Language Processing or Deep Learning. In the end, we used some keywords connecting them in various ways using the Boolean Operators AND and OR. The search words we used were "service desk", "support ticket", "support tickets", "Machine Learning", "ML", "Artificial Intelligence", "AI" and "classification".

Depending on each database, we found another combination of those keywords to be useful to find relevant literature. For example, in Scopus we needed more keyword to confine the search, whilst in the Web of Science database fewer keywords lead to relevant results. The keywords finally used for each database are shown in Table 1. Unfortunately, it was not possible to find one general search string working in all databases due to the limits and specifications of these databases.

In total, we initially received 122 hits, from which we regarded 60 hits as relevant. Those 60 hits regarded as relevant comprised several duplicates. Eliminating duplicates, we received 41 relevant papers. Doing Forward and Backward Search [13] we additionally found 2 papers we regarded as relevant. Google Scholar was consulted during both.

The criteria for a relevant hit comprised: *The main topic of the paper was a STS (the keyword "ticket" sometimes produced hits in the field of tickets for air travel or festivals); Machine Learning, or at least Data Science was actually used to improve a STS; Technical solutions were developed, discussed, tested or deployed.*

4. Findings of the Literature Review

4.1. Paper Type and Publisher

A first finding of our Literature Review was that a broad spectrum of publishers published the papers found. Most papers (26 papers) found in the literature search were published on conferences.

Table 1: Overview of search keywords and search results

Databases	Search Keywords	Hits	Rele-vant Papers
Scopus	("support ticket" OR "support tickets" OR "service desk") AND ("Machine Learning" OR "AI" OR "ML") AND ("classification")	53	26
IEEE Explore	("All Metadata":support ticket) AND ("All Metadata":classification) AND ("All Metadata":Machine Learning)	17	13
	("All Metadata":service desk) AND ("All Metadata":classification) AND ("All Metadata":Machine Learning)	5	5
Ebsco	TI = (ticket and classification)	1	1
	TI = (incident management AND Machine Learning OR artificial intelligence)	2	2
Web of Science	TI=(support ticket*)	14	6
	TI=(service desk *)	30	8
In Total		122	60
Without Duplicates			39
Articles identified in Forward & Backward Search			2
Final Selection			41

Table 2: Overview of paper type and publishers

Paper type	Number of Hits	in Percent
Journal	17	39,5%
Conference	26	60,5%
Publisher	**Number of Hits**	**in Percent**
IEEE Conference	11	26,8%
IEEE Journal	5	12,3%
ACM	2	4,8%
Other Journal	11	26,8%
Other Conference	12	29,3%

Table 3: Concept matrix

Paper	IM Tool	Requ. Escal.	Sent. Pred.	Other
Asres, Mengistu [17]	x			x
Nayak, Rai [15]	x			x
Yang [18]	x			
Amin, Lancaster [19]	(x)			
Baresi, Quattrocchi [20]	x			x
Han and Sun [21]	x			
Mour, Dey [22]	x			x
Revina, Buza [23]	x			
Xu, Mu [24]	x			x
Al-Hawari and Barham [8]	x			
İşcen and Gürbüz [2]	x			
Lo, Tiba [25]	x			x
Mukunthan and Selvakumar [26]	x			
Nayebi, Dicke [6]	x	x		
Mandal, Agarwal [11]	x			
Misra and Podder [27]	(x)			x
Palacios, Morillas [28]	(x)			x
Shanmugalingam, Chandrasekara [29]	x			
Werner, Li [30]	x	x	x	
Gajananan, Loyola [31]		(x)	x	
Gupta, Asadullah [1]	x			x
Han, Goh [32]	x			
Koehler, Fux [33]	x			x
Lyubinets, Boiko [34]	(x)			x
Meng, Xu [35])				x
Montgomery, Damian [5]	(x)	x	(x)	
Paramesh, Ramya [36]	x			
Parmar, Biju [37]	x			
Patidar, Agarwal [38]				x
Saberi, Theobald [39]				x
Silva, Pereira [9]	x			
Stein, Flath [7]				x
Qamili, Shabani [3]	x		x	x
Zhou, Zhu [40]				x
Zuev, Kalistratov [10]				x
Xu, Zhang [41]	x			
Chagnon [4]	x			
Giurgiu, Wiesmann [42]	x			
Montgomery and Damian [43]		x		
Reddy, Reddy T [44]				x
Goby, Brandt [45]	x			x

(x) connotes a mentioning as minor topic or only implicitly. Categories in the columns: development/ deployment/evaluation of an incident management tool; customer request escalation prediction; sentiment prediction, and Other. Papers are sorted according to their publishing date from recent to prior.

Over 35% of all found papers (16 papers) were published on IEEE conferences or in IEEE Journals. As written later below, most papers pertain a certain, practical use case. We therefore conclude that the topic of STSs automation is mainly interesting for the application-oriented community that prefers to publish at conferences. All findings pertaining paper type and publisher are shown in Table 2.

4.2. Topic Analysis

In a next step, we analyzed and categorized the topics dealt with according to Corbin and Strauss [46]. Especially, we were interested in technical features that are developed, tested, and deployed in the present research. These findings are presented in Table 3.

The topics discussed the most throughout the relevant literature are: *development, deployment and evaluation of an incident management tool, customer request escalation prediction,* and *customer sentiment prediction*. Apart from that, the topics examined in the literature were quite individual. Papers that examined further individual topics were (also) categorized as *Other*.

4.3. Machine Learning algorithms used

One of the questions that we were the most interested in was *Which Machine Learning algorithms are already in use and which have already proven themselves to work?* 17 of the examined 41 papers implemented and evaluated at least one ML algorithm or solution known from literature that is more theoretical. 15 of these compared at least two of such algorithms/solutions with each other.

To determine the "best performing" algorithm in those papers, we first looked if the authors of each paper identified one of their algorithms as "best performing". If this was not the case, we compared the algorithms presented by their reported accuracy, precision and recall. In case of similarly performing results, we weighted the reported accuracy as tiebreaker. A tabular overview of our findings regarding used ML algorithms and how often they were the best performing ones in a paper is presented in Table 4.

The ML algorithm most used (10 papers) for support ticket classification was (sometimes a modified version of) the Support Vector Machine (SVM) approach, which was also the algorithm that best performed in most papers (6 papers). In these 6 papers accuracies between 63% [37] and 98% [18] were reached. Those results are heavily depending on the data set used, but overall the papers an accuracy between 80-90% for SVM seems reasonable [1, 2, 8, 9, 18, 37].

The second place (7 papers) is taken by (versions of) the Random Forrest (RF) approach that performed best in 3 papers. Here, we found maximal accuracies of 78% [30], 90% [6] and 92 % [36].

As it is shown in Table 4, RF performed especially well in request escalation prediction and sentiment prediction, while in ticket classification SVM performed better.

Table 4: ML algorithms and ML solutions evaluated and compared in the literature

ML Algorithm	Ticket Class.	Requ. Escal	Sent. Pred.
K-Nearest Neighbor classification (KNN)	4 (0)		
Support Vector Machine (SVM)	9 (6)	2 (0)	1 (0)
Decision Tree (DT)	4 (0)		
Rule-Based	3 (0)		
Naïve Bayes	6 (0)	2 (0)	1 (0)
Random Forrest (RF)	6 (2)	2 (2)	1 (1)
DNN	6 (3)		
unsupervised learning	1 (0) <60% accuracy		

The number in each cell indicates how often an algorithm was evaluated for a specific task. The number in brackets behind indicates how often the algorithm was graded the best performing in the paper. The abbreviations for the tasks indicate (from left to right): ticket classification, customer request escalation prediction, customer sentiment prediction.

Additionally, Naïve Bayes approaches were presented and evaluated in the literature (7 papers) but no paper was found in which a Naïve Bayes algorithm could beat an SVM or an RF algorithm.

In addition, a trend can be identified over the years. According to Revina, Buza [23], the earliest approaches of support ticket classification were ruled-based approaches that were outperformed by RF and SVM. This probably is the reason that mostly in the years 2018 and 2019, RF and SVM were the approaches most implemented and evaluated and also best performing [2, 6, 8, 9, 30, 36]. Since 2019, some authors also propose artificial-neuronal-network-based solutions, because of the good performance of neuronal network approaches in other fields of ML [19, 21, 23, 29]. While there is some promising evidence, that Deep Neuronal Network (DNN) based solutions might be able to outperform present ML approaches like RF or SVM [29] (reported accuracies: 77% [29], 83% [19], 99% [21]), this point seems not to be fully reached yet [23]. Revina, Buza [23] find in their conclusion (page 11) “that simple algorithms work well if using appropriate linguistic features” and can be equally performative.

When analyzing the 4 papers presenting DNN-based approaches [19, 21, 23, 29], we found that two of those papers used specific LSTM-networks, which were both the best performing approaches in that papers [19, 29]. Since two papers are not a sufficiently large sample, no general statement can be derived here. Nevertheless, LSTM DNNs seem to be a promising approach to which further attention should be paid.

4.4. Incident Management Tool

To develop, implement, test and deploy an automated tool for incident management is by far the topic most treated in the found literature. 31 of the in total 41 examined papers presented a complete, or at least components for a practical support ticket classification artifact. The remaining 10 papers discussed or evaluated only parts of such tools or treated support ticket classification in a more theoretical approach. We take this finding as evidence that the majority of research in the field is focused on creating practical technical applications.

In most cases, a practical use case was on hand, such that the authors mainly developed such an artifact for or optimized an existing one in a distinct situation (for example in [1]; [8]; [6]; [29]; or [17]). This of course leads to a problem of generalizing their results as already ascertained by Misra and Podder [27]. The accuracy data presented by those papers should therefore also be read with caution.

Misra and Podder [27] also stated that theoretical ML knowledge obtained by the scientific literature often has problems in unfolding its full potential when being applied in practical technical solutions, because in reality either training data is not accessible like in experiments or real-world data is more noisy or diverse than polished data for university experiments.

Across all literature analyzed, there is a consensus, that training data is the linchpin of developing a well-functioning ML tool. In such, it is not surprising that every practical use case of developing, deploying, testing and optimizing a real support ticket is massively impacted by the training data accessible to the authors and the ticket data that are meant to be classified.

4.5. Request Escalation Prediction

The term “Request Escalation” denotes the phenomenon that during a support process sometimes customers are not satisfied and escalate their request by pressuring the support agent or the agents’ supervisors. An escalation of a customer request mostly occurs when the customer creating the request or a support manager is dissatisfied with the way a support ticket is processed and requests its escalation. This request escalation leads to a concentration of extensive human resources for solving the customer request [30] within the support providing company. This concentration of human resources means more stress for the involved support agents and support managers and additionally, these escalations are mostly expensive for the company and significantly lower customer satisfaction [30, 43].

Five papers of the analyzed literature deal with ML-supported or even ML-driven escalation process optimization. The primary goal is often to predict the risk of a support request to escalate, such that a company can concentrate resources on a customer request before the customer gets angry and escalates his/her request [5, 43]. Montgomery, Damian [5] developed a prototype that is able to predict escalation probability in percent using Random Forest and XBoost algorithms reaching an accuracy of 81% on their deployed use case. The artifact ESMMArT presented in Nayebi, Dicke [6] is also able to predict an escalation probability using Random Forrest reaching 90% accuracy. Further, Werner, Li [30] developed a cost-based mechanism to train and evaluate ML algorithms for request escalation prediction.

4.6. Sentiment Prediction

The field of customer request escalation is closely connected to the field of customer sentiment prediction. As mentioned above, customer satisfaction is a very important good when running an IT related business. As also mentioned, the quality of support services heavily influences customer satisfaction [1, 30]. Therefore, it would be very useful for companies to be able to predict the sentiments of their customers, while a support incident is open [30]. In this context, the term sentiment describes the general feelings of a customer and his overall attitude towards the company and its support services. Mostly in the literature, the customer’s sentiment is classified in the three categories “positive”, “neutral” and “negative” [30].

Some of the analyzed papers deal with prototypes for customer sentiment prediction. The main goal is hereby to predict the sentiment that a customer is experiencing during a support process, mostly after he/she created a support ticket [3].

The predicted sentiment of a customer can then be used for a wide scope of application: Gajananan, Loyola [31] developed an artifact predicting customer sentiment from support ticket data for predicting the probability if a customer would renew his subscription at a cloud business service; Werner, Li [30] use customer sentiment prediction to predict the risk of support escalation; and Qamili, Shabani [3] want to use customer sentiment prediction for internal evaluation of the support system.

The results in the field of emotion/sentiment prediction are mixed. Qamili, Shabani [3] report low accuracy (<45%) due to the difficulties in accumulating labeled data for sentiment prediction. Gajananan, Loyola [31] report high accuracy for their subscription renewal prediction, but do not directly predict customer sentiment. Instead, they only use sentiment polarity as a parameter of their model. Werner, Li [30] report pleasing successes using the Watson NLU model, but give no values for their emotion prediction. Overall, the impression arises that so far only the surface of ML-based customer sentiment prediction has been scratched.

4.7. Specific topics according to the special use case - the category *other*

Every paper analyzed during this Literature Review is heavily influenced by the practical problem it was written around. The presented artifacts and prototypes were always developed and adjusted for a special use case and therefore, many papers at hand face, present and solve very special problems. As a result, in addition to the dominant topic of automated incident management tools, a large number of other topics are dealt with or broached. A selection of these reads:

- Automate labeling of tickets, either for training or for easier ticket resolution by a support agent [34]
- Chatbots [15, 33, 44]
- Spam detection [3]
- Performance optimization [24]
- Automated analysis of pictures attached to a support ticket [11]
- Business/process/text mining for better support system architecture [25, 45]
- AI explainability in support ticket automating [22]
- Ticket resolution time prediction [10]
- Automated STSs in context of Internet of Things (IoT) [28]
- Using answering bot (Microsoft LUIS) for automated request responses [29]

The findings and results in these topics were as diverse as the topics itself. Nevertheless, we were able to carve out some general findings:

- As in nearly any research in the field of Machine Learning the accessibility and quality of training data importantly influences the outcome of the project [1, 3, 25]
- The metrics precision and recall are by far the most-used metrics for evaluating ML ticket classification tools [5, 6, 18, 36].
- Classification tools work more precisely the fewer classes they have to classify to [1, 3, 25].

4.8. Data sets used

Most papers analyzed in this Literature Review used own datasets consisting of ticket data acquired in their specific use case. Often, papers were written in cooperation with companies providing ticket data for training and automated STSs were developed for these companies. Because of non-disclosure agreements, these data sets are not available for the community and also not described in detail in the papers analyzed.

Within the small group of remaining papers, we could not identify any publically available data set used in more than one paper.

5. Discussion

The primary goal of this Literature Review was to provide an overview over the field and to identify the present state of the art of automating STSs using Machine Learning. Our first finding was that several prototypes of ML-automated STSs were developed, deployed and evaluated in the past 5 years. Indeed, to develop, deploy and often evaluate a practical prototype for a specific use case is the dominating topic within the field of ML-automated STSs. At this juncture, every prototype presented is heavily shaped by the specific use case it was developed for. This leads to an overall problem of generalizability and validity of the results presented in each paper [1, 6].

Therefore, on the one hand, more comparative research and meta-analysis of prototypes developed is needed. On the other hand, we understand the specific constraints every practical use case brings along. Technical limitations like the volume of accessible training data or the volume of topics discussed within a STS; special requirements like the number of categories to classify tickets in, the number of support teams working at a company, a maximal response time requested by the management or technical standards within a company; and individual features like a request escalation pipeline, a spam detector, a company-specific evaluation metric or the like all have to be engaged individually. Therefore, we expect to see more prototypes for specific use cases presented in future research work.

Regarding the variability of algorithms applied and tested in the papers, it appears clear that the specific use case each paper deals with heavily influences which algorithms performs the best in each paper. Therefore, in order to find the "best performing" algorithm in one specific use case always fine-tuning and dealing with the specific use case will be necessary. Nevertheless, a statement can be made about which ML algorithms and solutions have proven themselves during the past 5 years not in one paper only, but in various papers and hence in various use cases. Generally, ML algorithms like SVM and RF have proven to be more accurate and precise than "older" rules-based approaches or Decision Trees [2, 8, 23]. Also, there is some evidence that DNNs could outperform the present best-performing algorithms SVM and RF, especially in the case of large training data and many classes to classify tickets in [21, 29]. SVM and RF can therefore be considered standard approach algorithms for support ticket classification.

Nevertheless, so far there is only pioneering work in the field of Deep Learning approaches for support ticket classification. Only 4 of the 41 analyzed paper treated Deep Learning in the context of support ticket classification. More research is needed to evaluate if Neuronal Networks can outperform present ML algorithms in support ticket classification. Although, there is a lot of research ongoing in the field of Deep Learning, the application of those results in STSs should be intensified. Closing, we want to argue that there is also a need for more prototype-based research developing DNN-based, automated support ticket desks.

Respective the topic of customer sentiment/emotion prediction, we see a great potential for further research work. Knowing customer sentiment has big economical potential [30] either by helping customer requests not to escalate [5], by helping to prioritize requests [6] or by helping to predict if customers would renew their subscriptions [31]. But, multiple other use cases for knowing customer sentiment can be thought. Nevertheless, research in the field of predicting customer sentiment based on their created support tickets is only few (see above). Therefore, more research in this field is needed.

Pertaining the topics categorized as Other, most of the presented topics in the category Other above are relevant and highly researched topics in the field of AI and ML in general. Especially, topics like AI explainability (226 hits in IEEE database), chatbots (606 hits), IoT (57k hits) and performance optimization (132k hits) are highly investigated. Therefore, it is not surprising that these topics were also researched in the context of support ticket automating.

Closing, we want to point out four topics that we considered being research gaps after analyzing the found literature:

First, we found that in most papers, automating an IT support desk is regarded as a process that reduces labor cost and increases customer satisfaction [1-4].

However, the perspective of support agents and support managers and how ML-optimized support desks can improve their working life is missed in the literature. Often, problems like angry customers, false ticket allocation, frequently asked questions, etc. stress or annoy support agents and support managers the most. Unfortunately, there is no research investigating how ML-driven automating can increase the job satisfaction of support agents and support managers, especially in the case of smaller companies with only limited human resources and the challenge of using their skilled employees as 2nd level support agents and productive workforce at the same time.

Second, in most papers analyzed the ML models were trained by human-guided training with data manually labeled, e.g. in [1], [2], [6], or [8]. Manually labeling thousands of tickets is mostly an unthankful work, therefore it would be nice to apply unsupervised learning solutions for training data labeling. Actually, Lo, Tiba [25] did some pioneer work in this field. Unfortunately, their unsupervised Kmeans and DBSCAN algorithms did not work very well. Also, the unsupervised algorithms tested by Nayebi, Dicke [6] did not perform very well. Revina, Buza [23] used a semi-supervised algorithm and comes to the finding: "semi-supervised learning (SSL) allows inducing a model from a large amount of unlabeled data combined with a small set of labeled data" [23, page 4] that was promising but still needed some labeled data. We conclude that more research in this field is needed to lower the effort for labeling training data enormously.

Third, we found the approach presented by Shanmugalingam, Chandrasekara [29] very interesting, where the authors of the paper created a ticket classification tool, in which the determined class information are transmitted to a solution bot in order to generate a STS that automatically answers customer requests. As already said, there is a lot of research going on pertaining chatbots and question-answer-systems are also a ML field that is researched in Jiang, Su [47] and Shevchenko, Eremin [48]. Surprisingly, most papers analyzed during this Literature Review do not use the class information received from their incident management tools to automate ticket answering. Especially in the case of commonly asked questions or so-called first-level support requests, automated ticket answering can save costs for the company and resolution time for the customers. Therefore, we argue that more research in the field of automated ticket answering should be performed.

Fourth, we recognized that all papers examined only focus on the positive impact that ML-automated STSs have on companies and customers. Meanwhile, several studies have revealed skepticism of people towards the usage of ML or AI (for example Gherheș and Obrad [49] or Morikawa [50]) and there are also reports about negative effects of the usage of ML solutions in companies, like customer being around 80% less satisfied and less likely to purchase when they realize they communicate with a chatbot [16]. Sometimes, there is a report of models not performing that well yet (for example in Lo, Tiba [25]), but we miss a treatment of negative aspects of ML-automated STSs, like customers not trusting the systems, support agents not understanding the decisions of the system, etc. Therefore, we argue for more attentiveness regarding possible negative effects of ML-driven automation of STSs and especially we suggest publishing such negative experiences and not only the promising results.

Additionally, we argue that STSs are an essential part of IT-related business and the potential of cost-reduction, raise in customer satisfaction and raise in employee satisfaction through ML-driven automating of these STSs makes the field not only interesting for Computer Science Research or Engineering but especially for Information Systems Research.

Finally, in our personal research in the field, we found that customer guidance while ticket creation can help improve the quality of trainings data, while also increasing customer satisfaction and increasing the accuracy of the ticket classifiers. Regarding the papers analyzed in this Literature Review, this topic seems not investigated yet.

For the future the authors of this paper aim at doing more research in the topics of customer guidance for better data, unguided machine learning and (semi-)automated question answering. In particular, the authors aim at creating an own STS for their own specific use case.

6. Limitations of this Literature Review

We intentionally narrowed the scope of this Literature Review to the topic of STSs automation. This means that we did not consider literature in the broader fields of ML-driven text classification, text understanding, text generation or text processing. In addition, literature pertaining ML question answering systems were not considered. This of course means that research gaps identified above might be solved theoretically by research in those other fields. Nevertheless, the practical application of such theoretical known solutions to the field of STSs is a promising task not realized yet.

The Literature Search did only comprise searching scientific databases. We deliberately did not search google, patent literature or other non-scientific publications. For this reason, we did not examine the present state of the art of commercial industry solutions like ServiceNow or SAP Service Ticket Intelligence, hence these were not examined in the analyzed literature.

7. Conclusion

Support ticket help desks are an essential part of modern companies. As developments in the field of ML advance, it becomes interesting to automate support ticket help desks using ML solutions to lower error rate and cost within the support department. In order to investigate the present state of the art in the field of ML-driven support ticket automating we conducted this Literature Review.

We found that every paper analyzed is heavily influenced by the practical use case it was written about and that this is due to the individual nature of support ticket systems, the companies they are used in and the special, technical requirements raised by ML algorithms. We also found that the majority of the papers in the field is published on conferences.

We found the topic of creating an automated incident management tool is the dominating topic in this field of research. ML algorithms like RF and SVM are currently best performing in ticket classification, while there is evidence that Deep Learning algorithms will take the lead in the near future. Two other important topics in the field found in this Literature Review were Request Escalation prediction, in which it is the goal to optimize processes within the support system handling customers escalating their requests, and customer sentiment prediction, in which it is the goal to predict the sentiment of a customer based on the tickets he/she created. Additionally, we found a great variety of topics individually dealt with in single papers.

We provide an overview over the current technological state of the art, and give suggestions in which directions the research scope can be expanded, and identify research gaps.

8. References

[1] Gupta, M., A. Asadullah, S. Padmanabhuni and A. Serebrenik, *Reducing user input requests to improve IT support ticket resolution process.* Empirical Software Engineering, 2018. **23**(3): p. 1664-1703.

[2] İşcen, E.S. and M.Z. Gürbüz. *A Comparison of Text Classifiers on IT Incidents Using WEKA*. in *2019 4th International Conference on Computer Science and Engineering (UBMK)*. 2019.

[3] Qamili, R., S. Shabani and J. Schneider, *An Intelligent Framework for Issue Ticketing System Based on Machine Learning*. 2018. 79-86.

[4] Chagnon, C.T., A.C.; Djamasbi, S., *Creating a decision support system for service classification and assignment through optimization*, in *AMCIS 2017 - America's Conference on Information Systems: A Tradition of Innovation*. 2017, Worcester Polytechnic Institute, Germany.

[5] Montgomery, L., D. Damian, T. Bulmer and S. Quader, *Customer support ticket escalation prediction using feature engineering.* Requirements Engineering, 2018. **23**(3): p. 333-355.

[6] Nayebi, M., L. Dicke, R. Ittyipe, C. Carlson and G. Ruhe, *ESSMArT way to manage customer requests.* Empirical Software Engineering, 2019. **24**(6): p. 3755-3789.

[7] Stein, N., C. Flath and C. Böhm, *Predictive Analytics for Application Management Services*. 2018.

[8] Al-Hawari, F. and H. Barham, *A machine learning based help desk system for IT service management.* Journal of King Saud University - Computer and Information Sciences, 2019.

[9] Silva, S., R. Pereira and R. Ribeiro. *Machine learning in incident categorization automation*. in *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*. 2018.

[10] Zuev, D., A. Kalistratov and A. Zuev, *Machine Learning in IT Service Management.* Procedia Computer Science, 2018. **145**: p. 675-679.

[11] Mandal, A., S. Agarwal, N. Malhotra, G. Sridhara, A. Ray and D. Swarup. *Improving IT Support by Enhancing Incident Management Process with Multi-modal Analysis*. in *International Conference on Service-Oriented Computing (ICSOC)*. 2019.

[12] Paré, G., M.-C. Trudel, M. Jaana and S. Kitsiou, *Synthesizing information systems knowledge: A typology of literature reviews.* Information & Management, 2015. **52**(2): p. 183-199.

[13] Watson, R.T. and J. Webster, *Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0.* Journal of Decision Systems, 2020. **29**(3): p. 129-147.

[14] Stone, P.B., Rodney; Brynjolfsson, Erik; Calo, Ryan; Etzioni, Oren; Hager, Greg; Hirschberg, Julia; Kalyanakrishnan, Shivaram; Kamar, Ece; Kraus, Sarit; Leyton-Brown, Kevin; Parkes, David; Press, William; Saxenian, AnnaLee; Shah, Julie; Tambe, Milind; Teller, Astro;, *"Artificial Intelligence and Life in 2030."*. Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016, 2016.

[15] Nayak, S.P., A. Rai, K. Vankataramanappa, J.A. Pansuriya and J. Singler. *Adaptive MoD Chatbot: Toward Providing Contextual Corporate Summarized Document as Suggestions and Reported Issue Ticket Routing*. in *Advances in Artificial Intelligence and Data Engineering*. 2021. Singapore: Springer Singapore.

[16] Luo, X., S. Tong, Z. Fang and Z. Qu, *Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases.* Marketing Science, 2019. **38**(6): p. 937-947.

[17] Asres, M.W., M.A. Mengistu, P. Castrogiovanni, L. Bottaccioli, E. Macii, E. Patti and A. Acquaviva, *Supporting Telecommunication Alarm Management System With Trouble Ticket Prediction.* IEEE Transactions on Industrial Informatics, 2021. **17**(2): p. 1459-1469.

[18] Yang, L., *Fuzzy Output Support Vector Machine Based Incident Ticket Classification.* IEICE Transactions on Information and Systems, 2021. **E104.D**(1): p. 146-151.

[19] Amin, K., G. Lancaster, S. Kapetanakis, K.-D. Althoff, A. Dengel and M. Petridis. *Advanced Similarity Measures Using Word Embeddings and Siamese Networks in CBR*. in *Intelligent Systems and Applications*. 2020. Cham: Springer International Publishing.

[20] Baresi, L., G. Quattrocchi, D.A. Tamburri and W.-J. Van Den Heuvel. *Automated Quality Assessment of Incident Tickets for Smart Service Continuity*. in *Service-Oriented Computing*. 2020. Cham: Springer International Publishing.

[21] Han, J. and A. Sun. *DeepRouting: A Deep Neural Network Approach for Ticket Routing in Expert Network*. in *2020 IEEE International Conference on Services Computing (SCC)*. 2020.

[22] Mour, V., S. Dey, S. Jain and R. Lodhe. *Feature Store for Enhanced Explainability in Support Ticket Classification*. in *Natural Language Processing and Chinese Computing*. 2020. Cham: Springer International Publishing.

[23] Revina, A., K. Buza and V.G. Meister, *IT Ticket Classification: The Simpler, the Better.* IEEE Access, 2020. **8**: p. 193380-193395.

[24] Xu, J., J. Mu and G. Chen, *A multi-view similarity measure framework for trouble ticket mining.* Data & Knowledge Engineering, 2020. **127**: p. 101800.

[25] Lo, D., K.K. Tiba, S. Buciumas and F. Ziller. *An Emperical Study on Application of Big Data Analytics to Automate Service Desk Business Process*. in *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*. 2019.

[26] Mukunthan, M.A. and S. Selvakumar, *Multilevel Petri net-based ticket assignment and IT management for improved IT organization support.* Concurrency and Computation: Practice and Experience, 2019. **31**(14): p. e5297.

[27] Misra, J. and S. Podder. *Mining Text in Incident Repositories: Experiences and Perspectives on Adopting Machine Learning Solutions in Practice*. in *2019 34th IEEE/ACM International Conference on Automated Software Engineering Workshop (ASEW)*. 2019.

[28] Palacios, D., C. Morillas, M. Garcés and P. Tapia. *Big data-empowered system for automatic trouble ticket generation in IoT networks*. in *2019 IEEE 2nd 5G World Forum (5GWF)*. 2019.
[29] Shanmugalingam, K., N. Chandrasekara, C. Hindle, G. Fernando and C. Gunawardhana. *Corporate IT-Support Help-Desk Process Hybrid-Automation Solution with Machine Learning Approach*. in *2019 Digital Image Computing: Techniques and Applications (DICTA)*. 2019.
[30] Werner, C., Z.S. Li and D. Damian, *Can a Machine Learn Through Customer Sentiment?: A Cost-Aware Approach to Predict Support Ticket Escalations.* IEEE Software, 2019. **36**(5): p. 38-45.
[31] Gajananan, K., P. Loyola, Y. Katsuno, A. Munawar, S. Trent and F. Satoh. *Modeling Sentiment Polarity in Support Ticket Data for Predicting Cloud Service Subscription Renewal*. in *2018 IEEE International Conference on Services Computing (SCC)*. 2018.
[32] Han, R., V.K.H. Goh, A. Sun and M. Akbari, *Towards Effective Extraction and Linking of Software Mentions from User-Generated Support Tickets*. 2018. 2263-2271.
[33] Koehler, J., E. Fux, F.A. Herzog, D. Lötscher, K. Waelti, R. Imoberdorf and D. Budke. *Towards Intelligent Process Support for Customer Service Desks: Extracting Problem Descriptions from Noisy and Multi-lingual Texts*. in *Business Process Management Workshops*. 2018. Cham: Springer International Publishing.
[34] Lyubinets, V., T. Boiko and D. Nicholas. *Automated Labeling of Bugs and Tickets Using Attention-Based Mechanisms in Recurrent Neural Networks*. in *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*. 2018.
[35] Meng, F.J., J. Xu, X. Zhang, L. Yang, P. Chen, Y. Wang, X. Liu, N. Ayachitula, K. Murthy, L. Shwartz, G. Galambos, Z. Su and J. Zheng. *Opportunities and Challenges Towards Cognitive IT Service Management in Real World*. in *2018 IEEE Symposium on Service-Oriented System Engineering (SOSE)*. 2018.
[36] Paramesh, S.P., C. Ramya and K.S. Shreedhara. *Classifying the Unstructured IT Service Desk Tickets Using Ensemble of Classifiers*. in *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS)*. 2018.
[37] Parmar, P.S., P.K. Biju, M. Shankar and N. Kadiresan. *Multiclass Text Classification and Analytics for Improving Customer Support Response through different Classifiers*. in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 2018.
[38] Patidar, M., P. Agarwal, L. Vig and G. Shroff, *Automatic Conversational Helpdesk Solution using Seq2Seq and Slot-filling Models*, in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. 2018, Association for Computing Machinery: Torino, Italy. p. 1967–1975.
[39] Saberi, M., M. Theobald, O.K. Hussain, E. Chang and F.K. Hussain, *Interactive feature selection for efficient customer recognition in contact centers: Dealing with common names.* Expert Systems with Applications, 2018. **113**: p. 356-376.
[40] Zhou, W., X. Zhu, T. Li, L. Shwartz and G.Y. Grabarnik. *Multi-view feature selection for labeling noisy ticket data*. in *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*. 2018.
[41] Xu, J., H. Zhang, W. Zhou, R. He and T. Li, *Signature based trouble ticket classification.* Future Generation Computer Systems, 2018. **78**: p. 41-58.
[42] Giurgiu, I., D. Wiesmann, J. Bogojeska, D. Lanyi, G. Stark, R.B. Wallace, M.M. Pereira and A.A. Hidalgo, *On the adoption and impact of predictive analytics for server incident reduction.* IBM Journal of Research and Development, 2017. **61**(1): p. 9:98-9:109.
[43] Montgomery, L. and D. Damian. *What do Support Analysts Know About Their Customers? On the Study and Prediction of Support Ticket Escalations in Large Software Organizations*. in *2017 IEEE 25th International Requirements Engineering Conference (RE)*. 2017.
[44] Reddy, Y., C. Reddy T and P. Dayaker, *To dialogue with chatbot using machine learning.* International Journal of Mechanical Engineering and Technology, 2017. **8**: p. 729-739.
[45] Goby, N., T. Brandt, S. Feuerriegel and D. Neumann. *Business Intelligence for Business Processes: the Case of IT Incident Management*. in *ECIS*. 2016.
[46] Corbin, J.M. and A. Strauss, *Grounded theory research: Procedures, canons, and evaluative criteria.* Qualitative Sociology, 1990. **13**(1): p. 3-21.
[47] Jiang, C., Q. Su, L. Zhang and B. Huang, *Automatic Question Answering System Based on Convolutional Neural Network and Its Application to Waste Collection System.* Journal of Circuits, Systems and Computers, 2021. **30**(01): p. 2150013.
[48] Shevchenko, A.P., R.A. Eremin and V.A. Blatov, *The CSD and knowledge databases: from answers to questions.* CrystEngComm, 2020. **22**(43): p. 7298-7307.
[49] Gherheș, V. and C. Obrad, *Technical and Humanities Students' Perspectives on the Development and Sustainability of Artificial Intelligence (AI).* Sustainability, 2018. **10**(9): p. 3066.
[50] Morikawa, M., *Firms' Expectations about the Impact of AI and Robotics: Evidence from a Survey.* Economic Inquiry, 2017. **55**(2): p. 1054-1063.