Research Article

Customer Credit Risk Scoring Using Natural Language Processing (NLP): A News Analysis Approach

Authors: Hayri Sarı, Metin Turan

To cite to this article: Sarı, H., Turan, M., (2025). Customer Credit Risk Scoring Using Natural Language Processing (NLP): A News Analysis Approach, International Journal of Engineering and Innovative Research 7(2), p 138-152.

DOI: 10.47933/ijeir.1814441

To link to this artice: https://dergipark.org.tr/tr/pub/ijeir/archive

International Journal of Engineering and Innovative Research

http://dergipark.gov.tr/ijeir

Customer Credit Risk Scoring Using Natural Language Processing (NLP): A News Analysis Approach

Hayri SARI[1*], Metin TURAN[2]

[1] Istanbul Ticaret University, Department of Computer Engineering, Maltepe,İSTANBUL.
[2] Istanbul Ticaret University, Department of Software Engineering, Maltepe,İSTANBUL.

*Corresponding Author: hayri.sari@istanbulticaret.edu.tr
(**Received:** 31.10.2025; **Accepted:** 25.11.2025)

https://doi.org/10.47933/ijeir.1814441

ABSTRACT: The banking industry must make swift and reliable assessments in credit-granting processes that inherently involve high risk and must be grounded in information about the customer. In credit risk analysis, it is important to consider not only the customer's past performance and financial position but also the news circulating in the media environment. However, it is not feasible to manually review the large volume of news produced on a daily basis.
This study introduces a novel, automated framework that utilizes a two-stage approach—Named Entity Recognition (NER) followed by sentiment analysis using Turkish BERT-based models—integrated with a scoring algorithm to evaluate customer-related news. The research was conducted using a dataset comprising over 300 RSS news items related to 18 distinct customers. By collecting news from RSS feeds and sequentially processing them, the developed framework enables the classification of news items as positive, negative, or neutral, allowing these scores to be integrated directly into credit evaluation processes. Experimental results indicated that the proposed model reliably identifies and classifies the sentiment of customer-related news items in a manner consistent with business experts' assessments. Specifically, the model outputs serve as actionable insights; a high frequency of negative news acts as an early warning signal to trigger deeper risk investigation, whereas consistent positive sentiment can expedite the approval process for low-risk applications. The evaluation confirmed that the system can automatically collect and analyze news content, supporting faster and more objective decision-making compared to manual review processes.

Keywords: Credit risk analysis, sentiment analysis, named entity recognition, BERT, MongoDB.

1. INTRODUCTION

Today, the banking sector conducts a multidimensional credit evaluation process that relies not only on financial indicators but also on unstructured data sources related to large-scale customers. In this process, external textual data such as media content is becoming increasingly critical for gauging public perception about the customer. Yet it is not possible to manually examine and make sense of the hundreds of news items produced on a daily basis. In this context, NLP- and machine-learning–based methods play a strategic role in achieving this goal [1].

This study covers the design and implementation of an automated system that performs customer-level sentiment analysis on news content and contributes to credit risk evaluation processes. The system is structured by collecting and analyzing more than 300 news items relating to 18 different customers from RSS sources. Through Turkish BERT-based models [2] specifically selected for this task, the news texts undergo both named entity recognition (NER) and sentiment analysis. While the NER stage identifies which customer a news item refers to, the sentiment analysis module determines each news item's label (positive, negative, neutral), and these signals are integrated into the scoring model.

Within the data-processing pipeline, a .NET Core infrastructure and a MongoDB-based data store are used; news texts are collected and analyzed automatically at regular intervals. Splitting long texts into segments so as not to exceed the model's 512-token input limit has been one of the key preprocessing steps that improves analytical accuracy. Model outputs are evaluated together with confidence scores; the emotional impact of each news item is transformed into a numeric value, and customer-level average scores are computed.

In the scoring model developed in this study, the sentiment labels associated with the text segments of each news item are evaluated by weighting them with the model's confidence score. Confidence scores for positive and negative labels are accumulated, and the label with the highest total confidence determines the overall sentiment of the news item. The resulting values are expressed as scores close to +1.00 for positive sentiments and to –1.00 for negative sentiments; the average of these scores at the customer level is used as the overall media sentiment score.

The contribution of this study to the literature can be considered under two main headings. First, there is a limited number of multi-news analysis studies conducted in Turkish. In this respect, it constitutes an important example for NLP applications in low-resource languages. Second, integrating two distinct NLP components—sentiment analysis and named entity recognition—and making them usable in the context of credit risk represents a hybrid approach. In addition, making the developed scoring model usable as a numerical decision-support tool enables news analysis to move beyond purely qualitative assessments and produce measurable outputs. In these respects, the study provides a concrete contribution to the development of AI-based tools for financial analysis.

The integration of NLP methods into financial decision-making processes has attracted growing interest in the literature in recent years [3]. In particular, the analysis of unstructured text sources such as news content, social media posts, and press releases has made it possible to discover new risk indicators beyond traditional financial metrics. Accordingly, methods such as sentiment analysis and named entity recognition have begun to play important roles in client-focused media tracking and credit risk evaluation processes [4]. A large share of studies on credit risk prediction have remained limited to historical financial data; however, recent research has shown that such systems can be enriched by the analysis of text-based data [5]. By quantifying positive or negative judgments derived especially from news texts, sentiment analysis adds a new dimension to decision-support systems. That said, many studies in the literature focus only on English-language datasets, while the number of applications for morphologically rich languages such as Turkish remains quite limited [6]. For example, it is possible to evaluate credit applications of prospective customers using NLP methods in credit risk assessment. Combined with information obtained from social media and other text-based data, sentiment analysis provides a more comprehensive picture of creditworthiness. Analyses employing NLP can surface emotional and social signals about the financial status of corporate clients [7]. Emotional factors are known to have significant effects especially in financial markets. In the research by Çetin and Eryiğit, how consumer perception is shaped through analysis of the sentiment of social media comments is examined. Such studies provide critical information for banks and financial institutions seeking to understand customer behavior.

Analyzing customer sentiment supports banks in making more effective and informed choices in lending decisions [8].

Sentiment analysis stands out as a method that can be used in banking customer-experience research. By systematically examining customer feedback in big data, sentiment analysis allows us to understand how customers feel about banking services. Studies on sentiment analysis for Turkish texts have contributed to deepening this field. In this context, there appears to be a strong link between sentiment analysis and customer satisfaction [9]. On digital platforms, sentiment analysis helps evaluate the performance of products and services. For example, comments collected via social media can provide an overall sentiment view related to banking services. Tokcaer emphasizes the importance of examining the usability of digital services based on sentiment analysis. Thus, banks can better understand which areas receive negative feedback as they improve their services [10]. The impact of social media becomes more pronounced especially with big data and data-mining applications. User comments and posts can be processed through sentiment analysis to capture the current situation. Sentiment analysis allows content shared on social media to be classified as positive, negative, or neutral. For instance, social media analyses can be used to assess market perceptions that influence a bank's lending decisions [11]. İlhan and Sağaltıcı performed sentiment classification of user posts on Twitter using machine-learning methods such as Naive Bayes and Support Vector Machines (SVM). Such studies present applicable models for interpreting social media data and help evaluate a company's image on social media by analyzing user sentiment [12].

When credit risks faced by banks are measured effectively, this helps maintain financial stability. İldaş's work emphasizes the importance of accurate measurement methods as the first step in risk management. In this context, combining the data obtained from sentiment analysis with different factors that determine creditworthiness lays the groundwork for banks to make healthier and more efficient credit decisions [13]. As banks develop sentiment analysis methods by leveraging social media data, they particularly benefit from machine-learning techniques. For example, in the study conducted by Demirbilek and Demirbilek, machine-learning methods that analyze user sentiments using data obtained from user comments are discussed. Such techniques help banks evaluate the financial behavior and credit-usage potential of individual customers [14]. The evaluation of social media data in lending processes provides important information not only for retail loans but also for corporate loans. Kabaklarlı's work examines the impact of credit card usage on monetary policy, revealing the relationship between individuals' borrowing behavior and markets. Accordingly, sentiment analysis of social media can be used to better understand individuals' credit card usage and borrowing behavior [15]. Similarly, by analyzing users' posts on social media platforms, banks can observe levels of customer satisfaction and potential risk factors. Çelik and Akbulut's study analyzed user comments using sentiment analysis techniques on Twitter and emphasized the importance of this information in banks' strategic decision-making processes [16]. Sentiment data obtained from social media provides indirect information about users' financial situations. For example, Sel's work notes that individual and corporate perceptions on social media can be made measurable through sentiment analysis. Such analyses provide valuable information that banks can use in lending processes, laying the groundwork for better evaluations leading to credit approval or rejection [17]. Sentiment analysis techniques are applied using different machine-learning methods, thereby enabling the derivation of meaningful insights from large datasets. Yıldırım's research, for instance, addresses the development of AI-based methods proposed for sentiment analysis. Such automated systems allow banks to rapidly analyze complex and large datasets to better understand important factors affecting lending [18]. Furthermore, research on sentiment analysis on social media has increased in recent years. Demirel and Tapan's study examined the relationship between social media addiction and purchasing behavior and emphasized how such analyses can be directive in financial decisions. Considering social media trends and user sentiment when evaluating loan applications offers banks a broader perspective [19].

2. MATERIAL

The dataset used in this study consists of more than 300 RSS news items relating to 18 different customers. The objective is to systematically collect the content published about customers in the online media environment and store it in a form suitable for analysis using NLP methods in subsequent stages.

2.1. Data Sources and Coverage

Data are obtained from RSS (Really Simple Syndication) [20] feed endpoints provided by national and local news sites. RSS is preferred because (i) it offers a standardized metadata schema (title, link, description, date), (ii) it allows new content to be captured without delay, and (iii) it enables the establishment of a scalable collection mechanism for high-volume flows. In this study, topic/page feeds deemed relevant to the customers were monitored, and content published over the period was pulled at regular intervals.

2.2. Collection Frequency and Scheduling

Feeds were queried periodically via a scheduled-job mechanism. At the application level, this job was activated with short-interval triggers (on the order of minutes; e.g., 180 seconds), and new RSS items were checked at each trigger. Periodic execution ensures that hourly/daily fluctuations in news production are not missed and that the dataset remains up to date.

2.3. RSS Parsing Process

The collected news data were parsed in XML format and converted into a structured dataset. From each news record, the following fields were obtained:

- title: news headline
- link: URL that points to the original content
- description: summary/body text of the news
- pubDate: publication date-time

An example JSON record of a news item stored in the database is shown below:


```
{
  "_id": "ObjectId('66f1d4ce4e5be6568e2c6ac7')",
  "title": "THY’s credit rating upgraded",
  "link": "https://www.ahaber.com.tr/ekonomi/2024/09/19/thynin-kredi-notu-yukseltildi",
  "description": "<img src='https://iaahbr.tmgrup.com.tr/77319e/1200/675/0/20/850/498?...'>",
  "pubDate": "Thu, 19 Sep 2024 12:33:10 +0300"
```

}

During parsing, date fields were converted into a standard format; text fields (title and description) were stored as UTF-8. These fields constitute the primary input of the subsequent text-processing stages.

2.4. Storage Structure

The data obtained were stored in a NoSQL database (MongoDB). Source information and news content were kept in separate collections. In this way, source management (add/remove/update) and content management (search, deduplication, reprocessing) were made independent and scalable. Each news document contains the following information: source ID, title, link, description, and publication date. This design is conducive to associating text preprocessing and model outputs (e.g., named entity recognition results) with separate collections in later stages.

2.5. Language and Content Characteristics

The dominant language of the content in the dataset is Turkish. The news covers customer-oriented themes such as the economy, corporate announcements, investment, and sectoral developments.

2.6. Integrity and basic quality checks

During collection, link-level duplication risk was monitored; inserting the same link repeatedly was prevented. Care was taken to store timestamps consistently, thereby improving the reliability of time-based queries in period analyses.

2.7. Ethics and access

All data were obtained automatically from publicly available RSS feeds. The content was used solely for text mining and statistical analysis within the scope of the research; copyright and access conditions were observed.

3. METHODS

This section describes the method by which data obtained from news feeds are processed end to end and transformed into a numeric sentiment indicator at the customer level. The diagram (Figure 1) summarizes the data pipeline within the scope of the study—from collecting data from RSS sources to storing them in MongoDB, applying preprocessing steps, running Turkish BERT-based NER and sentiment analysis, then computing the average sentiment at the customer level via the Scoring Model, and finally integrating the results into credit-risk decision support.

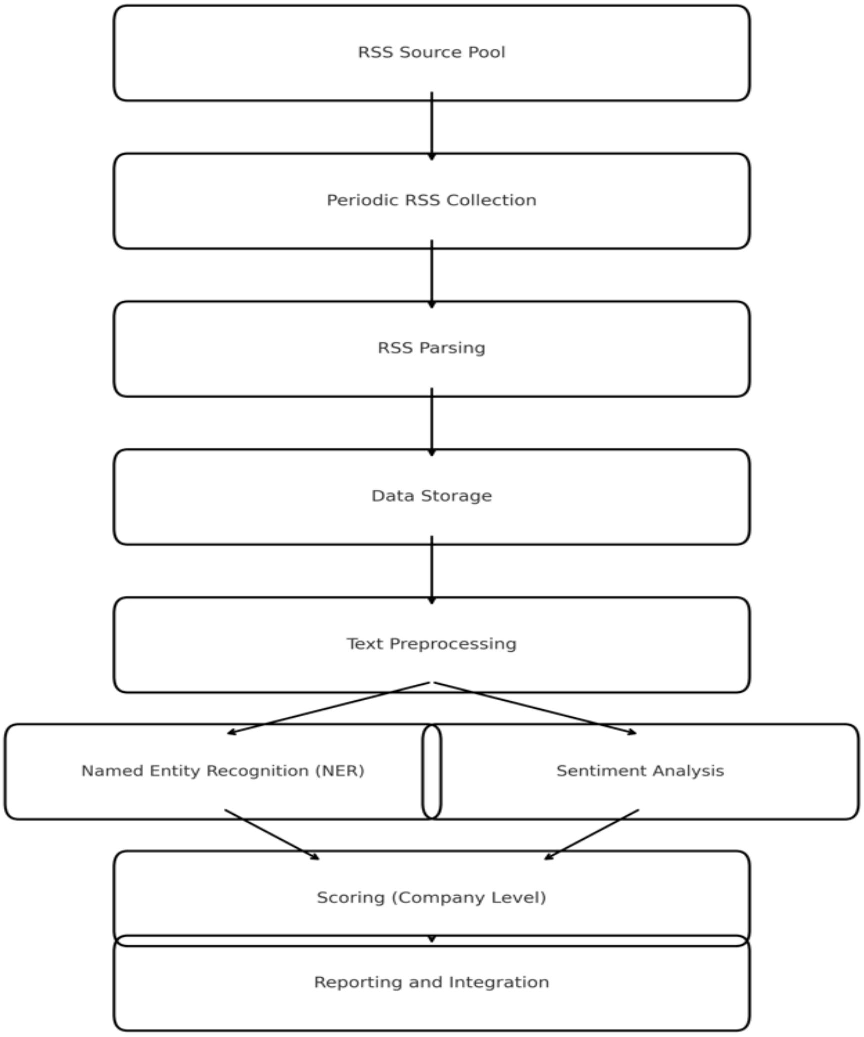

Figure 1. Data set and methodology flowchart

3.1. Scheduled Data Acquisition and RSS Parsing

News sources are triggered at specified intervals via a job scheduler. In the application, a Quartz-based scheduler is started, a job and a trigger are created, and the workflow is configured to run at short periods (e.g., 180 seconds). Each trigger initiates content retrieval from RSS endpoints and the processing of new items. RSS responses are parsed as XML; from each item node, the fields title, link, description, and pubDate are extracted.

3.2. Persistent Storage and Data Management

Parsed items are persisted on MongoDB. Source endpoints and content items are tracked in separate collections to support flexible querying, deduplication, and reprocessing scenarios.

3.3. Text Preprocessing

Before the sentiment analysis and NER stages, HTML markup in RSS descriptions is cleaned to obtain plain text. This step aims to make model inputs consistent and to reduce noise. Before the subsequent segment-based processing to be applied to long texts, content normalization is completed.

3.4. Named entity recognition (NER)

A BERT [21]-based NER model was applied to Turkish texts to extract named entities such as persons, organizations, and locations. The implementation utilized the Hugging Face Transformers library [22] and a Turkish BERT model fine-tuned for NER [23] via the pipeline("ner", grouped_entities=True) function. This configuration allows adjoining entities (e.g., “Ahmet Yılmaz”) to be merged as a single item while producing label, confidence score, and positional information for each entity. The NER outputs were stored in a separate collection linked to the corresponding news ID, thereby improving the accuracy of customer mapping and ensuring reliable input for subsequent scoring processes.

3.5. Sentiment Analysis

At the implementation stage, the Turkish BERT model fine-tuned for sentiment analysis [24] was run via the Hugging Face Transformers framework [22]. Unlike standard rule-based approaches, this model is built upon the BERT-Base architecture (12 layers, 768 hidden units, 12 attention heads, and 110M parameters). The model was fine-tuned on a diverse dataset comprising approximately 49,000 labeled samples, including movie reviews, product comments, and tweets. During the fine-tuning phase, the AdamW optimizer was employed with a learning rate of 2e-5 and a batch size of 32 over 3 epochs to minimize the cross-entropy loss. Regarding performance metrics, the model demonstrates high reliability with an accuracy of approximately 95% and an F1-score of 94% on the test set [24]. The model predicts one of three possible classes (positive, negative, or neutral) for each news text and produces a confidence score for that class.

To accommodate the model’s input limitation of 512 tokens, long news texts were divided into overlapping segments. Although the BERT architecture supports sequences up to 512 tokens, processing inputs at maximum length substantially increases computational complexity and inference latency due to the quadratic nature of the self-attention mechanism [21]. Conversely, using overly short segments (e.g., 128 tokens) risks weakening long-range semantic dependencies, which play a critical role in interpreting multi-sentence news narratives [25]. Therefore, a window size of 256 tokens was selected in this study to achieve a practical balance between computational efficiency and the preservation of semantic coherence. This configuration ensures that the emotional tone of the text is captured consistently without incurring unnecessary processing overhead.

The sentiment predictions for each segment were consolidated using a cumulative-confidence approach. In this method, confidence values for each class were summed, and the class with the highest total confidence was assigned as the overall sentiment label of the news item. In this way, sentiment variations observed across different parts of the text were unified under a single overall assessment.

The resulting ⟨label, confidence⟩ pair was matched with the corresponding news record in the database and passed directly as input to the module where the customer-level media sentiment score is calculated. This structure enables the sentiment information extracted from text to be integrated into analytical processes in a traceable and measurable manner.

Example:

URL: https://www.haberturk.com/bim-ceza-hukuksuz-ve-hakkaniyetsiz-3767136-ekonomi

Model outputs:

- Window 1 → positive, score: 0.70
- Window 2 → negative, score: 0.98
- Window 3 → negative, score: 1.00

Aggregated result (cumulative confidence):

- Class totals — positive: 0.70; negative: 1.98; neutral: 0.00
- Overall sentiment label: Negative
- Total confidence (negative): 1.98
- Number of segments (negative): 2
- Confidence score (mean of negative windows): 0.99 (= 1.98 / 2)

In this example, segment-level predictions produced by the model are combined based on the cumulative confidence value, and the result is determined with a high level of confidence in the negative direction.

3.6. Scoring Model

To compute a general sentiment score, for each news item we have:

- Sentiment label (positive or negative)
- Confidence score (between 0.51 and 0.99)

Weighted Overall Score

We score each news item as positive or negative:

- Add confidence score to sum for a positive emotion
- Subtract confidence score from sum for a negative emotion

Then we divide sum by the number of news items (Equation 1):

$$Overal\ Score = \frac{\sum (\text{semantik\_score})}{total\_news} \tag{1}$$

This score gives the following comments:

- close to +1.00 : Highly positive news
- around 0.00 : Neutral or mixed
- close to –1.00 : Highly negative news

3.7. Reporting and Integration

Final customer scores are reported with table and chart outputs; when needed, they can be exported in formats such as CSV. The results are integrated as a complementary indicator into credit risk evaluation processes. The continuously running scheduler and parsing components support the continuity of data flow in production environments.

4. EXPERIMENTS

In this section, the overall view of media-based sentiment scores for 18 anonymized customers (Figure 2) and evaluations for three specific customers are presented. News texts were processed with a Turkish BERT-based classifier; segment results were reduced to the news level, recorded in the form ⟨label, confidence⟩, and converted into a sentiment score at the customer level via a weighted average. The score range is [–1, +1]: positive values indicate positive media perception, negative values indicate negative media perception. The dataset used in this analysis is openly available on Zenodo [26].

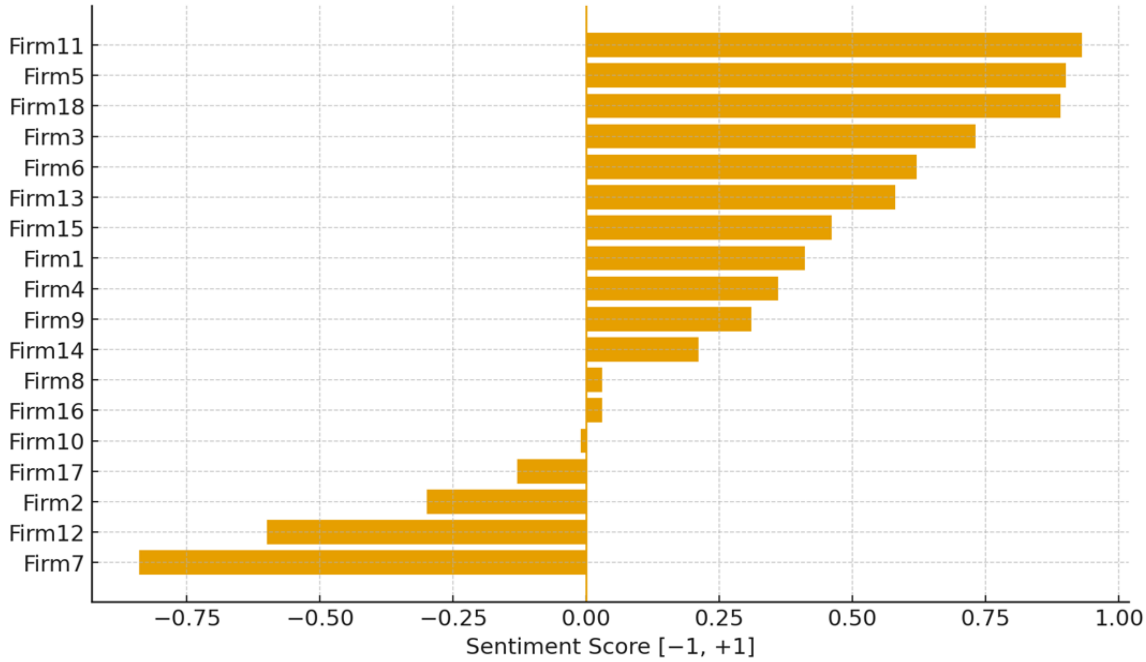

Figure 2. Overall view of media sentiment scores for 18 anonymized customers

4.1. Firm1

The overall view (Figure 3) indicates a composition in which positive-toned content is dominant in number and magnitude, while a small number of negative items pull the average down only to a limited extent.

- Number of news items: 11
- Class distribution: 8 positive, 3 negative
- Sectors: Construction, Energy, Tourism
- Average directional score: +0.412

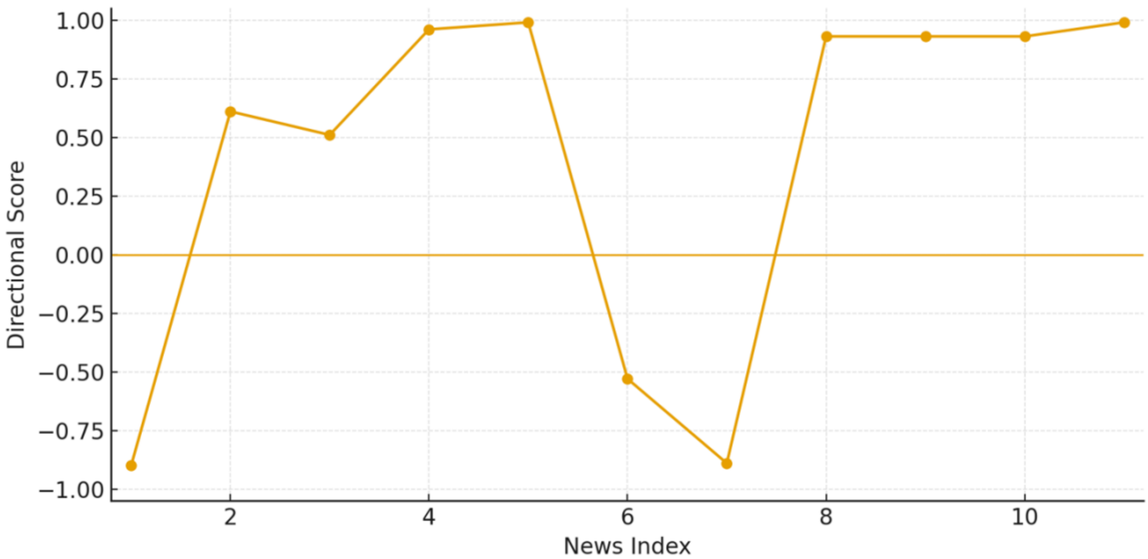

Figure 3. Directional score series by news item for Firm1

Credit Risk Commentary: Perception is moderately positive. Standard monitoring should be maintained; a cautious approach is advised against sector-driven volatility. Validation with operational performance and cash-flow indicators is recommended. Sector: Construction, Energy, Tourism. Short-term fluctuations due to sectoral external shocks and news intensity should also be taken into account in the decision process.

4.2. Firm17

The view (Figure 4) indicates a mildly negative tendency in which a small number of negative items with high absolute magnitude exert downward pressure on the average.

- Number of news items: 7
- Class distribution: 3 positive, 4 negative
- Sectors: Energy, Textile, Technology
- Average directional score: –0.139

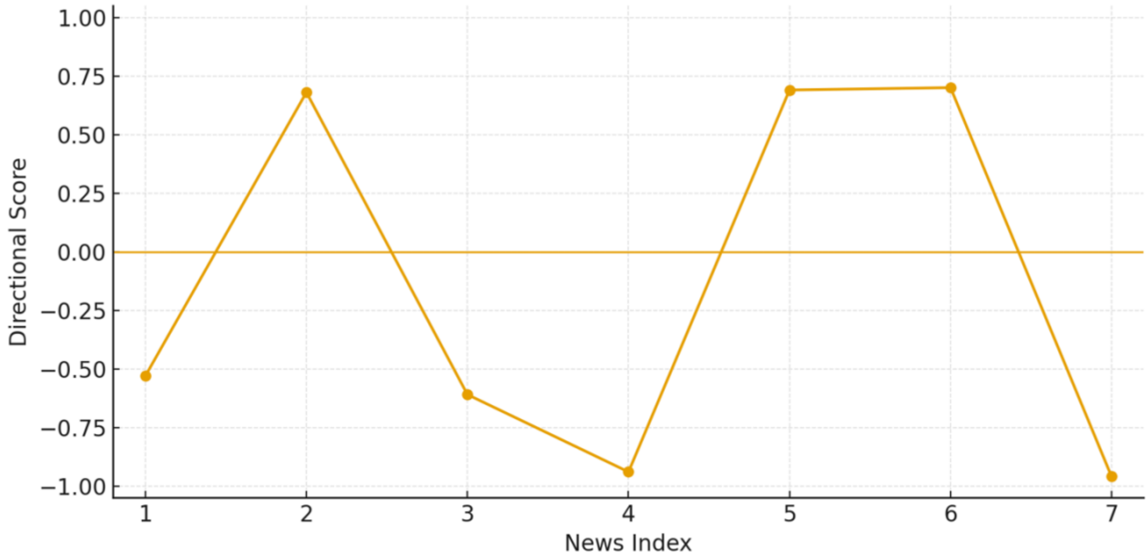

Figure 4. Directional score series by news item for Firm17

Credit Risk Commentary: Perception appears neutral/mixed. Because the news composition is mixed, the standard monitoring framework should be maintained; an approach sensitive to subsequent news flow should be adopted. Sectors: Energy, Textile, Technology. Short-term fluctuations due to sectoral external shocks and news intensity should also be taken into account in the decision process.

4.3. Firm18

The fact that all news items fall in the positive class and that the lower bound remains high indicates a strong and one-directional positive perception over the period (Figure 5).

- Number of news items: 8
- Class distribution: 8 positive, 0 negative
- Sector: Mining
- Average directional score: +0.891

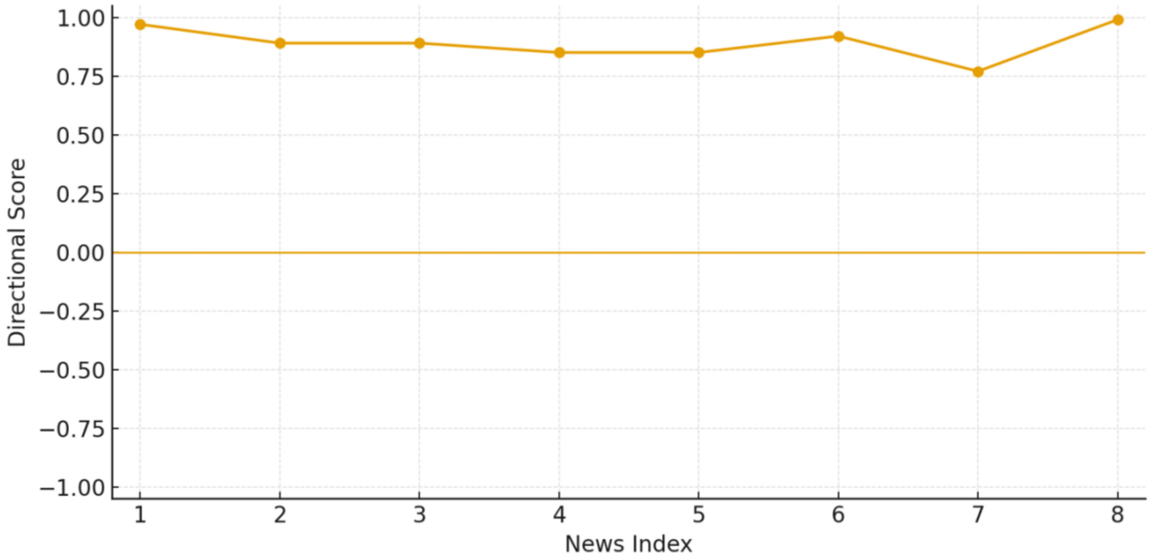

Figure 5. Directional score series by news item for Firm18

Credit Risk Commentary: Media sentiment is strongly positive. In the short term, it provides a supportive signal on reputation and demand; the standard monitoring period appears sufficient. A window of opportunity may emerge for new products/limits.

5. CONCLUSIONS

This study shows that texts obtained from news can be processed within an end-to-end pipeline and transformed into a unique and interpretable sentiment indicator at the customer level. Thanks to the combined operation of RSS-based data acquisition, basic text cleaning, named entity recognition, and Turkish BERT–based sentiment classification, it has been possible to produce decision-ready outputs in the form of label + confidence from high-volume and irregularly flowing content. In an application setting consisting of 18 customers and more than 300 news items, the method operated with low latency and in a traceable manner; carrying decisions at the news level into a customer-level weighted score simplified reporting and integration processes.

The findings stand out along three axes. First, the segmentation and aggregation steps limited context loss in long texts and increased label consistency, while mapping news items to the correct customers worked reliably. Second, customer-level sentiment scores diverged markedly by period and event type; this suggests that media perception carries a complementary signal independent of classical financial magnitudes. Third, the asymmetric effects of positive/negative event clusters on the score offer a practical early-warning function in decision areas such as adjusting monitoring periods and revising limit and collateral policies.

Nevertheless, the method has natural limitations. The fact that RSS feeds sometimes provide only summary content can weaken the sentiment signal and cause some content to fall outside the scope. The model’s sensitivity to thresholds and class balance and linguistic phenomena such as irony/sarcasm remain points of vulnerability. In multi-entity news items containing organizations with the same or similar names, customer disambiguation may require additional rules. In this context, periodic threshold calibration, expansion of coverage, and enrichment of disambiguation logic are important.

Future work aims to advance along the dimensions of coverage, models, and institutional integration. On coverage, the controlled addition of channels beyond RSS (full-text extraction, regulatory announcements, corporate filings, and appropriate social sources) will increase signal diversity and coverage ratio. On models, testing up-to-date architectures for Turkish (e.g., RoBERTa/DeBERTa variants and multilingual approaches) and adding aspect-based sentiment together with event type and severity extraction will strengthen the interpretability of the method's outputs. In terms of institutional integration, it is planned to link the sentiment score with quantitative risk components such as PD/LGD/EAD, to support it with uncertainty calibration and explainability techniques, and to mature MLOps-oriented monitoring and versioning processes.

In conclusion, news-based sentiment analysis and customer-level scoring appear to be addable to credit risk monitoring and reputation management frameworks as a scalable, traceable, and actionable layer. Enriching the method with different data sources, updating the model family, and deep integration with risk models will increase both predictive power and business-value creation. Thus, it becomes possible to transfer media-derived signals into institutional decision-making processes in a systematic and reproducible manner; with future enhancements, both the coverage and the decision-support capacity of the method are expected to grow.

REFERENCES

[1] SU, J., LAU, R. Y. K., YU, J., NG, D. C. T. ve JIANG, W. (2025). A multi-modal data fusion approach for evaluating the impact of extreme public sentiments on corporate credit ratings. Complex & Intelligent Systems, 11, 436. https://doi.org/10.1007/s40747-025-02067-5

[2] Savasy. (2020). *Turkish BERT models for NLP tasks* [Computer software]. Hugging Face. https://huggingface.co/savasy

[3] DU, K., ZHAO, Y., MAO, R., XING, F. ve CAMBRIA, E. (2024). *Natural Language Processing in Finance: A Survey*. Information Fusion, 115, 102755. https://doi.org/10.1016/j.inffus.2024.102755

[4] DEVEIKYTE, J., GEMAN, H., PICCARI, C. ve PROVETTI, A. (2020). *A Sentiment Analysis Approach to the Prediction of Market Volatility*. arXiv preprint arXiv:2012.05906. https://arxiv.org/abs/2012.05906

[5] *How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm* (2024). *Journal of Financial Stability*, 73, 101284. https://doi.org/10.1016/j.jfs.2024.101284

[6] Aydin, C. R. and Gungor, T. (2021). Sentiment analysis in Turkish: Supervised, semi-supervised, and unsupervised techniques. NATURAL LANGUAGE ENGINEERING, 27(4), 455–483. https://doi.org/10.1017/S1351324920000200

[7] GÖKTÜRK, İ. E. and Soydan, E. (2021). Zihinsel muhasebe eğilimi çerçevesinde bireylerde finansal karar alma süreci: finansal sosyal hizmet üzerine bir değerlendirme. Toplum Ve Sosyal Hizmet, 32(4), 1543-1564. https://doi.org/10.33417/tsh.903821

[8] Çetin, F. S. and Eryiğit, G. (2018). Türkçe hedef tabanlı duygu analizi için alt görevlerin incelenmesi – hedef terim, hedef kategori ve duygu sınıfı belirleme. Bilişim Teknolojileri Dergisi, 11(1), 43-56. https://doi.org/10.17671/gazibtd.325865

[9] Tokcaer, S. (2021). Türkçe metinlerde duygu analizi. Journal of Yaşar University, 16(63), 1516-1536. https://doi.org/10.19168/jyasar.928843

[10] TAŞKIN, K. (2022). Sanal pazar uygulamalarinin kullanilabilirliğinin duygu analizi temelli değerlendirilmesi. Premium E-Journal of Social Science. https://doi.org/10.37242/pejoss.4239

[11] Sarıman, G. (2020). Covid-19 sürecinde twitter mesajlarinin duygu analizi. Euroasia Journal of Mathematics Engineering Natural and Medical Sciences, 7(10), 137-148. https://doi.org/10.38065/euroasiaorg.149

[12] İlhan, N. and SAĞALTICI, D. (2020). Twitter'da duygu analizi. Harran Üniversitesi Mühendislik Dergisi, 5(2), 146-156. https://doi.org/10.46578/humder.772929

[13] İLDAŞ, T. (2021). Kredi riski ölçüm modellerinin değerlendirilmesi. Finansal Araştırmalar Ve Çalışmalar Dergisi, 13(25), 516-547. https://doi.org/10.14784/marufacd.976580

[14] DEMİRBİLEK, M. and DEMİRBİLEK, S. Ö. (2023). Google yorumları üzerinden makine öğrenme yöntemleri ve amazon comprehend ile duygu analizi: iç anadoluda bir üniversite örneği. Üniversite Araştırmaları Dergisi, 6(4), 452-461. https://doi.org/10.32329/uad.1383794

[15] Kabaklarlı, E. (2015). Türkiye'de kredi kartı kullanımının, para politikasındaki rolü ve etkileri. Sosyoekonomi, 23(26). https://doi.org/10.17233/se.90476

[16] ÇELİK, C. and Akbulut, M. (2023). Online pazaryerlerinin sosyal medya kanalları: twitter duygu analizi örneği. ODÜ Sosyal Bilimler Araştırmaları Dergisi (ODÜSOBİAD). https://doi.org/10.48146/odusobiad.1309113

[17] SEL, A. (2022). Pandemi sürecinde toplum görüşünün duygu analizi yöntemiyle incelenmesi: türkiye örneği. Beykoz Akademi Dergisi, 10(2), 134-154. https://doi.org/10.14514/beykozad.882353

[18] Yıldırım, M. (2022). Film yorumları kullanılarak önerilen yapay zekâ tabanlı yöntemle duygu analizinin gerçekleştirilmesi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 34(2), 751-760. https://doi.org/10.35234/fumbd.1138128

[19] DEMİREL, A. C. and TAPAN, M. G. (2023). Üniversite öğrencilerinin sosyal medya ve kompulsif çevrimiçi alışveriş bağımlılığı arasındaki ilişkinin incelenmesi. İnsan Ve Toplum Bilimleri Araştırmaları Dergisi, 12(1), 60-78. https://doi.org/10.15869/itobiad.1147669

[20] Winer, D. (2003). *RSS 2.0 specification*. http://www.rssboard.org/rss-specification

[21] 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.* https://arxiv.org/abs/1810.04805

[22] Hugging Face. (n.d.). Transformers [Computer software]. https://huggingface.co/transformers

[23] Savasy. (2020a). *Turkish BERT model fine-tuned for Named Entity Recognition (NER)* [Computer software]. Hugging Face. https://huggingface.co/savasy/bert-base-turkish-ner-cased

[24] Savasy. (2020b). *Turkish BERT model fine-tuned for sentiment analysis* [Computer software]. Hugging Face. https://huggingface.co/savasy/bert-base-turkish-sentiment-cased

[25] D C. Sun, X. Qiu, Y. Xu, and X. Huang, "How to Fine-Tune BERT for Text Classification?" in Chinese Computational Linguistics, Cham: Springer International Publishing, 2019, pp. 194–206. https://doi: 10.1007/978-3-030-32381-3_16

[26] Sarı. (2025). Dataset - Customer-based news [Data set]. Zenodo. https://doi.org/10.5281/zenodo.17429383