Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 246 (2024) 128–137

Procedia Computer Science

www.elsevier.com/locate/procedia

28th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2024)

Enhancing Customer Support in Banking: Leveraging AI for Efficient Ticket Classification

Simone Bonechi[a,*], Giulia Palma[a], Mario Caronna[a], Massimiliano Ugolini[b], Alessandra Massaro[b], Antonio Rizzo[a]

[a]Department of Social, Political and Cognitive Science, University of Siena, Siena, Italy
[b]Monte dei Paschi di Siena, Siena, Italy

Abstract

In an era characterized by rapid technological advancement and rising customer expectations, accurate ticket classification in banking customer service emerges as a critical necessity. In this context, we designed a comprehensive ticket classification pipeline, leveraging a real-world dataset comprising 4,243 chat-based user requests, classified into ten distinct classes, provided by MPS Bank. Our approach proposes a complete data processing pipeline with an exploration of two text classification methodologies: BERT (Bidirectional Encoder Representations from Transformers) and TF-IDF (Term Frequency - Inverse Document Frequency) with SVM (Support Vector Machine). The experiments highlight that both models have considerable potential, promising substantial improvements in the operational efficiency of customer support, ultimately increasing the overall quality of service.

Keywords: Deep Learning; Natural Language Processing; Ticket Classification; Customer Service Efficiency

1. Introduction

In these years, dominated by web and smartphone applications that allow direct interaction with a multitude of public and private services, the importance of support tickets and incident reports in fostering communication between customers and service providers cannot be overstated [1]. On the one hand, these channels are designed to offer customers a simplified tool to promptly report any problems encountered while using a product or service. At the same time, service providers can leverage these mechanisms to quickly respond to user concerns, ensuring a higher level of customer satisfaction, greater overall productivity, and precise adherence to Service Level Agreements (SLAs). Fast and effective problem resolution not only strengthens customer loyalty and satisfaction but also minimizes any

* Corresponding author.
E-mail address: simone.bonechi@unisi.it

10.1016/j.procs.2024.09.235

potential negative impact on the business. By proactively managing and resolving issues, service providers can maintain a positive brand image, fostering long-term customer relationships. Furthermore, this approach demonstrates a commitment to excellence and responsiveness, which ultimately strengthens the service provider's reputation in the marketplace. This means that a proactive approach to handling support tickets and incident reports not only benefits the customer but also contributes significantly to the overall success and growth of the business.

In this context, automatic ticket classification could allow to organize and categorize customer service inquiries according to the type of issue that is being addressed. This helps businesses prioritize their customer inquiries, ensuring that the most urgent requests are addressed quickly and accurately. In fact, ticket classification allows a company to quickly identify the best team to respond to a customer's request, without the need to manually analyze unstructured ticket data. The interest in this topic has increased in recent years, driven by the exponential growth of online stores. As consumers increasingly rely on online stores as an alternative to physical stores, the importance of effective customer service in this environment has become fundamental. Indeed, most companies strive to raise the quality of their support services, ensuring rapid response times and simplified problem resolution with minimal procedural complexities. Thus, virtual customer support systems have emerged as vital components in the support operations of the organization. Among them, the automatic classification of tickets, opened by customers, is of significant importance in dealing effectively with end-user issues. Furthermore, in addition to e-commerce, large organizations such as universities and municipal offices could also benefit from an efficient ticket management system. Solving the problem while we are at work or at home has become crucial in our age and efficiently handling user requests has great importance in terms of user satisfaction and cost reduction. A reliable ticketing system is essential also in this context to enable organizations to deliver world-class support services while meeting the ever-changing preferences and demands of their customer base.

Service providers offer customers a range of communication channels to connect with them and open a ticket, including emails, web forms, phone calls, live chats, and social media platforms [25]. These interactions yield valuable textual data which, however, due to their conversational style, often present a challenge when automatically parsed, as they are typically composed of concise and noisy descriptions of the problem. To speed up the ticket management system, over the years, experts have suggested automating the various steps involved in this process. These include:

- Ticket classification — Categorizing tickets into general topics [2])
- Expert search — Assigning the problem to an expert who can solve it [23]
- Ticket resolution — Automatically resolve tickets [33])

In the context of banking customer service, there are common challenges to face. These include handling high volumes of requests, ensuring quick and accurate responses, and managing the complexity of customer interactions. These factors underlined the need for the development of advanced ticket classification models to improve operational efficiency and improve the complex customer experience.

In this scenario, the bank Monte dei Paschi di Siena (MPS) has recently developed a new innovative internal assistance model, based on an integrated cognitive system (machine-human) to interpret natural language and provide support to the network of branch offices on commercial, application, process, or technological issues. The new model started with two main Artificial Intelligence (AI) components based on Machine Learning market products. The first component, named "Assistant", is focused on the semantic analysis of the user request to understand the content of the text and check the availability of a solution to be provided directly to the user. The second component, named "Classifier", is designed to receive requests not recognized by the Assistant and address the requests to the correct service team.

Over time, the evolution of the classifier has led to changes in the model and hyperparameters, negatively affecting the performance of the new service process. In the meantime, the knowledge managed by the Assistant has grown significantly, allowing to increase the number of requests managed directly by this component and consequently reducing the number of requests passed to the classifier. However, to handle the remaining requests not covered by the Assistant, the project team decided to replace the Classifier with a solution where the end user chooses the scope of the request (e.g. payments, PCs, printers, credit cards, etc.). The purpose of this work is to verify the effectiveness of different approaches, to automatically classify incoming requests, reduce the manual effort of the end user, and improve the overall efficiency of the system. For this reason, we mainly focused on ticket classification comparing various classification approaches based on Machine Learning.

Machine Learning and Deep Learning have recently achieved tremendous success across various fields, including computer vision [15, 13, 7], biomedicine [10, 6, 28, 8], and natural language processing [31, 24]. In particular, transformer-based encoder models like BERT [12] drastically outperform the previous state-of-the-art for text understanding. In our work, we first evaluate different methods taking into consideration the nature of the available dataset and the related literature of this field. Based on this preliminary study, we decided to compare two distinct approaches. The first approach is based on BERT, a powerful language model, to represent and classify the text of the ticket. Indeed, BERT's deep contextual understanding makes it well-suited for this task. In contrast, the second approach employs a TF-IDF (Term Frequency, Inverse Document Frequency) technique to represent the text, complemented by an SVM (Support Vector Machine) [9] for classification. The selection of BERT and SVM models for experimentation was driven by their established effectiveness in handling natural language processing tasks, particularly in domains with complex and varied linguistic patterns such as banking customer service.

Data augmentation techniques have been employed to artificially enlarge the training dataset. The effectiveness of the trained models was evaluated in terms of prediction accuracy and experiments reveal that BERT enables ticket classification with better performance than the SVM approach. However, it is interesting to note that the difference between these two methods is less pronounced than our initial expectations. This could likely be due to the intrinsic characteristics of the problem and the size of the dataset. Indeed, in our specific case study, the sentences are relatively concise and the correct classification of a category may strongly depend on the presence or absence of specific keywords. Additionally, it is worth noting that our dataset includes only 4,243 tickets, which may not be enough to properly train a BERT classifier.

This work is organized as follows: Section 2 revised the literature connected to text classification and in particular ticket classification. Section 3 describes the dataset and employed classification methods and Section 3 describes the dataset and the experimental setups. Instead, Section 4 presents and discusses the obtained results. Finally, Section 5 draws conclusions and discusses possible future developments.

2. Related works

This section collects the literature related to the two main topics studied in this work. In particular, in Section 2.1 the main contributions in text classification are briefly described; instead, Section 2.2 reviews the main approaches for ticket classification.

2.1. Text Classification

Numerous approaches to text classification have been proposed over the years, including vector space-based methods [29, 30, 32] and topic-based models [5, 22, 4]. Vector space models involve reducing each document in the corpus to a vector of real numbers, where each dimension represents the counts of the words in a dictionary. The TF-IDF approach [29] is the most famous and also the most used vector space method. In this approach, each dimension of the feature vector computes the term frequency count adjusted by the inverse document frequency count. In this way, the most common words in the entire corpus assume a low value allowing us to effectively identify a subset of discriminative terms for documents within the corpus. Although simple and efficient, TF-IDF has limitations in capturing correlations between terms.

In recent years, thanks to the great advances in Natural Language Processing (NLP), automatic Text Classification (TC) has dramatically increased its performance. This breakthrough was largely driven by the introduction of the Transformer architecture [31], which revolutionized many language understanding tasks, including TC. By leveraging contextualized language models, such as BERT [12], words can now be represented more meaningfully, allowing better context comprehension and leading to substantial performance boosts in TC benchmarks. BERT-based LMs have found applications in various NLP tasks, including user-generated content analysis such as tweets [27] and user reviews [17]. While BERT has shown sensitivity to text noise, such as spelling mistakes [16], it remains widely utilized and studied. Additionally, it has been shown to be resistant to label noise, and it was demonstrated that trying to mitigate the effects of label noise can actually reduce BERT performance [35].

2.2. Ticket Classification

Recently, there has been growing interest in the task of topical Ticket Classification (TiC). Traditional methods, like Support Vector Machines (SVMs) [9] applied on TF-IDF representations, have been extensively explored for TiC. However, more sophisticated text representations, such as Word2Vec [20], proved to be able to enhance TiC performance [2]). Furthermore, deep learning approaches to TiC, which employed MultiLayer Perceptron [14], Convolutional Neural Networks [36, 26], and Recurrent Neural Networks [19, 18], proved to be successful in capturing complex patterns from the ticket data. Another study [3] proposes using a two-phase classifier based on machine learning to categorize and sub-categorize support tickets in an Issue Tracking Systems, automatically assigning them to the right team. The COTA system [21], leveraging deep learning using an encoder-decoder architecture, integrates with the Uber support platform to improve service by identifying ticket types and recommending appropriate response patterns. In the context of car dealership customer service, [34] presents a deep learning approach to automatically classify phone call requests, facilitating faster routing and problem resolution.

In this article, we aim to address these gaps by exploring the possibility of leveraging Large Language Models (LLM), particularly BERT-based LLMs, for Ticket Classification in the domain of banking customer support. By integrating state-of-the-art machine learning techniques with domain-specific knowledge and practical insights, we seek to advance the field and provide tangible solutions for improving the efficiency and effectiveness of customer support operations in banking institutions.

3. Materials and Methods

The dataset utilized in this study and the data pre-processing steps are presented in Section 3.1 and Section 3.2, respectively. While Section 3.3 describes the text representation employed as input for the machine learning tools. Finally, the proposed classification methods are detailed in Section 3.4.

3.1. Datasets

In this study, we used a private dataset comprising 4,243 tickets provided by the MPS bank. The dataset has been made available to us in CSV format where all references to the user have been removed to maintain privacy. Each ticket represents a chat-based assistance request written in Italian that is composed of two key components:

- Dialog — the description of the issue made by the users. This is the unstructured text that we want to categorize
- Topic — the actual scope of the ticket. This is the target that we want to predict. In the dataset, we have a total of 10 topics, Figure 1 shows the distribution of samples among the 10 classes.

It should be noted that since the data represent a real case scenario it may contain errors, repetitions, and symbols.

The dataset was split into training, validation, and test sets, with 75% of the samples assigned to training, 10% to validation, and the remainder to the test set.

3.2. Data pre-processing

To effectively employ a machine learning approach that automatically classifies text, some preprocessing steps are usually required. Depending on the type of approach we want to use, we could do different types of text preprocessing.

- Misspelled Word Correction — Identify and correct misspelled words to improve data accuracy and consistency
- Stop word removal — Filter out common words such as “a”, “the”, and “of” that have little semantic meaning, reducing noise in the data
- Tokenization — Breaking sentences or paragraphs into individual tokens (words or subwords) for further analysis and processing
- Stemming — Reduction of words to their root form to ensure that variations of the same word are treated as a single entity

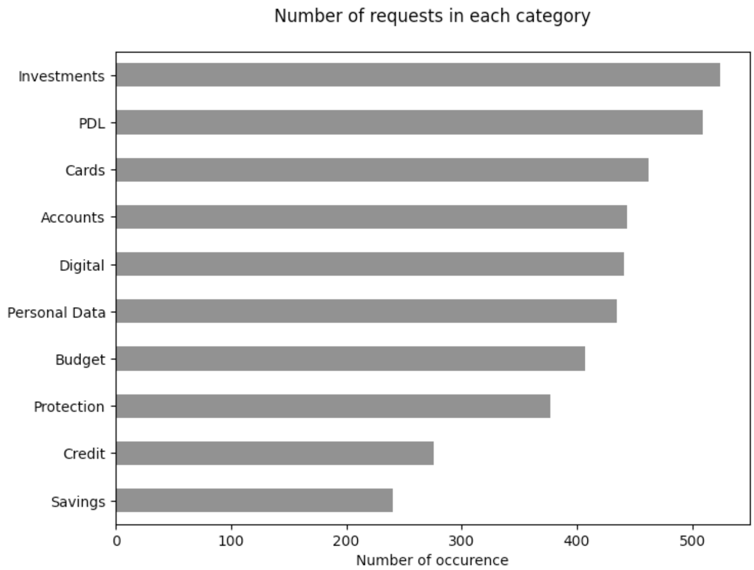

Fig. 1. Distribution of tickets among the 10 classes.

- Lemmatization — Similar to stemming, lemmatization aims to transform words into their base form but ensures that the resulting word is valid, offering better semantic interpretation

BERT requires minimal pre-processing to be employed, indeed we mainly addressed out-of-context symbols, which resulted from typing errors or conversion issues between ASCII and UTF-8 formats, and we converted all texts to lowercase.

On the other hand, traditional approaches to text classification can also benefit from the removal of punctuation and stop words. In the approach using TF-IDF and SVM, we apply stop word removal using the Italian stop word list provided by the python nltk library. This process eliminates common and uninformative words from text data, resulting in more effective and efficient classification.

3.3. Text Representation

3.3.1. BERT tokenizer

The BERT model requires tokenized input sentences for its processing. To achieve this, we used the pre-trained BERTTokenizer provided by the Huggingface library. This tokenizer is based on WordPiece's tokenizer, which allows it to effectively split a sentence into multiple tokens. The resulting textual representation is presented as a list of the IDs corresponding to each token. In this study, we conducted experiments using various numbers of tokens to represent sentences, specifically exploring the use of 64, 128, and 256.

3.3.2. TF-IDF

The TF-IDF is used to create the input vector for the SVM classifier. In particular, TF-IDF allows to represent the text of a ticket as a histogram, capturing the normalized frequency of each word in the text. To achieve this, we first create a dictionary and then we compute the frequency of each word of the dictionary in the training data. For each ticket, we calculate the number of occurrences of each word from the dictionary and then normalize it. Normalization is performed by dividing the occurrences of words in the text by the relative frequency calculated on the training set. This process ensures that the resulting histogram for each ticket reflects the importance of each word in relation to the entire training set.

To compute the TF-IDF vector representation we used the TfidfVectorizer and TfidfTransformer methods from the Sklearn library. To build the dictionary we set the minimum word frequency to 2 ensuring that words appearing in less

	Category	Most Correlated Unigrams	Most Correlated Bigrams
0	Accounts	pacchetto, corrente, conto	un conto, mps mio, conto corrente
1	Budget	voce, partita, partitario	tra banche, una scrittura, una partita
2	Cards	prepagata, credito, carta	una carta, carta di, di credito
3	Credit	fido, pef, elise	di mutuo, un mutuo, in elise
4	Digital	pao, digital, banking	codice sia, il digital, digital banking
5	Investments	fondi, consulenza, athena	su athena, gestione patrimoniale, in athena
6	PDL	aziendale, stampante, pc	posta elettronica, di filiale, cellulare azien...
7	Personal Data	poteri, firma, kyc	un kyc, poteri di, di firma
8	Protection	auto, protezione, polizza	polizza auto, mia protezione, una polizza
9	Savings	attiva, riscatto, previdenza	previdenza per, fondo pensione, previdenza attiva

Fig. 2. Identifying Highly Correlated Unigrams and Bigrams by Category.

than 2 documents were not considered, this way we were able to discard rare words that potentially do not interest our purpose. Furthermore, we conducted experiments with dictionaries of various sizes. Specifically, we explored four dictionaries, each containing a different number of words: 7,000, 8,000, 9,000, and 10,000 respectively.

An advantage of the TF-IDF representation is its interpretability. In fact, for example, it is possible to easily identify the most distinctive words for each label (see Figure 2), facilitating a more in-depth analysis of the dataset.

3.4. Classification approaches

In this work, we experiment with two text classification approaches to categorize the tickets. The first approach is based on BERT, a cutting-edge deep learning model, while the other explores classical machine learning algorithms such as SVM.

On the one hand, we decided to use BERT, a transformer-based neural network model, which normally requires a huge dataset and computational power to be trained. Transformers are a major breakthrough that leverages a self-attention mechanism, assigning varying weights to different parts of the input data based on their significance. Transformers find wide applications in Natural Language Processing (NLP) as they are designed to process sequential data. During our experiment, we used the HuggingFace Transformers library and we explored the use of two different pre-trained models, BERT base Italian XXL uncased and BERT base Multilingual uncased.

On the other hand, we chose SVM, a classic machine learning classifier, less powerful than transformers but faster and computationally lighter. SVMs are kernel-based algorithm, which seeks to find an optimal boundary between possible outputs by transforming input data into a higher-dimensional space.

Both methods offer distinct advantages and disadvantages when applied to multiclass classification for support requests. SVM stands out for its computational efficiency, offering speed in processing. In contrast, BERT excels at acquiring semantic and contextual information but its implementation may require more computational resources and data availability than SVM. By carefully evaluating the strengths and limitations of each approach, we aim to determine the most appropriate method for our ticket evaluation task.

4. Experiments and Results

In this section, we detail the experimental setups for both the BERT and TF-IDF with SVM approaches in Sections 4.1 and 4.2, respectively. Section 4.3 then presents and compares the performance of these two approaches across the various setups.

4.1. BERT setup

The experimental setup was meticulously designed to encompass a comprehensive exploration of various methodologies and techniques in ticket classification. Specific parameters and configurations were carefully chosen to ensure the robustness and reproducibility of the experiments. To alleviate the lack of training data and improve the BERT generalization ability to avoid overfitting, we implemented various data augmentation techniques. Specifically, we leveraged the Python nlpaug library to perform the following augmentations:

- Synonym: Randomly replaced some words in the sentence with synonyms.
- Swap: Randomly swapped the positions of some words in the sentence.
- Delete: Randomly deleted some words in the sentence.
- Contextual Word Embedding: Randomly replaced words with others that had similar embeddings obtained using GPT-2.

Furthermore, we used an additional translation-based augmentation strategy. The selected data augmentation techniques for BERT training were chosen to enhance the model's generalization, enabling better handling of the variability in natural language present in banking customer service data. In this case, we used the Meta NLLB (No Language Left Behind) translator [11] to first translate the sentence into English and then back into Italian. This process produced a sentence that retained its semantic meaning but used different words. These strategies allow us to create an augmented version of the training dataset containing 18,498 samples. Both the Italian and multilingual BERT models were trained using both versions of the training dataset (i.e. the augmented and non-augmented dataset). The training also involved experimenting with various input lengths (i.e. 64, 128, and 256 tokens). All models were trained with the following setup and evaluated on the validation set to select the optimal configuration for our application.

- Optimizer: AdamW
- Learning rate: 2e-5 (0.00002)
- Epochs: 500
- Patience: 30
- Batch size: 30

The validation set was also used to early-stop the training process to avoid overfitting.

4.2. SVM with TF-IDF setup

We employed the SVM implementation provided by the Python library scikit-learn. In this configuration, we combined training and validation sets, conducting hyperparameter selection through a grid-search approach using 5-fold cross-validation. Specifically, we defined the following hyperparameter grid:

- Kernel: rbf, poly, linear, sigmoid
- Gamma: auto, 1e-1, 1e-2, 1e-3, 1e-4, 1e-5
- C: 0.00001, 0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000
- degree: 1, 2, 3, 4
- Vocabulary size: 7,000, 8,000, 9,000, 10,000
- Class weight: balanced

4.3. Results

We trained the BERT models following the experimental setup outlined in Section 4.1. The results, obtained with various configurations, on the validation set were shown in Table 1.

Although the performance improvement is modest, with an increase of less than 1% in accuracy across all setups, it appears that the augmentation strategies contribute positively during the training process. Indeed, as can be observed

Dataset	Non Augmented			Augmented		
Input dim.	64	128	256	64	128	256
Italian	66.66%	77.13%	85.95%	66.29%	77.77%	**86.19%**
Multilingual	65.37%	76.06%	84.52%	67.22%	75.85%	84.52%

Table 1. Validation accuracy of Italian and multilingual BERT as the input size varies.

from these results, the highest performance is achieved when training the pre-trained Italian BERT model on the augmented training set with an input size of 256 tokens.

Instead, we trained the SVM model using the TF-IDF representation, following the experimental setup detailed in Section 4.2. Table 2 presents the results obtained at the conclusion of the grid-search hyperparameters selection based on 5-fold cross-validation, showing changes in performance as the dictionary size varies.

Dictionary Size	Mean Accuracy	Standard Deviation
7,000	**84.70%**	**0.016%**
8,000	84.62%	0.021%
9,000	84.62%	0.018%
1,0000	84.45%	0.022%

Table 2. Cross-Validation accuracy of the SVM model as the dictionary size varies.

All models showed comparable performance during cross-validation. As a result, we chose to use the dictionary containing 7,000 words, as it allowed us to achieve the best performance while keeping the dictionary size modest. This reduces the size of the input and consequently alleviates the computational overhead. The best hyperparameters of the model trained with the selected dictionary are the following:

- Kernel: rbf
- Gamma: 0.1
- C: 10
- Class weight: balanced

Finally, we tested the best BERT model and the best SVM model on the test set. The results, in terms of accuracy, are reported in Table 3, instead the confusion matrices are depicted in Figures 3.

Model	Accuracy
BERT	**85.88%**
SVM	82.42%

Table 3. Accuracy of the SVM and of the BERT model on the test set.

The results obtained on the test set validate the superiority of the BERT-based classification approach, demonstrating its ability to achieve higher accuracy compared to TF-IDF and SVM. In particular, the integration of this approach into MPS's customer support pipeline has the potential to significantly improve their responsiveness to user issues, which signifies a promising advancement in customer support effectiveness.

The increased accuracy of the BERT model suggests that it can significantly improve ticket classification. This results in faster response times and greater customer satisfaction. Furthermore, a deeper analysis of the findings reveals broader implications for MPS Bank's business practices. By leveraging advanced machine learning like BERT, the bank can optimize resource allocation, streamline operational efficiency and gain a competitive advantage in the rapidly evolving banking landscape. Conversely, a decline in classifier performance could lead to delays in resolving customer requests, potentially causing frustration. Additionally, misclassified tickets could lead to misallocation of resources, further prolonging resolution times and impacting customer satisfaction. These delays and errors have the potential to damage MPS Bank's reputation and erode trust in the brand.

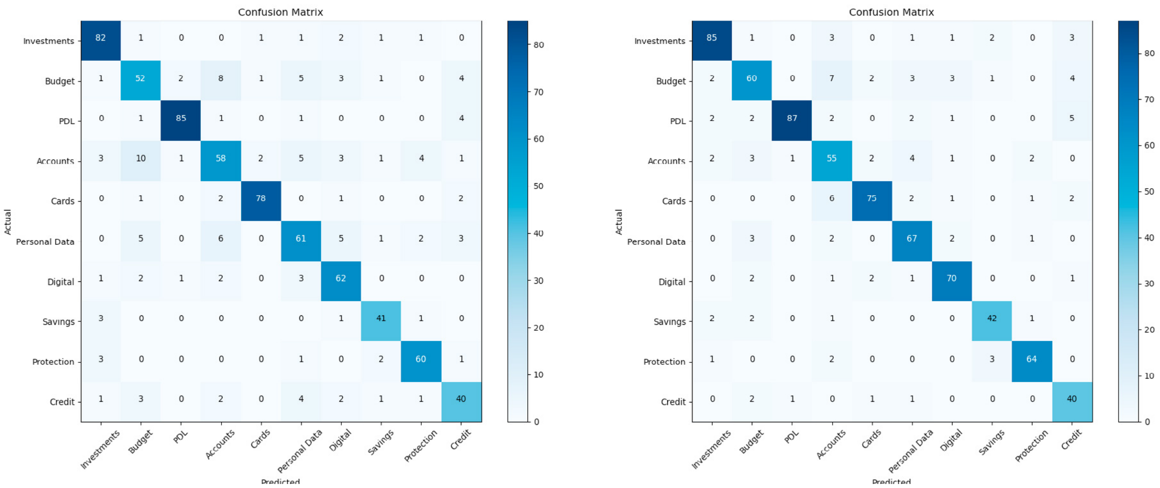

Fig. 3. Confusion matrix obtained with the best SVM model (left) and with the best BERT model (right) on the test set.

5. Conclusion

This study delves into the critical task of ticket classification, an indispensable aspect of modern enterprise customer service operations. Support ticket classification plays a central role in streamlining your organization and prioritizing customer requests, thus helping to improve customer satisfaction and operational efficiency. In this work, we examined various methods of text representation, including the use of BERT-based language models and traditional TF-IDF techniques. Through systematic experimentation, we evaluated the effectiveness of these approaches in accurately categorizing support tickets. As expected, the BERT-based method demonstrated superior performance compared to the TF-IDF approach, taking advantage of its ability to incorporate contextual information. Interestingly, however, the performance disparity between the two approaches is not overly pronounced. The simplicity and efficiency of the TF-IDF approach could, therefore, emerge as a pragmatic choice, particularly for organizations with limited computational resources. The effectiveness of the proposed approaches was validated in a real scenario presented by MPS bank. The results indicate that implementing these methods has the potential to increase the effectiveness of their ticketing system, reduce response times, and ultimately drive customer satisfaction to unprecedented levels.

However, the ever-evolving nature of banking customer service requires continuous adaptation of support models. It will be a matter of future research to optimize the performance of our model improving robustness and generalization. Fine-tuning the model with domain-specific financial terminology can improve its ability to handle different customer requests and linguistic variations. Additionally, incorporating customer feedback into the model can further refine its effectiveness and streamline banking customer service operations.

References

[1] Al-Hawari, F., Barham, H., 2021. A machine learning based help desk system for it service management. Journal of King Saud University - Computer and Information Sciences, 33(6):702–718 .

[2] Aleksandra Revina, K.B., Meister, V.G., 2020. It ticket classification: The simpler, the better. IEEE Access, 8:193380–193395 .

[3] Altintas, M., Tantug, A.C., 2014. Machine learning based ticket classification in issue tracking systems, in: Proceeding Int. Conf. Artif. Intell. Comput. Sci.(AICS 2014), pp. 195–207.

[4] Bai, B., Weston, J., Grangier, D., Collobert, R., Sadamasa, K., Qi, Y., Chapelle, O., Weinberger, K., 2009. Supervised semantic indexing, in: Proceedings of the 18th ACM conference on Information and knowledge management, pp. 187–196.

[5] Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. Journal of machine Learning research 3, 993–1022.

[6] Bonechi, S., 2023. Isic_wsm: Generating weak segmentation maps for the isic archive. Neurocomputing 523, 69–80.

[7] Bonechi, S., Andreini, P., Bianchini, M., Scarselli, F., 2018. Generating bounding box supervision for semantic segmentation with deep learning, in: Artificial Neural Networks in Pattern Recognition: 8th IAPR TC3 Workshop, ANNPR 2018, Siena, Italy, September 19–21, 2018, Proceedings 8, Springer. pp. 190–200.

[8] Bonechi, S., Bianchini, M., Bongini, P., Ciano, G., Giacomini, G., Rosai, R., Tognetti, L., Rossi, A., Andreini, P., 2019. Fusion of visual and anamnestic data for the classification of skin lesions with deep learning, in: New Trends in Image Analysis and Processing–ICIAP 2019: ICIAP International Workshops, BioFor, PatReCH, e-BADLE, DeepRetail, and Industrial Session, Trento, Italy, September 9–10, 2019, Revised Selected Papers 20, Springer. pp. 211–219.
[9] Boser, B.E., Guyon, I.M., Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers, in: Proceedings of the fifth annual workshop on Computational learning theory, pp. 144–152.
[10] Ching, T., Himmelstein, D.S., Beaulieu-Jones, B.K., Kalinin, A.A., Do, B.T., Way, G.P., Ferrero, E., Agapow, P.M., Zietz, M., Hoffman, M.M., et al., 2018. Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface 15, 20170387.
[11] Costa-jussà, M.R., Cross, J., Çelebi, O., Elbayad, M., Heafield, K., Heffernan, K., Kalbassi, E., Lam, J., Licht, D., Maillard, J., et al., 2022. No language left behind: Scaling human-centered machine translation. arXiv preprint arXiv:2207.04672 .
[12] Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 .
[13] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. Advances in neural information processing systems 27.
[14] Kallis, R., Di Sorbo, A., Canfora, G., Panichella, S., 2019. Ticket tagger: Machine learning driven issue classification, in: 2019 IEEE International Conference on Software Maintenance and Evolution (ICSME), IEEE. pp. 406–409.
[15] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 25, 1097–1105.
[16] Kumar, A., Makhija, P., Gupta, A., 2020. Noisy text data: Achilles' heel of bert. arXiv preprint arXiv:2003.12932 .
[17] Lu, Z., Du, P., Nie, J.Y., 2020. Vgcn-bert: augmenting bert with graph embedding for text classification, in: Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14–17, 2020, Proceedings, Part I 42, Springer. pp. 369–382.
[18] Lyubinets, V., Boiko, T., Nicholas, D., 2018. 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), IEEE. pp. 271–275.
[19] Mani, S., Sankaran, A., Aralikatte, R., 2019. Deeptriage: Exploring the effectiveness of deep learning for bug triaging, in: Proceedings of the ACM India joint international conference on data science and management of data, pp. 171–179.
[20] Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781 .
[21] Molino, P., Zheng, H., Wang, Y.C., 2018. Cota: Improving the speed and accuracy of customer support through ranking and deep networks, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 586–595.
[22] Nigam, K., McCallum, A.K., Thrun, S., Mitchell, T., 2000. Text classification from labeled and unlabeled documents using em. Machine learning 39, 103–134.
[23] Omayma Husain, Naomie Salim, R.A.A.S.A., Hassan, A., 2019. Expert finding systems: A systematic review. Applied Sciences, 9(20) .
[24] Otter, D.W., Medina, J.R., Kalita, J.K., 2020. A survey of the usages of deep learning for natural language processing. IEEE transactions on neural networks and learning systems 32, 604–624.
[25] Paolo Zicari, Gianluigi Folino, M.G., Pontieri, L., 2021. Discovering accurate deep learning based predictive models for automatic customer support ticket classification. Proceedings of the 36th Annual ACM Symposium on Applied Computing, SAC '21, pages 1098–1101, New York, NY, USA. Association for Computing Machinery .
[26] Pistellato, M., Cosmo, L., Bergamasco, F., Gasparetto, A., Albarelli, A., 2018. Adaptive albedo compensation for accurate phase-shift coding, in: 2018 24th International Conference on Pattern Recognition (ICPR), IEEE. pp. 2450–2455.
[27] Polignano, M., Basile, P., De Gemmis, M., Semeraro, G., Basile, V., et al., 2019. Alberto: Italian bert language understanding model for nlp challenging tasks based on tweets, in: CEUR Workshop Proceedings, CEUR. pp. 1–6.
[28] Rossi, A., Vannuccini, G., Andreini, P., Bonechi, S., Giacomini, G., Scarselli, F., Bianchini, M., 2019. Analysis of brain nmr images for age estimation with deep learning. Procedia Computer Science 159, 981–989.
[29] Salton, G., Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. Information processing & management 24, 513–523.
[30] Singhal, A., Buckley, C., Mitra, M., 2017. Pivoted document length normalization, in: Acm sigir forum, ACM New York, NY, USA. pp. 176–184.
[31] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Advances in neural information processing systems 30.
[32] Weinberger, K.Q., Chapelle, O., 2008. Large margin taxonomy embedding for document categorization. Advances in Neural Information Processing Systems 21.
[33] Wubai Zhou, Wei Xue, R.B.Q.W.C.Z.T.L.J.X.Z.L.L.S., Grabarnik, G.Y., 2017. A system for ticket analysis and resolution. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '17, pages 2181–2190, New York, NY, USA. Association for Computing Machinery .
[34] Zhong, J., Li, W., 2019. Predicting customer call intent by analyzing phone call transcripts based on cnn for multi-class classification. arXiv preprint arXiv:1907.03715 .
[35] Zhu, D., Hedderich, M.A., Zhai, F., Adelani, D.I., Klakow, D., 2022. Is bert robust to label noise? a study on learning with noisy labels in text classification. arXiv preprint arXiv:2204.09371 .
[36] Zicari, P., Folino, G., Guarascio, M., Pontieri, L., 2021. Discovering accurate deep learning based predictive models for automatic customer support ticket classification, in: Proceedings of the 36th annual ACM symposium on applied computing, pp. 1098–1101.