Gümüşhane University Journal of Science

GUFBD / *GUJS* (2024) 14(1): 45-60 doi: 10.17714/gumusfenbil.1361431 Research Article

# Analysis of customer reviews for digital banking applications with text mining methods

*Metin madenciliği yöntemleri ile dijital bankacılık uygulamalarına yönelik müşteri yorumlarının analizi*

**Burcu SAYIN OKATAN*[1], Handan ÇAM[1]**

[1]*Gümüşhane Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, 29100, Gümüşhane*

• Received: 16.09.2023 • Accepted: 23.10.2023

**Abstract**

Virtual services, which provide an important comfort area in today's digital world, are used by the majority of people. Accordingly, digital banking is one of the most used online financial services. In this research, it was aimed to analyze the digital banking services used by bank customers at a high rate and by using text mining methods using a data pool consisting of their comments. In the study, in the light of the data of the Banks Association of Turkey, the digital banking data of the 10 most used private banks and 3 state banks and a total of 13 banks constitute the population. The data covers the period from January 2020 to August 2022.In total, between 1,200,000-1,250,000 raw data were obtained from social media platforms where the relevant banks could be interpreted. Banks were examined one by one; Analyzes about word density were applied, wordcloud data visuals were created, and the perspective on banks was measured with individual sentiment analyses. As a result of the study, the most frequently cited by bank customers are The ease, usefulness, and service fees of digital applications are interpreted. Therefore, it has been understood that the digital services of private banks and public banks do not differ much, but the digital services of private banks are more efficient in terms of usefulness and self-renewal. As a result of the analysis, different suggestions were made to banks within the scope of customer satisfaction and quality service delivery in terms of digital banking services.

**Keywords:** Digital banking, Sentiment analysis, Text mining, Word cloud

***Öz***

*Günümüz dijital dünyasında önemli bir konfor alanı sağlayan sanal hizmetler, insanların çoğunluğu tarafından kullanılmaktadır. Bu anlamda Dijital bankacılık da en çok faydalanılan çevrimiçi finans hizmetlerindendir. Bu araştırmada da, banka müşterileri tarafından oldukça yüksek oranda kullanılan dijital bankacılık hizmetleri ile yapmış oldukları yorumlardan oluşan veri havuzu kullanılarak metin madenciliği yöntemleri ile analiz amaçlanmıştır. Çalışmada, Türkiye Bankalar Birliği verileri ışığında en çok kullanılan 10 adet özel banka ve 3 adet devlet bankası ile toplamda 13 adet bankaya ait dijital bankacılık verileri anakütleyi oluşturmaktadır. Veriler, Ocak 2020 ile Ağustos 2022 tarihleri arası baz alınarak elde edilmiştir. Toplamda 1.200.000-1.250.000 arasında ham veri ilgili bankaların yorumlanabildiği sosyal medya platformlarından elde edilmiştir. Bankalar tek tek incelenmiş olup; kelime yoğunlukları hakkında analizler uygulanmış, wordcloud veri görselleri oluşturulmuş ve yine tek tek duygu analizleri ile bankalara bakış açısı ölçümlenmiştir. Çalışmanın sonucunda, banka müşterileri tarafından en çok; dijital uygulamaların kolaylığı, kullanışlılığı, hizmet ücretleri yorumlanmıştır. Dolayısıyla özel bankalar ile kamu bankalarına ait dijital hizmetlerin çok fazla farklılık göstermediği ancak özel bankalara ait dijital hizmetlerin kullanışlılık ve kendini yenileme anlamında daha verimli olduğu anlaşılmıştır. Yapılan analizler sonucunda bankalara, dijital bankacılık hizmetleri açısından müşteri memnuniyeti ve kaliteli hizmet sunumu kapsamında farklı önerilerde bulunulmuştur.*

**Anahtar kelimeler**: *Dijital Bankacılık, Duygu Analizi, Kelime Bulutu, Metin Madenciliği*

*Burcu SAYIN OKATAN; burcusayinokatan@gumushane.edu.tr

## 1. Introduction

All people worldwide are affected by technological advances, people, societies, and economies are affected by technology development and renew their life routines to these developments. The majority of the sectors have gone to change in service, functioning and many more areas due to technological advances.

The banking sector is among the areas most affected by technological developments. With these financial developments, many different fields such as different banking terms, new different products, digital services and products, virtual money have entered the economy sector. Therefore, it is inevitable for the banking sector affected by technological developments to be digitized and become a digital sector in this field (Laukkanen, 2014). Especially digital service opportunities, the increase in the use of computers and mobile phones have become an increasingly important element as new digital services implemented in the banking sector. The share of smart mobile phones with a particularly high usage rate in this area is quite high. Mobile phones have an important place in this sector with digital banking applications. Digital banking services are now being used more instead of traditional banking services. Thus, the banking sector, which has survived the time limit, can serve 24 hours a day. Therefore, physical competition in the banking sector leaves its place to virtual competition in the sense of digitality (Johnsson, 2005; Coelho & Easingwood, 2003).

The developing digital banking sector has made more progress with the Covid-19 outbreak, which has completely taken the world from 2020. In other words, the digital banking sector, which is progressing at normal speed in terms of meeting the need, was among the most important digital services of the period by showing a development at a much higher speed with the pandemic period. In addition, people who already use digital banking have developed digital applications more in terms of use and started to perform their transactions faster. On the other hand, banks that maintain their traditionality have been compelled to move their services to the digital environment in order to keep up with increasing demand and developing technology and have developed well in terms of digitalization (Andrian et al., 2022). Competition in the banking sector in the world has made further progress with digitalization. This situation increased significantly during the Covid-19 Pandemic period. People who have to be independent from the outside world have received online banking services with internet banking and digital banking applications, so banks that have been able to offer this digital service in the best way have been the winners of this period. Therefore, digitized traditional banks and only virtual banks have increased competition by fueling competition (Onashabay, 2021).

Customers can be shown as the most important banking element in the digitalization and development of banks. One of the most important benefits of this changing banking sector can be shown as people's communication with banks. Much more comfortable and fast access is now available. Accordingly, customers who can evaluate this communication in the internet environment have access to service evaluation. These evaluations usually consist of comments that the bank has been criticized positively or negatively by the customer party. These customer comments are mostly made through banks' internet branches, mobile applications, various forum sites and many social media outlets. These comments provide an important repository in terms of the digitalization of the banking sector and its ability to provide a more efficient service in this field (Andrian et al., 2022). Today, the banking sector now, because it performs a large part of its services in Turkey and the world from digital media, development is still realized thanks to customer feedback.

In this research, in the analysis of customer comment data; Text mining methods were used using the python programming language. When the literature on research given in the other section was examined, many studies were found on this subject. The digital banking sector, which has become a particularly popular issue today, where digitalization has accelerated, has become the subject of most different areas. It was taken with customer comments obtained over the internet and text mining methods of different data of the bank and different analyzes were carried out. Accordingly, various forward estimates have been made. The research first mentioned the importance of working with the introduction and continued with the literature review. Later, methodology and implementation were mentioned, and analysis results were evaluated on the server, suggestions and discussion sections were organized on the subject.

## 2. Literature

Facing the large repository brings with it the problem of obtaining the desired data and arranging it. One of the important areas where the desired data acquisition is carried out is the text mining methods that make up

the methodology part of the study. Artificial intelligence techniques are used in text mining methods. There is not much work on digital banking, which is examined by text mining methods. The research on the closest to the research in this field is the 2021 study by Chang and his friends. In this research, artificial intelligence has been applied to provide information and find a viable process that will play a crucial role in providing innovative value in environmental education. In this research, which is stated as two main areas of artificial intelligence and natural language processing (NLP) intelligence agents leading the artificial intelligence trend. Between 2011 and 2020, NLP was adopted for the analysis of research topics of environmental education research journals in the Web of Science (WoS) database, and summary categories and features for environmental education articles were interpreted. In the research, ash data was selected from summaries and keywords of research journal articles analyzed by text mining, cluster analysis, secret Dirichlet allocation (LDA) and co-word analysis methods. By comparing the results of the research, most categories known by K-vehicles and LDA methods have been shown to be the same and share similar words; however, minor differences were obtained in the two categories. Participation of field experts in the research; helped consistency and accuracy of classified topics and documents (Chang et al., 2021). Again in 2021, emotion analysis (SA) techniques were used by Ogunleye and Oluwatoba to investigate customers' attitudes towards banking using Twitter data. With the bet they faced in Ogunleye and Oluwatoba studies; The unstructured nature of the data and word uncertainty stated that it complicates emotion analysis. In the context of this study, it was stated that it is more difficult than DA, Pidgin English in the bank area and the processing of English words in natural language, and it was stated that there was a problem in this regard (Ogunleye & Oluwatoba, 2021).

Another text mining study is the research of Park and his friends 2022. It is an emotion analysis study based on the shooting and processing of twitter shares of a particular region constructions in America with text mining methods. Park and his friends (2022) aimed to determine the impact of the land bank program on housing satisfaction using Twitter data. Approximately 300,000 Twitter posts with location information created in the city of Detroit in the research. It was collected to determine the degree of sensitivity to each tweet and categorized as positive and negative emotions with emotion analysis to determine the relationship between housing satisfaction and the land bank program. As a result of the research practice, the increase in ownership, the year of construction, the number of properties sold by home value and land banking; It has been determined that it has a negative impact on neighborhood satisfaction in Detroit. Although the results of the study showed that the land bank program did not significantly increase housing satisfaction in Detroit; It has been stated that it makes a partial contribution to the improvement of living standards (Park et al., 2022). Another research is the research of Andrian and his friends in 2022. In this research, customer satisfaction of digital banking in Indonesia was measured based on the emotion analysis on text data obtained by text mining from Twitter (Andrian et al., 2022). Mazinani and his friends aim to introduce the current literature in this field in 2022. In addition, in this research, each system was compared with the rest and its main distinguishing factors were determined. With this article, research that tries to estimate stock prices based on text mining and sensitivity analysis in a systematic review document has been summarized and systematized. Accordingly, research findings showed that text mining analyzes are often used in stock estimates and capital market issues (Mazinani et al., 2022). Again, most recently, Ngo and his friends (2023) contributed differently to the current literature on CBDC (Central bank digital currency). First of all, this study, which uses social media to investigate public discourses about CBDC worldwide, has identified the most frequently discussed issues related to CBDC among the public worldwide. Unlike this study, authors who comment on the main results of text mining, emotion analysis and regression models measure the attitude towards the currency specifically released (Ngo et al., 2023).

**Table 1.** The studies closest to the research in the literatüre

| | Author, Year | Subject |
|---|---|---|
| Studies examining digital banking with Text Mining methods | Chang et al., 2021 | "Applying Text Mining, Clustering Analysis, And Latent Dirichlet Allocation Techniques for Topic Classification Of Environmental Education Journals" |
| | Ogunleye & Oluwatoba, 2021 | "Statistical learning approaches to sentiment analysis in the Nigerian banking context" |

**Table 1** Continue

| | Author, Year | Subject |
|---|---|---|
| Studies examining digital banking with Text Mining methods | Park et al., 2022 | "Analysis Of Residential Satisfaction Changes by The Land Bank Program Using Text Mining" |
| | Andrian et al., 2022 | "Sentiment Analysis on Customer Satisfaction of Digital Banking İn İndonesia" |
| | Mazinani et al., 2022 | "A Comprehensive Review of Stock Price Prediction Using Text Mining" |
| | Ngo et al., 2023 | "Governance and monetary policy impacts on public acceptance of CBDC adoption" |

## 3. Method and application

In the research using text mining, text mining was briefly introduced, its varieties were mentioned, and the methods applied were specified.

### 3.1. Text mining

Text mining provides a framework for maximizing the value of information requested from a large text data; therefore, the rate of preference of text mining techniques has increased steadily in recent years and has created a variety of excess (Hassani et al., 2020). Also, due to the structural centrality of the natural language text, text mining; It also benefits from developments in other computer science disciplines related to the study of natural language. As an important element, text mining uses techniques and methods in computational linguistics based on information extraction, information retrieval and compilation of information obtained (Gonzales et al., 2016). Therefore, text mining refers to the extraction of implicit potentially valuable information and patterns from high-volume unstructured textual data such as natural language texts, automatically or semi-automatically unknown (Feldman & Sanger, 2007).

Text mining is a subfield of data mining that enables obtaining valuable new information from unstructured (or semi-structured) sources (Feldman & Sanger, 2007). The system scheme of text mining is shown in summary as follows.

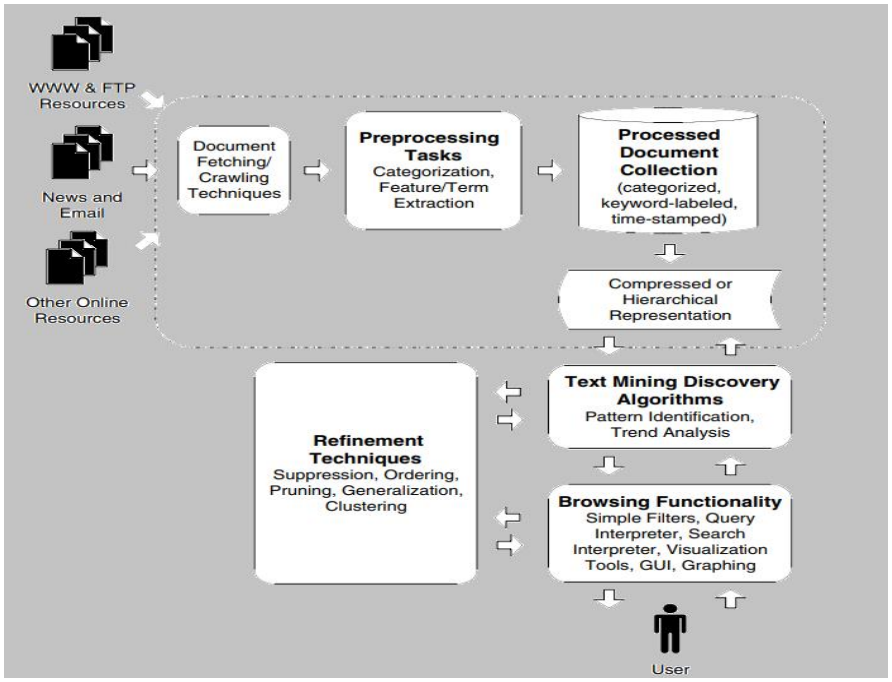


**Figure 1.** System architecture for generic text mining system (Feldman & Sanger, 2007).

### 3.1.1. Areas where text mining is applied

Text mining is an area that should be focused on as a general concept, similar to the field of data mining. Applications, techniques, and tools under this concept roof have different uses according to different fields of study and disciplines to which text mining is linked. The areas where Text Mining is used are general main topics; document classification and document clustering, information access and extraction, natural language processing, concept extraction and web mining (Miner et al., 2012).

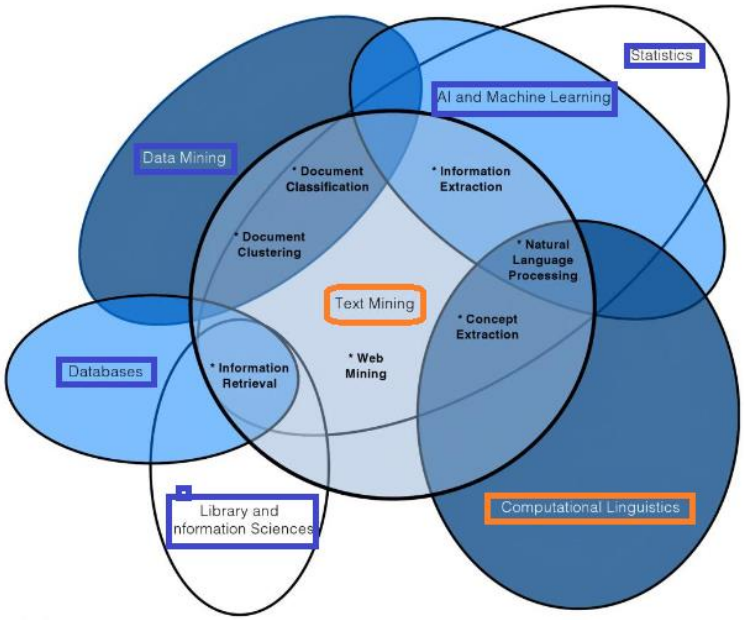


**Figure 2.** Areas where text mining is used (Miner et al., 2012)

Application areas where text mining is related and used directly are shown. Below are briefly stated in the definitions of these areas.

Document Classification: Using data mining classification techniques; are applications for grouping and categorizing documents, paragraphs, and terms. In these applications, classification model data is trained with labeled data. It is a controlled type of application (Miner et al., 2012).

Document Klustering: Using data mining clustering techniques; It includes grouping and categorizing documents, paragraphs, and terms. (Miner et al., 2012). It is especially used in text analysis discoveries. Application is carried out using unsupervised learning algorithms (Zhai & Massung, 2016).

Information Retrieval: It covers the search, storage and acquisition of text data by keyword (Miner et al., 2012: 35). It is the process of bringing many relevant information in connection with the user's search data to obtain many information needed (Zhai & Massung, 2016).

Information Extraction: is the process of obtaining information and information relationships from textual data that is unstructured. It is therefore to obtain structured text data from semi-structured or unstructured text data (Miner et al., 2012).

Natural Language Processing: a text mining area that is mostly synonymous with computational linguistics, where language processing and language understanding tasks are applied in the lower plan (Miner et al., 2012). Syntax applications obtained by semantic analyzes are examined within the scope of this field.

Concept Extraction: The process of dividing words and word groups into semantically similar groups (Miner et al., 2012). It includes the process of summarizing the text by identifying all conceptual possibilities in textual data and determining the most important as a result of algortimic values (Villalon & Calvo, 2009).

Web Mining: applications made in the web environment of text mining with a focus on the interconnection and scales of texts (Miner et al., 2012). If it is a different definition; It is the area where controlled and unsupervised text mining algorithms are used in the automatic discovery of documents in the web environment, extracting information from these documents and creating general patterns (Saini & Mohan Pandey, 2015).

Multiple text mining areas were used in the research. Information extraction and natural language processing areas have been particularly used.

### 3.2. Application

In this section, which could be the focus of the research, user comments were analyzed, and the data obtained as a result of this analysis was shared.

#### 3.2.1. Purpose and importance of the research

As stated in the literature section, a few studies conducted in the light of evaluation data of digital banking services were examined mostly on the basis of a single or a few banks. Additionally, a local research effort was made to determine customer attitudes towards digital banking. Unlike the above-mentioned studies, the purpose of this study is to analyze the digital banking interpretations obtained from freely accessible forum sites and Google Play applications as data sources, by examining the detailed word and emotion densities using text mining methods, and to analyze these digital applications. This study, which aims to look at digital banking in today's Turkey from the customer's perspective, aims to provide clearer inferences by analyzing realistic and first-hand data by applying text mining methods.

#### 3.2.2. Application model and tools used

The text mining application stages are illustrated in the diagram below.

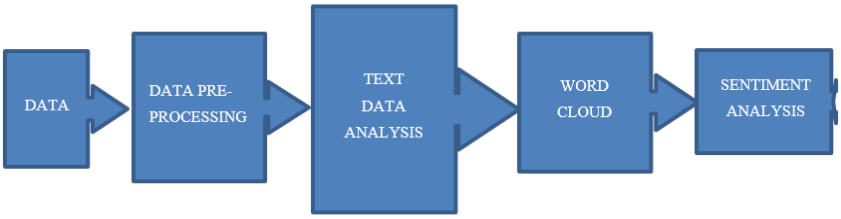


**Figure 3.** Text mining analytics application model (Sayın Okatan, 2023)

The tools used in the study are as follows;
• Many Social Sharing Forum Sites and Google Play constitute data sources
• Python Programming language
• Jupiter Notebook

Various data extraction operations were performed, and text mining was performed on Jupiter Notebook using the Python programming language used in the study. Data pre-processing, machine learning, data analysis and data visualization were performed with the obtained data. Jupyter Notebook is a type of working environment in which we analyze the texts obtained from the internet, run the codes written in Python language, and analyze the cleaned and preprocessed data, so all operations are carried out in this environment.

#### 3.2.3. Application analysis

After determining the working problem, Turkey Banks Association data in the light of the most used 13 Bank data created our general sample set. It was determined as the date range between January 2020 and July 2022.Digital banking comments from 12 banks in total, including 3 state banks and 10 private banks, created the dataset. Data extraction was done with the help of various codes using the Jupiter notebook. In the shooting process, Python programming language was used and completed using various key codes. The total number of uncleaned data is between 1,200,000-1,250,000, and as a result of pretreatments, the total data is determined

around 580,000. On average, there are close to 50,000 data per bank, and certain elements from this data are (News Texts, With the ready expressions and the data that the text general makes no sense, cleaning operations were made by making) and the analyzes were applied as such. Letters determined by the author for the reason that there is no ethical violation regarding the use of their data in the banks study and does not create any perception of bias; B, U, R, C, S, A, Y, I, N, Data outputs were shared, and their analysis was examined by specifying O, K, T, N, G.

The data obtained in the research were processed in the Jupiter notebook with the help of various text mining methods and encodings using the python programming language, taking into account the subject model of the text mining method. Accordingly, the application model and method review are given in detail in the previous sections. In addition, the tools used in the study were mentioned in the application model section. In the study, data preprocessing and machine learning, data analysis and data visualization were made using the Python programming language. Jupyter Notebook refers to a kind of working environment where we analyze texts and run codes written in Python and analyze text data. There are many coding environments where the Python language is used, as well as many non-python programming languages using the Jupiter notebook. However, in the use of the python language as a reason for preference, the Jupiter environment is clear and understandable in terms of design and can be shown to offer an interactive open source working environment. From the data collected, intermediate results were obtained for each bank, which were created in excel form separately. This part, which includes the removal of text from plug-ins like many shape video websites, is the first and very important stage of research. It is also the first stage of text mining, known as pretreatment, to purify this data from punctuation and helpful words - which will not work in analysis.

Preliminary processes, also known as extortion processes, can be shown as the most basic cleaning methods in text mining (Gupta & Lehal, 2009). These extration processes are the processes for editing the comment data we receive as data. So, configuring data is the primary goal at this stage. Text mining, preprocessing the textual data obtained, storing intermediate results for use when necessary, converting the data into structured data with this pretreatment, using some techniques to pass structured data through various analyzes and visualizing the final results (Vijayarani et al., 2015).

In addition, text preprocessing is a method that transfers text to a machine-readable format for further analysis. Text includes steps such as preprocessing, text normalization, rooting and symbolization (Chintalapudi et al., 2021). Text mining analyzes were applied because the data obtained about banks over the Internet are text data. Interpretations are syntactically in different formats. Therefore, they have been subjected to some pretreatments. Various libraries of the python program were used for the analysis of the data set. Selenium, which helps to extract text data from sites with certain input information, in processing the data drawn in the field of Natural Language Processing, NLTK (Natural Language Toolkit), a library containing open source Python modules and datasets to manage and develop our research; reading data in the project, which allows us to work with multi-dimensional arrays and matrices, NumPy ( Numerical Python), Pandas libraries are used, ensuring the availability of data preprocessing and data cleaning stages. String and Random libraries are used, providing complex variable substitutes, the ability to format values and generate random numbers with the method described in the text in the form of built-in string class (Desai, 2019).

Accordingly, each bank comments whose data were withdrawn at the stage of pretreatments (disconnection-extration operations) were handled separately. Using NLTK, Pandas, String and Numpy libraries; dividing the text into desired shapes (usually divided into words) and each bank comments are handled separately to save it to the series, words were cut into pieces and converted to string very format with tokenization i.e. Tokenization (Icification) processes. Punctuation (Nokaat) process was applied by removing the punctuation marks from the text produced by this conversion, and the frequencies of the words were calculated with the FreqDist function and the auxiliary words that were repeated very often. The FreqDist function used here and also in wordcloud analysis is an important step in determining word densities.

Then, the removal of words in the text will not make any difference in meaning; stopping words, pronouns, prepositions, words that do not 'add value' to the research context, etc. Stopwords (Meaningless word extraction) has been implemented to extract unnecessary words such as. In this way, the size of the text data is reduced, and faster processing of the text data is provided (Feldman & Sanger, 2007). Lemmatization (Rooting) has been processed to remove attachments of words, which are the continuation of these word extraction

processes and another step-in simplification. Then in the text with the help of various encodings; Hashtags, usernames, URL links and numerical expressions starting with '#' have been removed.

The texts were cleared of the above-mentioned characters and all elements that prevent analysis; In addition, the normalization process was made, and the regulation was made and made suitable for the analysis.

Word Cloud is a type of analysis that is the visualization of abstract data with schematic or graphical tools. The word cloud was determined as the first data visualization tool used in the study. Word Cloud is an image of many words that imply the content of the analysis document. The words are presented in the section where the application outputs in various shapes and colors are examined according to the frequency and importance of the document. The larger the text form, the higher the number of formations (Jadhav et al., 2022). For this reason, the validity of word cloud outputs is higher due to the excess of the processed text data.

In emotion analysis, Wordcloud is used to find the frequency of dominant words and then draw conclusions based on the subject and condition of the study (Mustaqi et al., 2020). So, it offers the impression of a kind of issue distribution map. Word intensities in certain periods are shown after the extraction of entries (text pretreated) received and edited from forum sites and google play over the Internet. In this analysis, the WordCloud library was used, and data visualized as a used data visualization provider that also tems text data that specifies the dimensional frequency and importance of each word. The wordcloud outputs for each bank are discussed separately, with emotion analysis outputs and sample comments in which the bank's general opinion is understood, in the ubular section.

In the last stage of implementation, it was tried to understand the issues that customers are general bleeding of banks and especially positive/ negative/ neutral from bank comments by analyzing emotions. Emotion analysis, an active research area that has emerged recently; It is a discipline that extracts people's emotions, thoughts, thoughts, and behaviors using the Natural Language Processing (NLP) methods from users' text data (Danneman & Heimann, 2014). Another method used in data visualization in the study is emotion analysis. Emotion analysis is defined as a field where written language analysis is conducted, which allows people to evaluate their emotions, thoughts, and opinions on a particular subject. The Internet has made the lives of people who are highly used in digital banking, especially in the pandemic process and after, and thus created an environment where emotions and thoughts are reflected through social networks. For this reason, emotion analysis has become one of the most preferred analysis tools in text mining and natural language processing due to increased data flow (Al-Hashedi et al., 2022).

Emotion analysis, also defined as idea mining, is often used with the help of python programming language NLTK (Natural Language ToolKit) and TextBlob libraries. TextBlob's output is polarity and subjectivity. Polarity score (-1 to 1); here -1, 'lousy', 'bad', 'useless' and 1, 'beautiful', 'good', 'useful'… defines the most positive words like. The subjectivity score is between (0 and 1), indicating the amount of personal thinking. If a sentence has a high subjectivity and close to 1, it appears that the text contains more personal opinions than real information (Sharma et al., 2014). Emotion analysis was made using textblob, matplotlib and seaborn libraries in the study; A diagram image of this sensitivity analysis was created with neutral, positive and negative variables. To be detailed, the polarity values of the comments were obtained with the textblob library, and these polarity values were visualized as a distribution graph using the seaborn library. The outputs obtained as a result of the analyzes made accordingly will be examined in the findings section below.

## 4. Results

Since the figures obtained as a result of the analyzes made in this section are used in Turkish data, their contents are Turkish, and their originality is preserved. Emotional analysis charts with word cloud outputs and polarities of comments are provided.

### 4.1. B public bank analysis results

As can be seen from the wordcloud output when the comments of Bank B are examined, many customer comments have been reached that there are serious problems with the entry processes regarding the Use of Mobile Bank and that the system has discarded them even if the entry is made. In general, it was noticed that there were serious problems related to the application. The fact that the R&D of the Mobile Banking

application, which is an important system of the digital foot of the bank in the word cloud and in this regard, has not been able to respond to customer needs, means that the use of the mobile application is difficult. If the results of the emotion analysis are to be looked at, the intensity is at minus polarity values, mostly involving customers' negative comments about mobile banking. When the comments are examined; It is noteworthy that the comments describing the entry problems in relation to the introduction have negative polarity values. When the (-1,1) interval and (-0.5, 0.5) range setting is examined separately, the negative occurrence of the majority of comments is remarkable, as the negative polar comments show intensity.

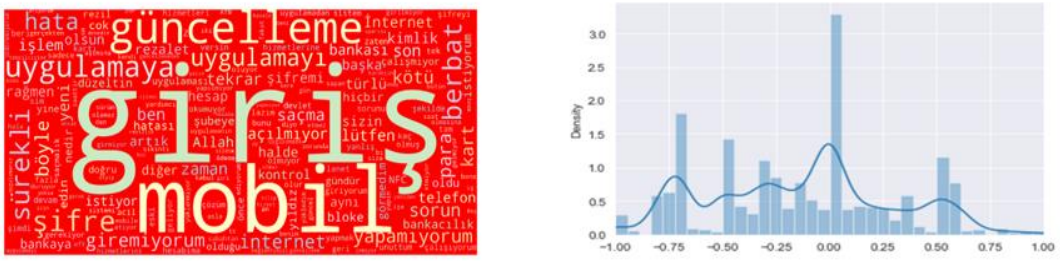


**Figure 4.** Wordcloud and sentiment analysis chart of B public bank

### 4.2. Analysis results of U public bank

Looking at the comments of the U bank, the status determinations related to the update are clearly included in the banking transactions over the internet and in the mobile application.

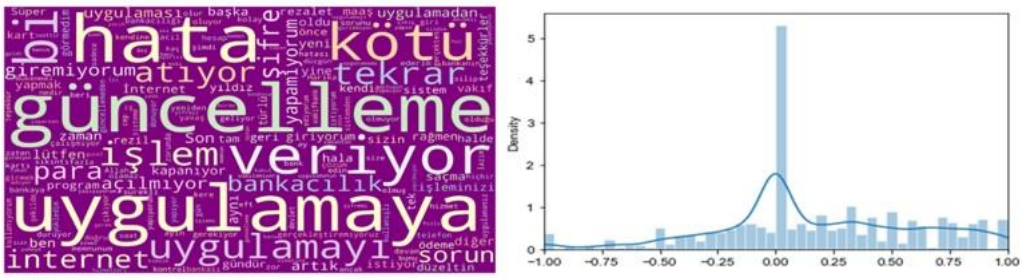


**Figure 5.** Wordcloud and sentiment analysis chart of U public bank

### 4.3. R public bank analysis results

R is a public bank and many difficulties have been expressed in the comments. It is mentioned that there were many comments especially in 2020 and early 2021 and that there was a victimization caused by the constant error of the application. As seen from the word cloud, the overuse of the word "error" is due to comments made about the R bank's mobile banking, which has made a lot of mistakes. The trouble encountered in the practice of the bank in particular caused customers to write a lot of comments on this matter.

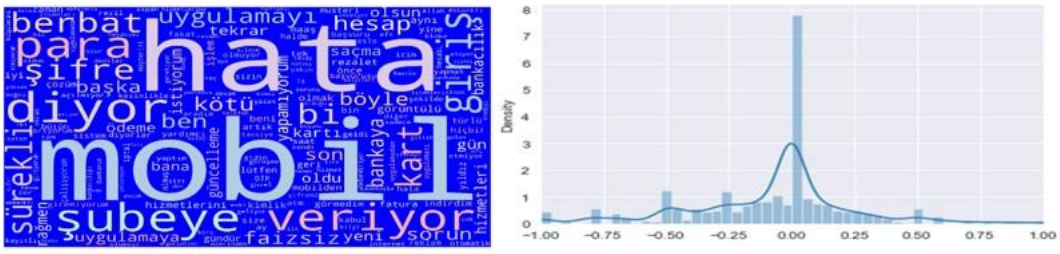


**Figure 6.** Wordcloud and sentiment analysis chart of R public bank

### 4.4. Analysis results of the C private bank

In the comments that the results of the analysis were analyzed, it was emphasized that it became much more complex with the updating of the system. It has been stated that its old applications are much more useful, and it has been shown that the update has been made in general, causing negative comments.

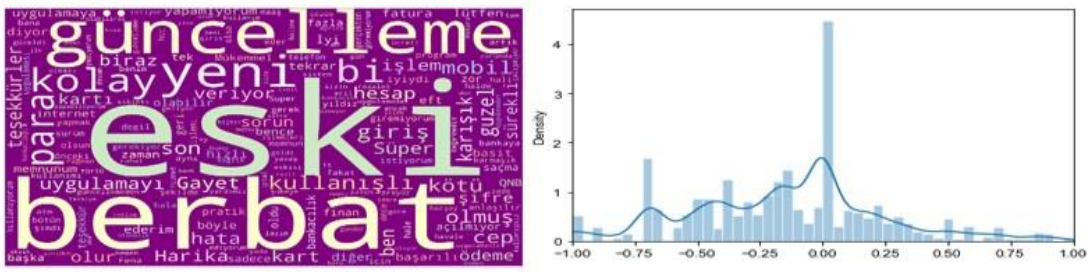


**Figure 7.** Wordcloud and sentiment analysis chart of C private bank

### 4.5. Analysis results of the S private bank

Looking at the comments of the S Private Bank, it is remarkable that the interpretations are seriously different and positive. In this sense, it is understood that private banks are different from each other.

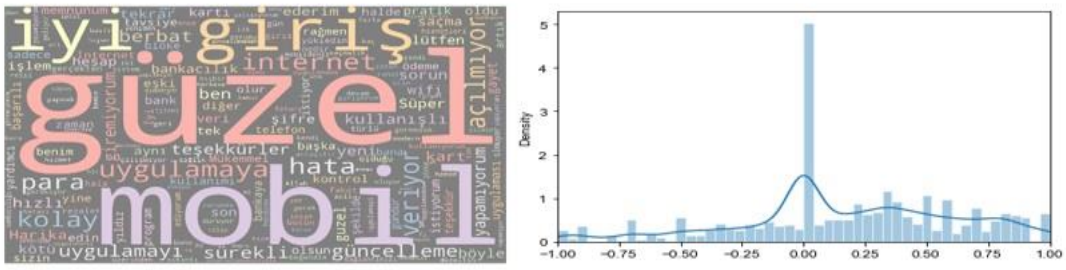


**Figure 8.** Wordcloud and sentiment analysis chart of S private bank

### 4.6. Analysis results of the A private bank

Considering the analysis results of bank, A, it is remarkable that the fees cut by the bank are criticized by customers.

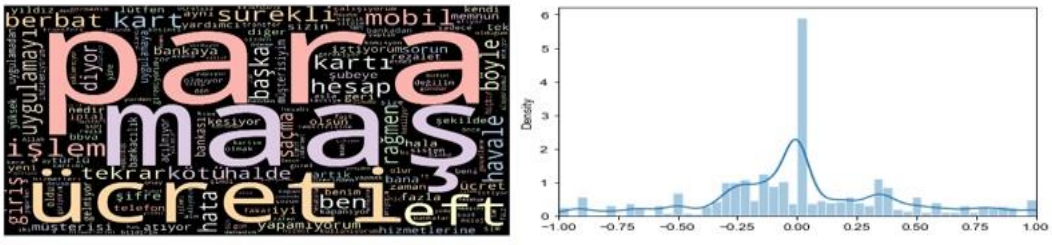


**Figure 9.** Wordcloud and sentiment analysis chart of A private bank

### 4.7. Analysis results of the Y private bank

Y Bank is a Private Deposit bank and is a bank that is generally liked by the customers in the first reviews. As seen at first glance, it consists of a good impression of digital banking from the word cloud.

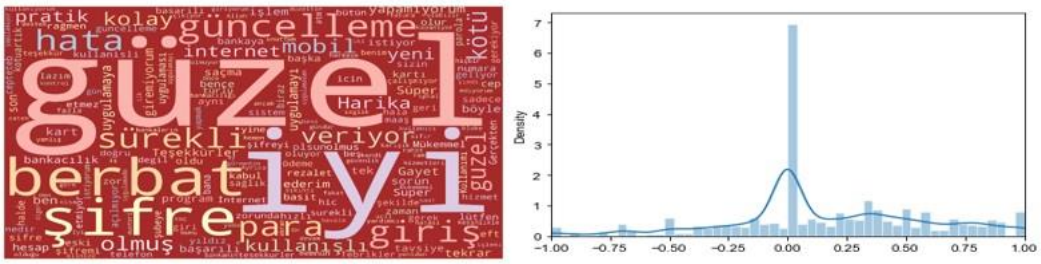


**Figure 10.** Wordcloud and sentiment analysis chart of Y private bank

### 4.8. Analysis results of the I private bank

In the majority of the comments made about Bank I, there is thanks and praise to mobile banking services, whether it be internet banking services or mobile banking services. Especially the usefulness, ease, practicality and openness of the application are praised. It is understood from the comments made that customers attach great importance to usefulness and convenience in digital banking. Because satisfaction in this matter generally praises the digitality of the bank.

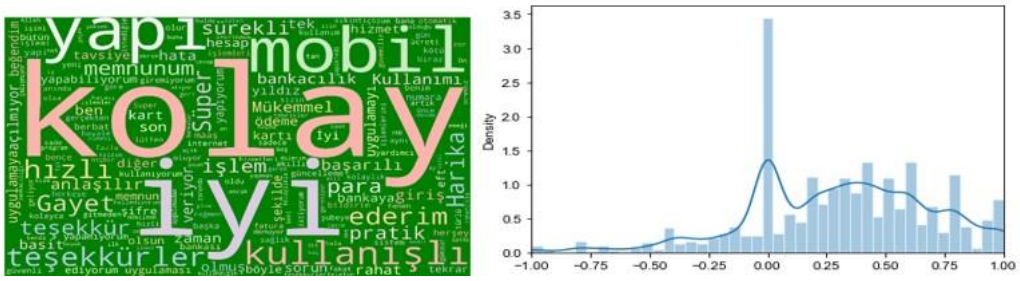


**Figure 11.** Wordcloud and sentiment analysis chart of I private bank

### 4.9. Analysis results of the N private bank

N private bank is a private deposit bank that has been serving for many years. The N bank, whose experience in banking is considered to be excessive, has a preliminary impression that contradicts the various analysis results made with customer comment data.

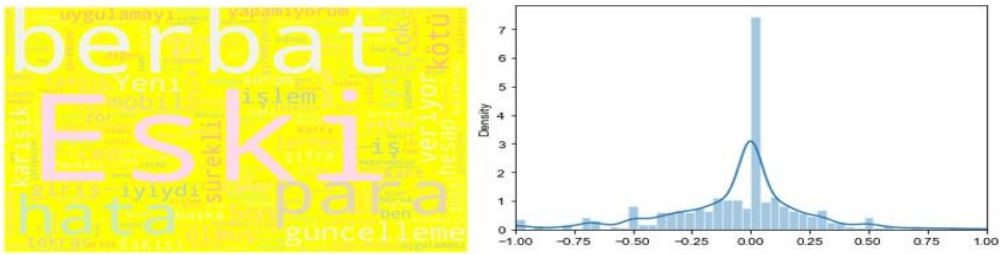


**Figure 12.** Wordcloud and sentiment analysis chart of N private bank

### 4.10. Analysis results of the O private bank

The results of the analysis from that bank reveal the bank's success in digital banking. Among the analyzes made, it became a bank that stands out with its positive positivity. Therefore, it can be said clearly that he is successful in digital banking related to the bank.

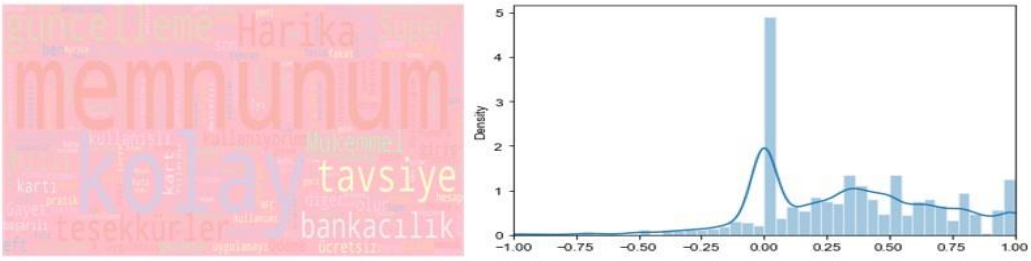


**Figure 13.** Wordcloud and sentiment analysis chart of O private bank

### 4.11. Analysis results of the K private bank

When the interpretative outputs of the K bank are examined, it shows an environment of satisfaction similar to that bank.

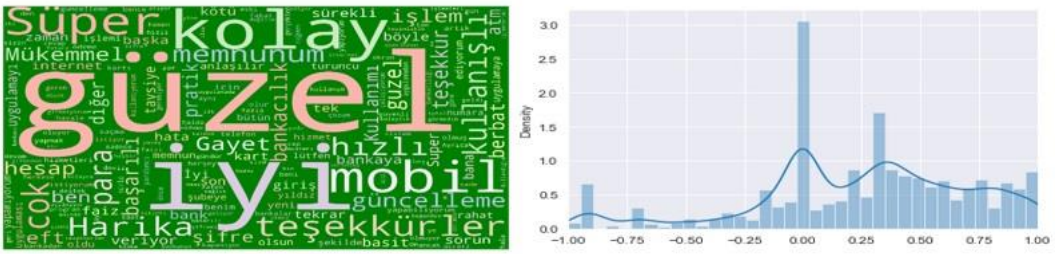


**Figure 14.** Wordcloud and sentiment analysis chart of K private bank

### 4.12. Analysis results of T and G private participation banks

Private participation banks are important in comparing their analysis with other banks and being compared in themselves. It has been tried to understand how customers' view of the digital banking of these participation banks and what direction their complaints are.

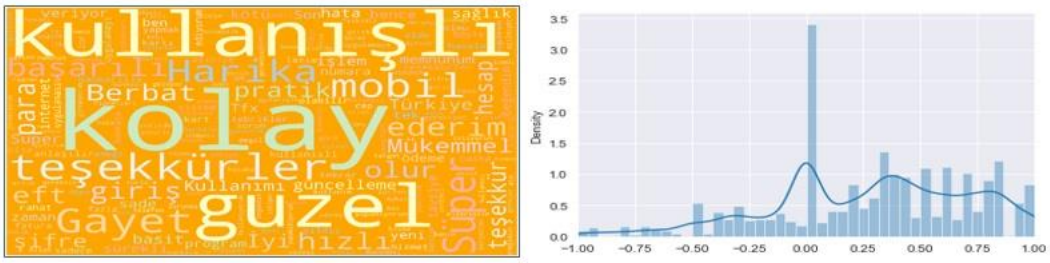


**Figure 15.** Wordcloud and sentiment analysis chart of T private participation bank

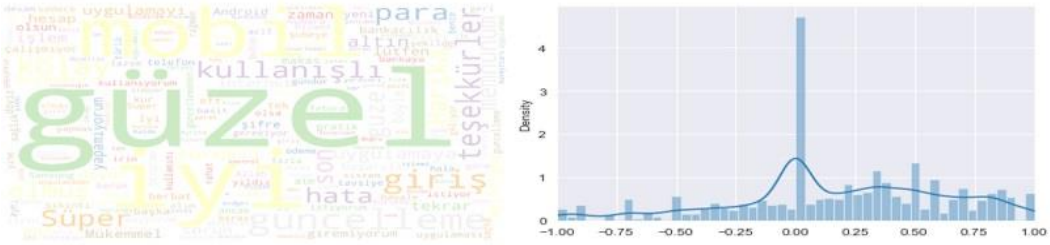


**Figure 16.** Wordcloud and sentiment analysis chart of G private participation bank

According to G Bank's T Bank, it shows the negative feeling more. In the emotion analysis, in which both banks evaluated approximately 25 thousand data analysis results, it was seen that the G bank was criticized more on some issues. It is understood that this criticism situation is due to the 'update' problem. People have made many comments regarding the inability to benefit from the services of the bank in question after the update.

## 4. Discussion

In the results of the analysis, data related to the development of banks and their own development were obtained. However, there are some limitations and points to be developed in the study. If the lack of forum sites that include complaints and comments about internet banking or if it is considered on sites that do not share data, less data on this subject is seen as a constraint of the study. In addition, the fact that the customer who has problems does not give a positive return that his problem has been fixed after writing a comment is another constraint. The analysis, which also compares to the public and private banks, is limited by the fact that the public participation banks are not analyzed. This study can be examined in more detail by adding public participation banks. Another point is Compared to the text data collected; it was understood that the number of comments on public banks was more limited than private banks. Although sufficient comment data has been obtained; this can be demonstrated as the constraint of the study on analyzing public banks.

With this study, an important contribution has been made to the literature in terms of evaluating the digital services of public and private banks through the eyes of customers and seeing the progress in digital banking services of recent years. With other similar studies, its difference was examined, and it was stated that they added it to the literature. It is stated that text data from open access forum sites and open access Google play sites are used over the Internet. It was the first digital banking study to use these resources. As a study close to research, Emotion analysis (SA) techniques were used by Ogunleye and Oluwatoba in 2021 to investigate customers' attitudes towards banking using Twitter data. The Pidgin English used in the research affected the analysis's proper results and it was stated that the desired efficiency could not be reached due to lack of data. The fact that the data density of this research is quite satisfactory, and the sample is wide shows its most important difference. Ngo and his friends (2023), CBDC (The central bank contributes in different ways to the current literature on the digital currency). Unlike this study, the authors who comment on the main results of text mining, emotion analysis and regression models measure the attitude to the currency that was released. Although this study is methodically similar, the evaluated variables differ. Customer satisfaction of digital banking in Indonesia was measured based on the emotion analysis on text data obtained by text mining from Twitter in the 2022 study of Andrian and others. In this article, which is very similar to this study, customers' overview of digital banking was tried to be measured. This research differs from that of Andrian and his friends, as banks are evaluated individually, densities are measured on a word basis, with detailed focal points specified, and data density is incomparably superior. This research, which is different in terms of the source and analysis of this text data with the textual data used, also differs from Cheng's research in 2021. In addition, Cheng and his Friends are concentrated on a single bank, and no information about rival banks has been obtained. The work done in terms of revealing digital banking differences with other banks that are important in terms of benchmarking differs. It is seen in the literature under review that it has been observed that studies trying to estimate stock prices based on the results of text mining word cloud and emotion analysis using together or in combination with different analyzes have gained intensity. With this study, which analyzed text data from the data source, where customers convey their ideas of direct application and internet banking, the perspective of more direct digital banking was revealed.

This study was determined by the word cloud from the dominant words, and it was understood that sufficient inferences could be made since the positive and gative perspectives were determined by the emotion analysis. However, by expanding the same study data, research can be conducted with different implications with LDA (Latent Dirichlet Allocation), which is the subject modeling method of text mining methods. In addition, by downloading the returns of banks in digital media, the return data can be examined accordingly and the developments in this sense can be explored.

## 5. Conclusions

The findings detailed in the application section show that; It has shown that customers use digital banking users much more than internet banking. When the comment data drawn over the internet is examined, it is

understood that customer returns are mostly related to mobile banking. Looking at the general review data, it seems that the mobile banking usage rate is much higher. In addition, it is understood from the comment data; It is understood that the pandemic period has a developmentally positive effect on digital banking. This inference is more about the height of the rate of comment after the pandemic period. The historical information of the comments received shows that the data for the period 2021-2022 are much more. This shows that the rate of digital banking use increased with the pandemic period. In particular, 2022 data shows that digital banking services continue to serve by seriously renewing itself and strengthening its digital infrastructure. However, it is understood from wordcloud outputs that; customers had problems with these renewal and updates of some banks. The fact that the website or application slows down, freezes, and problems with excretion from the system causes customers to give a revision by commenting on these banks, especially when logging in or performing a transaction. These criticisms, which express the negative side especially in the emotion analysis graph, did not leave their place to positive comments when the bank services improved. They are also frequently mentioned in comments in technical status assessments, corrections, renewals and updates. The renewal and updating of an application system that is especially accustomed to causes some criticism of the customer.

The most common in the application findings by customers; usefulness, convenience, service fees, especially in mobile banking, the application has been mentioned to be comfortable and fast. In this sense, in digital banking services, the distinction is noticed between private banks and public banks, even if there are no excessive differences. It has been understood that private banks have updated much more frequently, have more renewal in their digital services and have a solid infrastructure as it is stated to be faster and easier to reach. State banks follow private banks in the back, but it has been understood that it has closed this difference, especially according to the 2022 comment data. In addition, private participation drew attention to the positive difference that does not appear in the word cloud of banks and appears in emotion analysis. The difference between public banks and private banks can be seen as equivalent to the difference between private banks and private participation banks. If it is necessary to rank briefly from the best to the less good; private participation banks, private banks and public banks can be shown.

It is understood that mobile and internet baccarat is more preferred by mobile banking customers within the scope of digital banking. Mobile baccarat gives more confidence in this than being standard, and the fact that smartphone usage is high is the biggest factors. With the satisfactory size of internet banking by becoming more standard, the usage rate can be increased, if not as much as mobile banking.

**Acknowledgement**

This article was produced from Burcu Sayın Okatan's doctoral thesis.

**Author contribution**

Res. Ass. Dr. Burcu Mr. Okatan: Creation of data, analysis and evaluation, writing articles.
Assoc. Dr. Handan Çam: Interpreting the results, writing an article.

**Declaration of ethical code**

The authors of this article declare that the materials and methods used in this study do not require ethics committee approval and/or legal-specific permission.

**Conflicts of interest**

The authors declare that they have no conflict of interest. The authors declare that they have no conflict of interest.

**References**

Al-Hashedi, A., Al-Fuhaidi, B., Mohsen, A. M., Ali, Y., Gamal Al-Kaf, H. A., Al-Sorori, W., & Maqtary, N. (2022). Ensemble classifiers for Arabic sentiment analysis of social network (Twitter data) towards COVID-19-related conspiracy theories. *Applied Computational Intelligence and Soft Computing, 2022*, 1-10. https://doi.org/10.1155/2022/6614730

Andrian, B., Simanungkalit, T., Budi, I., & Wicaksono, A. F. (2022). Sentiment analysis on customer satisfaction of digital banking in Indonesia. *International Journal of Advanced Computer Science and Applications*, *13*(3). https://doi.org/10.14569/IJACSA.2022.0130356

Chang, I. C., Yu, T. K., Chang, Y. J., & Yu, T. Y. (2021). Applying text mining, clustering analysis, and latent dirichlet allocation techniques for topic classification of environmental education journals. *Sustainability*, *13*(19), 10856. https://doi.org/10.3390/su131910856

Chintalapudi, N., Battineni, G., & Amenta, F. (2021). Sentimental analysis of COVID-19 tweets using deep learning models. *Infectious Disease Reports*, *13*(2), 329-339. https://doi.org/10.3390/idr13020032

Coelho, F. & Easingwood, C. (2003). Multiple channel structures in financial services: a framework. *Journal of Financial Services Marketing*, *8*(1), 22-34. https://doi.org/10.1057/palgrave.fsm.4770104

Danneman, N., & Heimann, R. (2014). *Social media mining with R.* Packt publishing ltd.

Desai, R. (2019, December 26). *Top 10 python libraries for data science*. https://towardsdatascience.com/top-10-python-libraries-for-data-science-cd82294ec266

Feldman, R., & Sanger, J. (2007). *The text mining handbook: advanced approaches in analyzing unstructured data.* Cambridge University Press, 35-46.

Gonzalez, G. H., Tahsin, T., Goodale, B. C., Greene, A. C., & Greene, C. S. (2016). Recent advances and emerging applications in text and data mining for biomedical discovery. *Briefings in Bioinformatics*, *17*(1), 33-42. https://doi.org/10.1093/bib/bbv087

Gupta, V. & Lehal, G. S. (2009). A survey of text mining techniques and applications. *Journal of Emerging Technologies in Web Intelligence*, *1*(1), 60-76.

Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, *4*(1), 1. https://doi.org/10.3390/bdcc4010001

Jadhav, A., Kaur, M., & Akter, F. (2022). Evolution of software development effort and cost estimation techniques: five decades study using automated text mining approach. *Mathematical Problems in Engineering*, *2022*, 1-17. https://doi.org/10.1155/2022/5782587

Jonsson, S. & Gunnarsson, C. (2005). Internet technology to achieve supply chain performance. *Business Process Management Journal*, *11*(4), 403-417. https://doi.org/10.1108/14637150510609426

Laukkanen, T. (2014). Internet vs mobile banking: comparing customer value perceptions. *Business Process Management Journal*, 788. https://doi.org/10.1108/14637150710834550

Onashabay, N. (2021). *Effects of COVID-19 pandemic on the key profitability factors of digital challenger banks. Starling Bank case study* [Doctoral Dissertation, Central European University].

Ogunleye, B. O. (2021). *Statistical learning approaches to sentiment analysis in the Nigerian banking context* [Doctoral dissertation, Sheffield Hallam University].

Mazinani, M. T., Hassani, H., & Raei, R. (2022). A comprehensive review of stock price prediction using text mining. *Advances in Decision Sciences*, *26*(2), 1-36.

Miner, G., Elder IV, J., Fast, A., Hill, T., Nisbet, R., & Delen, D. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Academic Press.

Mustaqim, T., Umam, K., & Muslim, M. A. (2020). Twitter text mining for sentiment analysis on government's response to forest fires with vader lexicon polarity detection and k-nearest neighbor algorithm. *In Journal of Physics: Conference Series*, *1567*(3), 032024. IOP Publishing. https://doi:10.1088/1742-6596/1567/3/032024

Ngo, V. M., Van Nguyen, P., Nguyen, H. H., Tram, H. X. T., & Hoang, L. C. (2023). Governance and monetary policy impacts on public acceptance of CBDC adoption. *Research in International Business and Finance*, *64*, 101865. https://doi.org/10.1016/j.ribaf.2022.101865

Park, S., Lee, J., & Park, Y. (2022). Analysis of residential satisfaction changes by the land bank program using text mining. *Sustainability*, *14*(18), 11707. https://doi.org/10.3390/Su141811707

Saini, S., & Mohan Pandey, H. (2015). Review on web content mining techniques. *International Journal of Computer Applications*, *118*(18), 33–36. https://doi.org/10.5120/20848-3536

Sayın Okatan, B. (2023). Investigating customer reviews on digital banking applications through text mining methods, [Doctoral dissertation, Gümüşhane University Graduate Institute].

Sharma, R., Nigam, S., & Jain, R. (2014). Opinion mining of movie reviews at document level. *arXiv preprint arXiv:1408.3829.* https://doi.org/10.48550/arXiv.1408.3829

Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication networks*, *5*(1), 7-16.

Villalon, J., & Calvo, R. A. (2009). Concept extraction from student essays, towards concept map mining. In *2009 ninth IEEE International Conference on Advanced Learning Technologies*, 221–225. IEEE. https://doi: 10.1109/ICALT.2009.215

Zhai, C., & Massung, S. (2016). *Text data management and analysis: a practical introduction to information retrieval and text mining*. Morgan & Claypool Publishers. https://doi.org/10.1145/2915031