ARAŞTIRMA

Geliş Tarihi: 20 Kasım 2024
Kabul Tarihi: 10 Mart 2025

A SEMANTIC APPROACH TO REGULATION DETECTION IN BANKING DOMAIN

Bankacılık Alanında Düzenleme Tespitine Yönelik Anlamsal Bir Yaklaşım

DOI: 10.58307/kaytek.1587234

Rahim DEHKHARGHANI, Olla HUSSEİN, Ezgi CAN, Laura ORYNGALIYEVA, Hakan KANTAŞ

Özet

Bankacılık Düzenleme ve Denetleme Kurumu'nun (BDDK) bankacılık sistemlerinde önceden belirlenmiş regülasyonları takip etmek, performanslarını iyileştirmelerine yardımcı olabilir. Bir denetim sistemi, BDDK düzenlemelerini izleyerek bir bankada regülasyonların ne kadar takip edildiğini belirleyebilir. Bu izleme şimdiye kadar manuel olarak yapılmıştır ancak bu işlem çok zaman alıcı ve zahmetli bir operasyondur. Bu regülasyonların otomatik olarak izlenmesi önemli miktarda zaman ve emek tasarrufu sağlayabilir. Bu makale, Türkiye'deki bankacılık alanında regülasyon tespiti için bir Semantik yaklaşım önermektedir. Önerilen sistem, anlamsal analizi gerçekleştirmek için YAKE anahtar kelime çıkarıcı ve Cümle Dönüştürücülerle (sentence-transformer'lar) entegre edilmiş Türkçe dili için eğitilmiş bir BERT modelini kullanır. Önerilen sistem için girdi, büyük bir mevzuat koleksiyonu kapsayan bir metin dosyası ve denetlenecek belirli düzenlemeleri içeren bir hedef dosyadır. Analiz edilen metin, BDDK tarafından yayınlanan düzenleyici belgelerden elde edilmiş, hedef metin dosyası ise bir Türk bankası tarafından sağlanmıştır. Önerilen sistem, yeni düzenlemeleri belirlemeye olanak tanır; Ayrıca bir uzmanın önerilen çıktılardan hangisinin doğru hangisinin doğru olmadığını seçmesine olanak tanır, bu da önerilen yaklaşımı yarı-otomatik hale getirir. Sistem, temel görevleri otomatikleştirerek regülasyonları %86 Hassasiyet (Precision) ve %89 Kapsam (Recall) ile belirleyebilmiştir.

Anahtar Kelimeler: *Mevzuat tespiti, Mevzuata uygunluk, anlamsal analiz, Doğal dil işleme, Cümle dönüştürücüler, BERT.*

Abstract

Following pre-established regularities by the Banking Regulation and Supervision Agency (BDDK in Turkey) in the banking systems can help them improve their performance. A supervision system can monitor these BDDK regulatory to determine how much their regularities have been followed in a bank. This monitoring has been done manually so far, which is a laborious task. Automatic monitoring of these regularities can save a significant amount of time and effort. This paper suggests a Semantic approach for regulatory detection in the banking domain in Turkey. The proposed system utilizes a BERT model trained for the Turkish language integrated with the YAKE keyword extractor and Sentence Transformers to perform the semantic analysis. The input for the proposed system is a text file covering a large regulatory collection, and a target file containing certain regulations to be inspected. The analyzed text was sourced from regulatory documents issued by the Banking Regulation and Supervision Agency (BDDK) and the target text file was provided by a Turkish bank. The proposed system allows identifying new regulations and oddities; It also allows an expert to choose which one(s) of the suggested outputs is correct and which is not, which makes the proposed approach semi-automatic. The system could identify regulations with a Precision of 86% and a Recall of 89%, by automating the key tasks.

Keywords: *Regulatory detection, Regulatory compliance, semantic analysis, Natural language processing, Sentence transformers, BERT.*

Rahim DEHKHARGHANI, ahim.dehkharghani@khas.edu.tr, Kadir Has University
Olla HUSSEIN, ollamoe99@stu.khas.edu.tr, Kadir Has University
Ezgi CAN, ezgican@stu.khas.edu.tr, Kadir Has University
Laura ORYNGALIYEVA, laura.oryngaliyeva@stu.khas.edsksu.tr, Kadir Has University
Hakan KANTAŞ, hakan.kantas@ieee.org

INTRODUCTION

Checking financial regulations is a critical process in the banking sector and is considered a legal and mandatory task for the banks. Not to comply these regulations will cause major penalties for the bank. The bank regulations assess in reducing failure probabilities and increase the maturity of the bank system thereby increasing the confidence in it (Cyree, 2016). The Banking Regulation and Supervision Agency (BDDK in Turkey) establishes regulations that Turkish banks must follow. Following these regulations in the banking systems can help them increase their performance level.

Checking if a bank follows these regulations or not, used to be done manually so far, which is a laboring and expensive task. Generally, banks are facing challenges such as the annual updates on regulations in the industry and trying to be updated and consider all regulations (Cyree, 2016). Automatic or semi-automatic monitoring of these regularities can save a large amount of time and effort. To the best of our knowledge, no monitoring system accomplishes this process automatically for the Turkish banks. This issue was our main motivation for this paper.

This goal can be achieved using the state-of-the-art techniques recently proposed in the field of natural language processing (NLP) and deep learning. These techniques have completely transformed how machines process and interact with humans. The NLP field is classified into two main sections: Natural Language Understanding (NLU) and Natural Language Generation (NLG). The NLU branch, which is used in this paper, enables machines to understand and analyze human language, having the ability to extract keywords, ideas, ways of thinking, etc. (Khurana et al., 2023). In contrast, the NLG branch automatically generates text in natural language.

The problem attacked in this paper is to search for a regulation from BDDK, written in Turkish, inside the main document of a bank in Turkey—The bank's name cannot be disclosed due to data privacy and security issues. We used sentence transformers as the main semantic technique for this purpose. Sentence transformers is an embedding technique to semantically represent a sentence, paragraph, or a piece of text. Semantic analysis is necessary as the syntactic level search of regulations in the main document cannot find these regulations if they appear with different words (or in a different order than in the regulations). For example, a regulation such as 'Banka, yılda en az bir defa olmak üzere veya bilgi sistemlerinde meydana gelecek önemli değişikliklerden önce risk analizlerini tekrarlar' (The bank repeats risk analysis at least once a year or before any significant changes in information systems), can be written as 'geçen hafta risk anlizi toplantısı yaptık' (We held a risk analysis meeting last week) in the bank reports. Different forms of the same sentence or paragraph are the main challenges for the current problem. With NLU's different levels, this paper focuses on the semantic level analysis of text. Semantic analysis techniques such as sentence transformers help in the extraction of the deep meaning of text or understanding similar pieces of text that are written in different formats using different words (Khurana et al., 2023).

Our proposed system is similar to a research work which uses rule-based NLP methods for automatically extracting information from construction regulatory documents (Zhang and El-Gohary, 2013). This method involves both syntactic and semantic text features by using several pattern-matching-based and conflict-resolution rules. Their solution improves the extraction process and decreases the number of required patterns by utilizing an ontology for semantic text features and phrase structure grammar (PSG)-based phrasal tags. This method demonstrated its promise for automated regulatory compliance identification by extracting quantitative requirements from the 2009 International Building Code with a high Precision and Recall.

The integration of NLP models with YAKE keyword extractor was also done and explored in previous studies. For example, in (Gupta, etal., 2024) the authors used this combination in sustainability reports. This combination resulted in highly accurate and relevant results, proving the effectiveness of integrating several approaches.

According to Benkassioui et al. (2023), Rule-based methods (RBM) are based on established rules that have been developed by subject-matter experts. While the knowledge required to design the rules may be a barrier for these approaches, their Precision and Recall has historically been excellent, particularly for domain-specific tasks such as finance or law. While machine learning-based methods can be used to a wide range of activities and domains, their training process is mostly dependent on labeled data. Conversely, deep neural network based architectures such as BERT, enable the system to process unstructured data more efficiently and learn from the context. These techniques are appropriate for more complicated regulation documents since they can generalize over sizable datasets. Additionally, Campos et al. (2018) state that YAKE uses the statistical features of the text and be adjusted to different languages, domains, and document sizes.

Moreover, a recent work by Jain et al. (2024) focused on sanctions and Anti-Money Laundering (AML) attempts. The authors investigate how AI technologies, including NLP, machine learning (ML), and predictive analytics, might improve the monitoring, detection, and prevention of regulatory violations. This work utilizes the BERT model for text processing and highlights the accuracy brought by AI to compliance

procedures. The limitations of this study are mentioned as 1) Data privacy: since it is handling massive amount of sensitive data, 2) Ethical implications, 3) AI transparency: considering the black box nature that limits the human understanding of what is going on and the ability to justify the system decisions, and finally 4) Human oversight: to prevent and address issues related to the AI.

Another research work on this topic was accomplished by Aziz and Andriansyah (2023) which explores the use of AI in preventing fraudulent activities, enhancing fraud detection, and monitoring risk management. Deep neural networks are used as a tool for identifying complex patterns and forecasting fraudulent transactions. The authors covered how to use Natural Language Processing (NLP) to improve "Know Your Customer" (KYC) procedures and how to leverage graph analytics to spot questionable transactional interaction. Other AI-driven strategies that are highlighted in this study include geospatial analysis, improved biometric verifications, predictive analytics, and chatbots.

In another work by Al-Shabandar et al. (2019), the application of Artificial Intelligence (AI) in Financial Compliance Management has been investigated. The authors focus on applying different AI and Machine Learning methodologies to manage financial compliance by addressing various challenges such as data privacy. Their work includes creating models to predict potential compliance issues and simulate them as an anomaly detection practice.

Lastly, Calderón (2020), discusses integrating AI in the financial regulation compliance domain and highlights the challenges and potential success, especially in the European Union regulations. The author concluded that AI-driven innovation and beneficial regulatory frameworks working together can greatly enhance regulatory oversight and compliance. Moreover, adopting novel techniques such as data ethnography can also help with technological and legal issues.

As mentioned earlier, despite a few related works on regulation compliance in the literature, to the best of our knowledge, there exist no research work analyzing the same problem for the Turkish language, either on financial or other institutions.

METHOD

In this section, firstly the data, and then different modules of the proposed methodology have been explained in detail.

Data Collection

The input documents used in the proposed system have two types: 1) the target file that involves a specific regulation to be checked and examined, and 2) the main document of a specific bank including all tasks accomplished by the bank. The first file was generated by taking some specific regulations from BDDK: Bankacılık Düzenleme ve Denetleme Kurumu (Banking Regulation and Supervision Agency). For example in our tests, we examined the IT risk management regulations in the bank. The second file covers a wider range of regulations in the bank, and they need to check if they are compliant with the regulations published in BDDK.

Keyword Extraction

The proposed system in this paper utilizes the YAKE model (Campos et al., 2020) for automatic keyword extraction. This model is applied to the target file (the file containing specific regulations to be examined). The keywords are those words that represent the document the most. The YAKE model is an online tool for keyword extraction. The YAKE model uses an unsupervised algorithm without any need for training on a specific document (Campos et al., 2018).

Semantic Similarity Detection

To find the semantic similarities between a pair of sentences or paragraphs, sentence embeddings for each sentence in the target file are created by using the pre-trained BERT model (Devlin et al., 2018), specifically for the Turkish language imported from the 'sentence_transformers' library in Python. The encoding is performed for both input files (the target and the main document) generating dense vectors that are used for the similarity calculations. The mathematical formulations for 'sentence_transformers' are as follows.

1- Input representation:

$$\mathbf{e}_i = \mathbf{e}_{\text{token}}(x_i) + \mathbf{e}_{\text{segment}}(x_i) + \mathbf{e}_{\text{position}}(i) \tag{1}$$

2- Self-attention mechanism:

- Scaled Dot-Product Attention:

$$Q = XW^Q, \quad K = XW^K, \quad V = XW^V \tag{2}$$

- Attention scores:

$$\mathbf{A} = \text{softmax}\left(\frac{\mathbf{QK}^T}{\sqrt{d_k}}\right) \tag{3}$$

- Contextual embeddings:

$$\mathbf{Z} = \mathbf{A}\mathbf{V} \tag{4}$$

3- Multi-Head attention:

$$\mathrm{H}_i = \mathrm{Concat}(\mathrm{head}_1, \mathrm{head}_2, \ldots, \mathrm{head}_h)\mathbf{W}^O \tag{5}$$

4- Feed-Forward Network:

$$\mathrm{FFN}(\mathbf{h}) = \mathrm{ReLU}(\mathbf{h}\mathbf{W}_1 + \mathbf{b}_1)\mathbf{W}_2 + \mathbf{b}_2 \tag{6}$$

5- Layer normalization:

$$\mathbf{h}' = \mathrm{LayerNorm}(\mathbf{h} + \mathrm{MultiHeadAttention}(\mathbf{h}))$$

$$\mathbf{h}'' = \mathrm{LayerNorm}(\mathbf{h}' + \mathrm{FFN}(\mathbf{h}')) \tag{7}$$

In summary, first, a sentence or paragraph is tokenized and then converted to a matrix using embedding techniques. The attention mechanism is used to focus only on the related words of the given sentence for each word in the target sentence. The softmax function is used to produce similar embeddings for similar sentences, and dissimilar embeddings otherwise.

The measure of calculating the semantic similarity between two sentences is cosine similarity. This measure calculates the semantic distance between the sentences from the search file and the sentence embeddings of the target file. Distance = 1 – similarity. The similarity, as presented in equation (8) is a value between (and including) 0 and 1. The v and w vectors are sentence embeddings of a sentence or a piece of text. The higher the cosine similarity, the higher the semantic similarity of two sentences is. Each sentence or paragraph is represented as a vector in this method.

The final similarity score is adjusted; the adjustment which is related to the keyword extraction process explained above, is based on the presence of the extracted keywords from the target file in the sentences of the search file. The main idea is that if a word in the sentence exists in the target sentence, the cosine similarity of this pair of sentences gets a high score.

$$\text{cosine}(\vec{v},\vec{w}) = \frac{\vec{v}\cdot\vec{w}}{|\vec{v}||\vec{w}|} = \frac{\sum_{i=1}^{N} v_i w_i}{\sqrt{\sum_{i=1}^{N} v_i^2}\sqrt{\sum_{i=1}^{N} w_i^2}} \quad \textbf{(8)}$$

A Real-World Example

Here, we present an example from the risk management regulations to understand each step of the suggested method more easily. We'll employ the following phrase: "Banka bünyesinde gerçekleştirilen BS iç kontrol ve iç denetim çalışmalarının sonuçlarının veya tespit edilen bulguların risk envanterine girdi teşkil etmesi sağlanır." (It is ensured that the results or findings of IS internal control and internal audit studies carried out within the bank constitute input to the risk inventory).

- First, it starts with the initial text processing which includes tokenization and POS tagging: tokenization splits the sentence into tokens: ["Banka", "bünyesinde", "gerçekleştirilen", "BS", "iç", "kontrol", "ve", "iç", "denetim", "çalışmalarının", "sonuçlarının", "veya", "tespit", "edilen", "bulguların", "risk", "envanterine", "girdi", "teşkil", "etmesi", "sağlanır"]. Then, the POS tagger assigns each token to its most appropriate part-of-speech tag.
- The next step is the keyword extraction with YAKE model, it starts to extract the keywords from the target file which are ["iç kontrol", "iç denetim", "risk envanteri"].
- In the next step, after applying the embedding and semantic analysis to both files, the similar candidates to this specific sentence will be three options with their similarity scores:
    1. RİSK.03-BT Risklerinin İzlenmesi Alt Süreç Amacı: Bu alt süreç, kontrol prosedürlerinin test edilmesinin sağlanmasını, test kayıtlarının değerlendirilmesini, aksiyon planlarının güncellenmesinin sağlanmasını ve risklerin raporlanmasını amaçlar. (Similarity 0.75567). Which means "03-IT Risks Monitoring Sub-Process Purpose: This sub-process aims to ensure that control procedures are tested, test records are evaluated, action plans are updated and risks are reported."
    2. RİSK.04.02-BT denetim bulgularının BT bünyesinde raporlanması Yönetim ve/veya bulgu sahipleri/paydaşları tarafından talep edilen BT sahipliğindeki denetim bulgularının ve aksiyon planlarının güncel durumları BT bünyesinde raporlanır. (Similarity 0.72924). Which means "2- RISK.04.02-Reporting of IT audit findings within IT. The current status of IT-owned audit findings and action plans requested by management and/or finding owners/stakeholders are reported within

IT."

3. Alt Süreç Açıklaması Azaltma/Önleme yanıtı verilen riskler kapsamında belirlenen kontrol prosedürlerinin kontrol sahipleri tarafından test edilmesini, test kayıtlarının değerlendirilmesini, tespit edilmiş risk bulguları kapsamında tanımlanan aksiyon planlarının aksiyon sahipleri tarafından güncellenmesini ve risklerin raporlanmasını sağlar. (Similarity 0.6804). Which means "Sub-Process Description Ensures that the control procedures determined within the scope of the risks to which the Reduction/Prevention response is given are tested by the control owners, the test records are evaluated, the action plans defined within the scope of the identified risk findings are updated by the action owners and the risks are reported."

Although it might seem like the first one has the highest similarity but the second one is closer in meaning to the original sentence. Figure 1 and Algorithm 1 illustrate a flowchart and a pseudocode including the main steps of the proposed methodology.

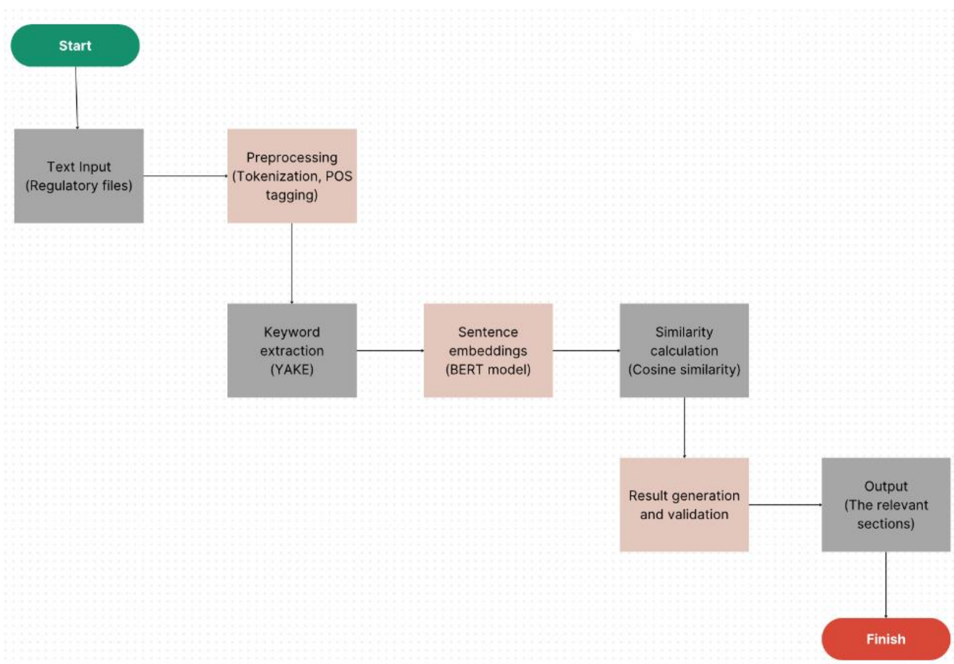

Figure 1. The flowchart of the proposed method

Algorithm 1 Regulatory compliance for the Turkish banks.
Input: A given regulation and the bank's main report
Output: The most similar regulations to the input regulation, extracted from the bank's main report

Initialize the YAKE parameters (ngram size and keyword numbers) and the number of candidate output sentences
Get input text files from the user
Extract keywords from the input regulation file (InReg)
Calculate the sentence embeddings of InReg
For sentence i in the main document:
  Calculate the sentence embeddings of sentence i
  Calculate the cosine similarity of InReg with sentence i as CoSim
  if sentence i contains one or more keywords, increase its score by 0.05
  If (CoSim > threshold(0.4))
    Append sentence i to the output candidate list
Ask an expert to approve or decline the sentences of the candidate sentence list
Return the approved sentence list as the final output

Implementation

The proposed system was developed using the Flask application, which is a micro web framework written in Python. Our project consists of an input upload form as shown in Figure 2. As soon as the user uploads the files, the system analyzes the sentences and prints the output sentences/paragraphs of the similarity-checking analysis in about 1-2 seconds on the output form as illustrated in Figure 3. Then, an expert user selects the desired ones (and deletes unrelated ones) which will be downloaded to a text file on the user's local device.

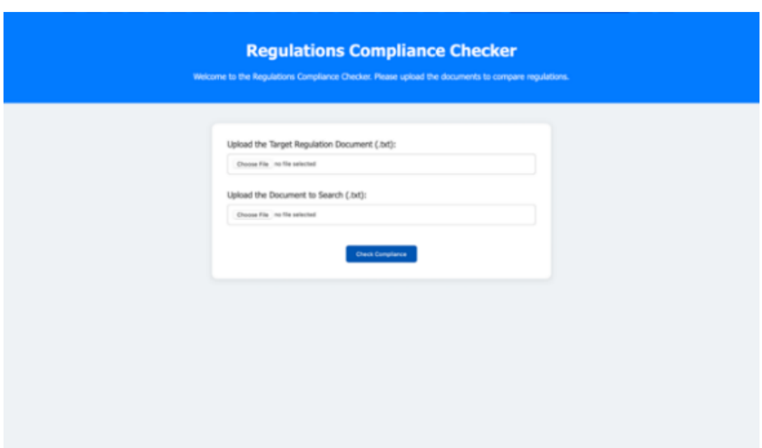

Figure 2. The first page for uploading files.

FINDINGS

This section provides the results achieved in the experimental evaluation of the proposed system sets that were tracked in the main document of a Turkish bank.

Performance metrics: The proposed system was examined by Precision and Recall metrics. These metrics are usually used to evaluate the performance of classification systems. Since the regulatory detection systems also accomplish a binary classification (a sentence in the main document is relevant or non-relevant to the given regulation), we used these metrics for evaluation. The system achieved 86.5% precision and 89.5% recall averages, showing the capability of the system to identify relevant regulations and minimize false negative cases. These metrics are calculated according to equations (9) and (10). Note that these statistics are the macro average of all regulations related to IT risk management regulations.

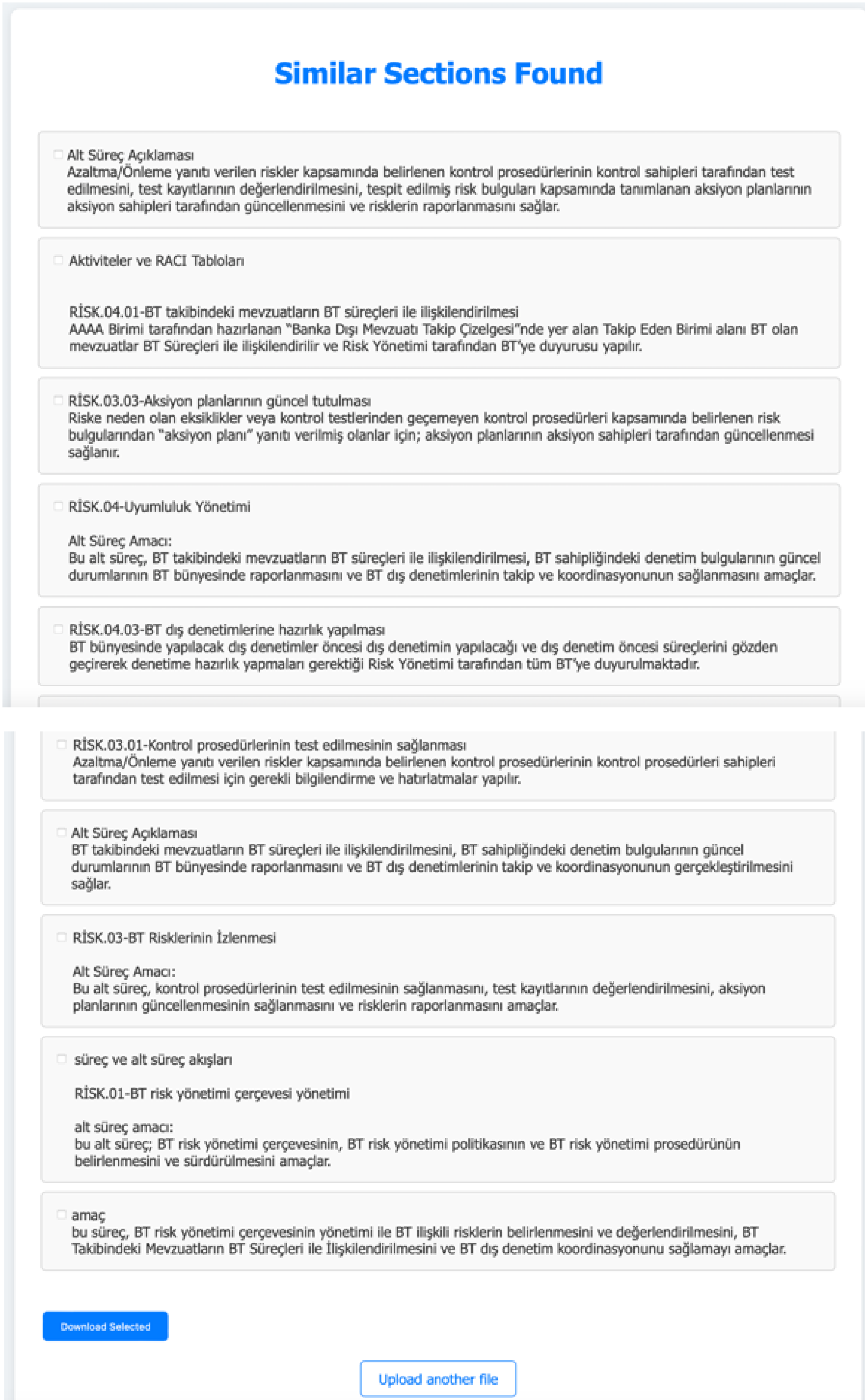

Figure 3. The results page

$$Precision = \frac{Number\ of\ sentences\ correctly\ estimated\ as\ relevant\ to\ a\ regulation}{number\ of\ all\ sentences\ estimated\ as\ relevant\ to\ a\ regulation} \tag{9}$$

$$Recall = \frac{Number\ of\ sentences\ correctly\ estimated\ as\ relevant\ to\ a\ regulation}{number\ of\ all\ sentences\ relevant\ to\ a\ regulation} \tag{10}$$

Streamlines: the initial tests and implementations of the system have proven the consuming time reduction for performing the regulation-checking process. Compared to the traditional manual approach, the system resulted in saving the time by approximately 89.47% reducing the time from 9-10 hours to 1 hour. The reason for this reduction is the process automation of extracting keyword and performing semantic analysis, which accurately and efficiently process large volumes of documents.

The compliance workflow: the system offers a feature of storing the uploaded files to be checked in a separate file, in this way the users can keep track of the files, which were checked and which were not. This feature impacts the overall compliance workflow by document management. In addition, the system can be adapted to the rapid changes in regulations by accurately detecting them without reprogramming them.

Table 1 provides the numerical results obtained as Precision and Recall, separately for the risk management and data management regulations. As can be seen, the Precision and Recall values are relatively high, and the proposed method is trustworthy especially because there will also be a manual check on the extracted sentences from the main document. This means that the mistakenly extracted sentences (false positives) can be manually deleted but sentences that were not extracted that should have been extracted (false negatives) cannot be identified by the system—about 10% of regulations. The proposed method works slightly better for the data management regulations probably because these regulations are more easily understandable for transformers compared to the risk management regulations. The time reduction as mentioned already is 87.5% compared to the manual case. In summary, we lose about 12.5% accuracy in contrast to earning 87.5% of the time spent for manual checks.

Table 1. Obtained results for risk management and data management regulations

Regulation type	Precision	Recall	F1-score	time reduction
Risk management regulations	0.86	0.89	0.87	87.5%
Data management regulations	0.87	0.90	0.88	87.5%
Average	0.865	0.895	0.875	87.5%

DISCUSSION AND CONCLUSIONS

Compared to traditional regulatory compliance procedures, which are mostly labour-intensive and prone to errors, our semantic approach provides an opportunity to understand and analyse the semantic content of regulations by utilizing Sentence Transformers and the BERT model. Traditional approaches generally comprise manual evaluations and basic keyword-based searches, which are insufficient for the comprehension of complex regulatory environments. Our system's capacity to comprehend and parse complicated regulatory language (especially for morphologically rich languages such as Turkish) lowers error margins and guarantees greater compliance. As mentioned earlier, we could not find any research work doing the same task as ours for the regulation compliance, therefore we were unable to provide a fair comparison.

The most similar projects to our system are the work done by Zhang and El-Gohary (2013) and Gupta, et al., (2024) which were explained in the Introduction section. The former employed a rule-based method to concentrate on extracting information from construction documents. The disadvantage of this work is the difficulty of maintenance for the rule-based methods since the rules have to be updated each year.

The latter is the most similar study to ours, but it focuses on sustainability reporting, while our study analyses the regulatory compliance in the financial domain, especially the Turkish banks.

The proposed system in this paper is semi-automatic; i.e., the automatically generated results as relevant sentences to a given regulation will be manually controlled by an expert in the bank to lower the mistakes and guarantee a greater level of compliance. However, this manual check for about 10-20 sentences (output of our system) will take much shorter time than the manual check over the whole document with about 20-30 pages. The average time consumed in financial institutions to perform manual regulatory compliance depends on the number of documents and the complexity of the regulations. For example, by the start of this project, the bank reported that 8-10 hours had been spent for the manual check of 10 documents with a moderate complexity; however, with our AI-driven system, this process takes a maximum of one hour to be completed, and the final manual check will take at most 10 minutes. After applying the above-mentioned techniques, we obtained 87% and 90% respectively as Precision and Recall for the automatic monitoring of risk regulations. F-score is the harmonic mean of these two metrics which is about 88% for the proposed system. The reason for using these metrics is their fairness in evaluating each class separately, which could not be done by some other metrics such as Accuracy.

One of the limitations of this system is its semi-automatic nature, as it requires manual validation for the output to ensure its accuracy. The proposed approach although reduces the time and workload, it does not totally eliminate human intervention. Another limitation is its specific domain and language, since this system is specifically developed for the Turkish language and the banking sector, it was not tested outside the banking domain and we are not sure if it will work on other regulations with a different language.

Future developments include adaptive learning, in which the system continuously modifies its model in response to new laws and anonymous user input, to increase the accuracy and responsiveness. The keyword extraction model could be replaced with another model trained particularly for extracting the bank regulations keyword (or the keywords of another domain). Furthermore, adding audio from regulatory meetings to our system to handle multimodal data could result in a more comprehensive compliance tool.

ACKNOWLEDGMENT

The authors of this paper express their gratitude to Professor Mehmet Nafiz Aydin for his guidance and co-supervision on this project.

Etik Beyanı: Yazarlar bu çalışmanın tüm hazırlanma süreçlerinde etik kurallara uyulduğunu yazar beyan eder. Aksi bir durumun tespiti halinde Kamu Yönetimi ve Teknoloji Dergisinin hiçbir sorumluluğu olmayıp, tüm sorumluluk çalışmanın yazarına aittir.

Yazar Katkıları: Rahim DEHKHARGHANI, Olla HUSSEİN, Ezgi CAN, Laura ORYNGALIYEVA, ve Hakan KANTAŞ çalışmanın tamamında birlikte katkı sunmuştur.

Çıkar Beyanı: Yazarlar ve herhangi bir kurum/ kuruluş arasında çıkar çatışması yoktur.

Teşekkür: Yayın sürecinde katkısı olan hakemlere teşekkür ederiz.

Ethics Statement: The authors declare that the ethical rules are followed in all preparation processes of thisstudy. In the event of a contrary situation, the Journal of Public Administration and Technology has noresponsibility and all responsibility belongs to the author of the study.

Author Contributions Rahim DEHKHARGHANI, Olla HUSSEİN, Ezgi CAN, Laura ORYNGALIYEVA and Hakan KANTAŞ have contributed to all parts and stages of the study.

Conflict of Interest: There is no conflict of interest among the authors and any institution.

Acknowledgement: We would like to thank the referees who contributed to the publication process.

KAYNAKLAR

Al-Shabandar, R., Liu, J., Lightbody, G., Wang, H., Browne, F., & Zheng, H. (2019). The Application of Artificial Intelligence in Financial Compliance Management. In AIAM 2019: Proceedings of the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (pp. 1-6). Association for Computing Machinery. https://doi.org/10.1145/3358331.3358339

Aziz, L. A.-R., & Andriansyah, Y. (2023). Innovative Approaches to Banking Fraud Prevention and Risk Management using AI. RCBA, 6(1), 110-130. https://doi.org/10.12345/rcba-2023-061

Bischof, J., Daske, H., Elfers, F., & Hail, L. (2021). A Tale of Two Supervisors: Compliance with Risk Disclosure Regulation in the Banking Sector. Journal of Accounting Research. https://doi.org/10.1111/1911-3846.12715

Benkassioui, B., Kharmoum, N., Hadi, M. Y., & Ezziyyani, M. (2023). NLP methods' information extraction for textual data: An analytical study. In International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD 2022) (pp. 515-527). Springer. https://doi.org/10.1007/978-3-031-15711-7_42

Calderón, A. (2020). Regulatory Compliance & Supervision in AI Regime: Banks and FinTech (Master's thesis, University of Helsinki, Faculty of Law).

Campos, R., Mangaravite, V., Pasquali, A., Jorge, A., Nunes, C., & Jatowt, A. (2020). YAKE! Keyword extraction from single documents using multiple local features. Information Sciences, 509, 257-289

Campos, R., Mangaravite, V., Pasquali, A., Jorge, A. M., Nunes, C., & Jatowt, A. (2018). A Text Feature Based Automatic Keyword Extraction Method for Single Documents. In G. Pasi, B. Piwowarski, L. Azzopardi, & A. Hanbury (Eds.), Advances in Information Retrieval: ECIR 2018 (Lecture Notes in Computer Science, Vol. 10772, pp. 684-691). Springer, Cham. https://doi.org/10.1007/978-3-319-76941-7_63

Campos, R., Mangaravite, V., Pasquali, A., Jorge, A. M., Nunes, C., & Jatowt, A. (2018). YAKE! Collection-independent automatic keyword extractor. In Advances in Information Retrieval. https://doi.org/10.1007/978-3-319-76941-7_63

Cyree, K. B. (2016). The Effects of Regulatory Compliance for Small Banks around Crisis-Based Regulation. Journal of Financial Research, 39(3), 289-309. https://doi.org/10.1111/jfir.12096

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 4171-4186. https://doi.org/10.18653/v1/N19-1423

Gupta, A., Chadha, A., & Tewari, V. (2024). A Natural Language Processing Model on BERT and YAKE Technique for Keyword Extraction on Sustainability Reports. IEEE Access, 12, 7942-7951. https://doi.org/10.1109/ACCESS.2024.3352742

Jain, V., Balakrishnan, A., Beeram, D., Najana, M., & Chintale, P. (2024). Leveraging Artificial Intelligence for Enhancing Regulatory Compliance in the Financial Sector. International Journal of Computer Trends and Technology (IJCTT), 72(5), 56-65. https://doi.org/10.14445/22312803/IJCTT-V72I5P116

Khurana, D., Koli, A., Khatter, K., et al. (2023). Natural language processing: State of the art, current trends and challenges. Multimedia Tools and Applications, 82(5), 3713–3744. https://doi.org/10.1007/s11042-022-13428-4

Zhang, J., & El-Gohary, N. M. (2013). Semantic NLP-Based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking. Journal of Computing in Civil Engineering, 27(6), 668-676. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000315