The Indonesian Journal of Computer Science
www.ijcs.net
Volume 14, Issue 2, April 2025
https://doi.org/10.33022/ijcs.v14i2.4803

Exploring AutoText Summarization Methods in Turkish: A Literature Review

Neda Alipour[1], Hadi Pourmousa[2], Mohammad Naserinia[3]
nalipour@gelisim.edu.tr[1], hpourmousa@gelisim.edu.tr[2],
mohammad.naserinia@ogr.sakarya.edu.tr[3]
[1,2] Faculty of Economics, Administrative and Social Sciences, Management Information Systems Department, Istanbul Gelisim University, Istanbul, Türkiye
[3] Business Faculty, Department of Management Information Systems, Sakarya University, Sakarya, Türkiye

Article Information

Received : 18 Mar 2025
Revised : 30 Mar 2025
Accepted : 15 Apr 2025

Keywords

Text summarization, Natural Language Processing, Abstractive summary, Extractive summary

Abstract

In recent years, the huge volume of textual data has become a challenge, as this challenge is seen in various fields, including scientific articles, legal documents, Internet archives, and even online product reviews. Given the limited data processing capacity of humans, processing large amounts of data is impractical and causes confusion; on the other hand, it requires a lot of effort, which ultimately results in a waste of time. To overcome this problem, the need to implement automated techniques such as automatic text summarization has emerged. Automated text summarization is an automated technique used to create a more condensed version of the original content that provides the same meaning and information. In fact, the generated output should contain important information from the original document. Various techniques for automatic summarization have been proposed in studies. Many studies have been presented on automatic text summarization methods, however, limited papers have contributed to reviewing different techniques of summarization methods in different languages, so this topic is evolving to reach maturity. This study focuses on different automatic text summarization methods in Turkish by reviewing the literature and previous studies, thus analyzing the performance of automatic text summarization methods.

A. Introduction

Today, with the increase in the amount of information, summarizing the text becomes important [52]. The large volume of information and texts may make users avoid reading significant and interesting texts. Therefore, summarizing the text is a necessary solution [22]. Since manual text summarization is a protracted process, automatic text summarization is considered another method [38]. Research on automatic text summarization was first described by Luhn more than 60 years ago [62], [28]. Text summarization plays an important role in automatic content generation, meeting minutes, assistance to the disabled, and also in fast reading of online documents [57].

The purpose of text summarization is to summarize large text documents [62], [38]; nevertheless, the most critical information in the text should also be included in the summary [29], [41]. As a consequence, the user can understand the main aspects of the document without reading the entire document [9]. In fact, automatic summarization works in such a way that the text is converted into a compressed version and the general meaning of the text is preserved [52]. This means that the summarized text must contain information related to the original text but also faces difficulties. In other words, the issues that cause complexity in summarization can be mentioned, such as time, redundancy, sentence order, etc., which must be taken into account when summarizing a large number of texts [22]. Reducing reading time can be stated as the main advantage of using summaries [31], [42], and also, reduces costs [1]. In terms of structural components, a text summarization process consists of three steps: diagnosis, interpretation, and summary [41], [31], [61], [15]. In the definition phase, the main and important points in the text are determined, and a summary of the prominent points is prepared [17]. After extracting the main points of the text in the previous stage, the integration process will take place in the interpretation stage. In addition, there may be corrections in the original sentences at this stage. The result of the previous step may not be understandable to the reader, so it should be more consistent and formulated. The final editing takes place in the summary stage and as a result, it will become intelligible to the reader [41].

The text summarization process identifies important and substantial information found in many documents, and this work is done in two ways: extractive and abstractive [21]. An extractive summary is achieved by obtaining important sentences from the relevant text. By knowing the main concepts of the document and interpreting them in a new way, an abstractive summary can be made [60], [2], [34]. This can be achieved through various supervised and unsupervised techniques such as convolutional neural networks (CNN), recurrent neural networks (RNN), and linguistics research center (LRC). Machine learning is also used to analyze reviews on any e-commerce website such as Amazon, helping to gain insight into user preferences and behavior for submitting eligible items, as well as providing relevant reminders [57]. There are challenges associated with abstractive summarization. The main trouble of abstractive summarization is representation. The capabilities of automated systems are limited by the multiplicity of their representations and their ability to produce representational structures—abstractive summarizers cannot produce summaries of text that their structure cannot represent. It is possible to formulate appropriate representations

under limited categories, but a general solution is not possible and depends on general domain semantic representations. It is not possible to create automatic systems that fully understand and represent human natural language [24]. The challenges of extractive summaries are as follows: 1) Compared to medium summaries, extractive summaries are usually long because they may include certain parts of the text that are not required in the summary. 2) In most cases, essential information is usually placed on different lines, and if it is not long enough to cover all these lines, extractive summaries usually cannot collect them [13], [24]. Today, most research focuses on abstractive and real-time summarization because it supports more complex structure [59], but in turn, extractive summarization has been widely used in research since 1958 [58].

Research on text summarization and thus summarization techniques is of considerable importance and continues to mature [59]. The progress in summarizing texts and increasing summarizing techniques in recent years is undeniable [5]. Although automatic text summarization is not a new topic, researchers still pay close attention to this topic [8]. As a result, extensive studies on text summarization have highlighted the importance of this scope. At the same time, most investigations have been limited to a small number of languages (mostly English, Chinese, Arabic, and Spanish) [23]. On the other hand, some studies examined a smaller subset of techniques [41], [59]. In addition, since the automatic summarization of text in different languages has been researched in past research, Turkish language summarization techniques have always been of interest to researchers, so this research centralizes on the investigation of Turkish language text summarization techniques. Therefore, the purpose of this research is to manifest an overview of Turkish text summarization techniques that have been used over the years. In addition, this research helps to reveal the comparison between different techniques.

In the second section of the article, the theoretical background is explained; in the third section, the techniques used for text summarization in Turkish are discussed; and in the fourth section, all existing automatic text summarization techniques and their performances are explained, and a comparison is made between the techniques. In the last section, the results and the work that can be done in future studies are explained.

B. Theoretical Background

Machine Learning Techniques for Text Summarization

Supervised and unsupervised learning have been introduced as two aspects of machine learning techniques [12], [50], [35]. Supervised learning is a tool in machine learning [55]. Supervised learning supports input-output pairs, and its algorithm works by detecting the output corresponding to each input [35]. In this technique, the machine is supervised by a supervisor, and then the machine learns with validated and correctly labeled data [14]. Also, Shafiq et al. [54] have stated that in supervised learning, labeled data are trained to provide predicted outputs. In this learning technique, the outputs appear as confirmed; in fact, when the machine is injected with real data, new unknown data is produced in the results, and it will produce better and desirable results. It can be said that this technique plays a significant role in providing a suitable solution to complex computational

problems [16]. Supervised learning is further divided into two techniques: classification technique and regression technique [27], [37], [43]. In the regression technique, it is used to evaluate and predict the relationship between predictor variables and response variables; in fact, the response variable has continuous values [53] and the output produced is continuous and based on the data it was trained on. In the classification technique, as the name suggests, outputs are classified and given class labels [6]. The classification technique can be called binary or multiple classification according to the output categories [49].

In the unsupervised learning technique, the data is unlabeled [33] and the structure is learned from the existing data, so it can be used if the classification is not known beforehand [51]. The process is that first the system must receive the required input, then the output is obtained using algorithms. Thus it is effective in finding classes or patterns. In this technique, the model is not checked. It is also classified into two different techniques: clustering and association [20], [48]. In clustering technique, different groups are created by putting similar unlabeled data into one group [56]. In other words, the clustering technique takes input data and, after processing, divides the data into different clusters according to common features [16]. In order to find some relationship in a huge dataset or database, the association technique helps to connect or correlate data elements [39], [4]).

Summary Evaluation Parameters

Performance evaluation plays a significant role in evaluating the effectiveness of the summary created. There are gold standard criteria for evaluating English summary performance. The Data Understanding Conference (DUC) has introduced metrics such as precision, recall, F1 score, similarity score, and Recall-Oriented Understudy for Gisting Evaluation (ROUGE).

- Precision: Precision is the ratio between correctly predicted positive predictions and the sum of all predicted positive observations.

1. Precision = $\frac{TP}{TP+FP}$

- Recall: Recall is the ratio between correctly predicted positive predictions and all observations in its class.

2. Recall = $\frac{TP}{TP+FN}$

- F1 score: The F1 score is the weighted average of precision and recall. Therefore, this score takes into account false negatives as well as false positives. F1 score is more useful if we have an uneven class distribution.

3. F1 score = $\frac{2*(Recall*Precision)}{Recall+Precision}$

The similarity score is used to compare how relevant the automated summary is to human-generated summaries.

ROUGE- ROUGE stands for Recall Oriented Understudy for Gisting Evaluation. A number of criteria are used to evaluate machine translation and automatic text summarization. It evaluates automatically generated summaries or translations by comparing them with a set of summary references [32], [36], [40].

C. Literature Review

Various automatic text summarization systems are widely used for natural languages, such as English. In the case of the Turkish language, yet automatic text summarization systems do not exist.

Altan [7] developed a system that utilized a single Turkish document as input and scored using features such as sentence position information and term frequency information and obtained summaries using a series of statistical methods. The content of all articles has been converted to HTML documents to ensure formal structures. The educational system includes a collection of 50 different articles on economic topics. A number of high-frequency words were searched in the document, and the heading was determined in the form of an HTML module. After all titles were tagged, they were recorded for use in the linguistic analysis stage. First, the estimated number of sentences was obtained by searching for punctuation marks. After separating the paragraphs and sentences, the title phrases were examined, and the positive or negative sentences were analyzed. Morphological analysis developed in Prolog has been tested for word extraction, however, the need to read at least 250-300 words from the pre-built database for each document made the parsing process too long and did not yield good results.

Kutlu et al. [32] proposed a general text summarization method through sentence ordering using surface features in Turkish texts. The system calculated sentence scores based on surface features and generated summaries by extracting top-ranked sentences from the original documents. This method used information retrieval techniques such as term frequency and natural language processing techniques such as keyword and centrality. In addition, the system uses features such as title similarity and sentence position. Sentence ranking is calculated using a score function using feature values and feature weights, and machine learning techniques are used to determine the optimal combination of feature coefficients. In this study, precision (P) and recall (R) are used to evaluate its performance by selecting internal evaluation. To ensure the quality of the machine-generated summary, an internal assessment compares the summaries generated by the machine with those generated by humans. In this study, to evaluate the performance, summarization outputs using the ROUGE evaluation technique were compared with manual summaries created by 25 independent human raters. When the average density ratio was obtained for the corpus test set, the recall values were 0.54 and the precision values were 0.809 with the ROUGE evaluation method.

Özsoy et al. [46] proposed two inferential Turkish text summarization algorithms based on new latent semantic analysis in their study. VT matrix was used to select sentences. First, the mean sentence score was calculated for each concept represented by a row of the VT matrix. If the value of a cell in that row is less than the average score calculated for that row, the cell score is set to zero. The main idea is that there may be sentences that are somehow related to the topic, rather than the main sentences that represent the topic. Two different sets of scientific articles in Turkish were used to evaluate the summary approach. Articles

were selected from different fields of medicine, sociology, and psychology, and there were fifty articles in each collection. Algorithms were evaluated on Turkish documents, and their performance was compared using the ROUGE-L score. One of the algorithms generates the best scores. The authors claimed that the crossover method designed in this study is better than other hidden semantic analysis methods. According to the results, the best results of all algorithms were obtained when the input matrix was created using the root type of words.

Pembe [47] proposed an automatic document summarization system for search engines based on information requests and text structure. In this system, in the first step, each document is processed structurally to reveal the headings and subheadings related to its hierarchy. Then, an approach was evaluated using machine learning based on tree representation and support vector machines and perceptron algorithms, developed with respect to title performance and hierarchy extraction operations. In the second step, the document structures that emerged in the first step were evaluated by sentence-by-sentence scoring and section-by-section scoring to develop automatic summarization methods. As a result, better results were found compared to Google summaries and information request summaries. The second dataset (Turkish dataset) contains Turkish web documents collected from Google results using TREC-like queries defined for Turkish. This has helped evaluate both structured processing and summarization methods using larger document datasets for machine learning algorithms. After scoring the sentences, points were given according to their importance, and sentence selection was done. Title extraction results for the Turkish dataset were obtained as 0.79 0.57 0.65 for Recall, Precision and F-score respectively. To test whether the methods generally work on Turkish documents, an accuracy of 71% was achieved by analyzing documents on a Turkish university website.

Güran [25] proposed a new weight value for inferential text summarization that can be used in text summarization methods based on hidden semantic analysis, and it was shown to increase the performance of all methods on four different datasets. In this study, a hybrid system with two different approaches was also proposed, which provides a combination of semantic and structural features to extract important sentences. Experimental results have shown that it achieves better success than using each feature individually. First, 130 news documents from different news sites were collected using the fuzzy hierarchical analysis process, which includes pairwise comparison of features, and summary documents of each were prepared by three people. The second approach consisted of 20 news documents with shorter texts and summarizing documents by 30 people, using a real-time, binary-coded genetic algorithm that allowed automatic determination of feature weights. F score was used to evaluate success. Experimental results have shown that combining features and using all features has better success than using each feature individually. According to the author, it has been found that the proposed hybrid system based on fuzzy analytical hierarchy process gives positive results in the Turkish dataset. According to the results, the genetic algorithm used in the second approach has a good effect on the dataset. Using a weighted value structure derived from decimal numbers, such as a fuzzy hierarchical analysis process, has produced better results than binary-coded genetic algorithms. Using real coded chromosome structures instead of binary

coded genetic algorithms has yielded higher results. It was also found that the "distributional feature of words" used for the first time was ranked higher on short documents. The F-score values are 0.552 for the BAHS-based hybrid system; 0.650 for GK-EVSD; 0.631 for İK-EVSD; 0.566 for GK-BHÇGD and 0.560 for İK-BHÇGD.

Güran et al. [26] developed a sentence scoring function based on Fuzzy Analytical Hierarchy Process (FAHP) based on genetic algorithm and used it for automatic extractive summarization based on 15 different sentence selection methods (sentence location, distribution features, similarity to main sentences, selection of sentences based on similarity to the first and last sentences and also the title sentence, term frequency information, sentence length, term frequency, word sentence score, average Tf-Idf, thematic features, numerical data, punctuation marks, positive keywords, noun phrases, semantic features, LSA-based features, centrality). Sentences were sorted based on their score function values, and summary documents were created by extracting sentences with the highest scores. In this study, Zemberek software was used. Two different Turkish datasets were used to observe the performance of the proposed system. The first set (Turkish130) contains 130 documents related to different fields and a human-derived extractive summary created with a summarization rate of 35%. The second set (Turkish 20) contains 20 documents and 30 extractive summaries prepared by 30 different evaluators (15 men, 15 women). The purpose of using Turkish20 is to demonstrate the stability of the result of the FAHP-based system. The proposed method was compared with a heuristic algorithm, the Genetic Algorithm (GA), and F-score was used as a performance measure. For the FAHP and GA algorithms, the results of 0.562, 0.565, 0.552, and 0.556 were obtained in the Turkish30 dataset and 0.552 and 0.556 in the Turkish20 dataset, respectively. According to the authors, the FAHP-based system produced better results than the genetic algorithm-based system.

Baydar [11] has been aimed in his study to increase the success of extractive summarization method in Turkish texts. Three different methods have been tried: namely general interval random scoring, intuitively determined random score interval evaluation and fixed score evaluation using genetic algorithm. "Dataset 2" included 20 news documents and was used as the dataset in this research. 30 people were asked to choose sentences that could summarize these documents. 35% of the sentence numbers of news texts were selected for inclusion in the summary text. In the first stage, software called Zemberek was used to find the root of the words, and after the intuitive scoring of the words, the scores of the sentences were calculated and summarized. Title, high frequency, introduction, conclusion, keywords, proper names, positive words, negative words, number, double quotation mark, end sign, and date are 12 features that are considered in words. In this research, to find the average length, the average of each sentence was calculated according to the number of words, and this number became the cut-off point. In addition, an accurate measurement value was used to test the success of the system. In the intuitive fixed integer scores, random integer values, and general interval random method, the success averages were 42.25%, 51.166%, and 59.25%, respectively.

Aysu [10] used automatic text summarization on Turkish news texts. To test the performance of this study, 20 different people were asked to choose 3

sentences that they think are important from among 50 different news texts. Then the results obtained from individuals were compared with the results obtained from the study. The results of the study show a summarization performance of approximately 0.36. In the study, a total of 100,000 news items were obtained for the dataset through the API (Application Programming Interface) provided by Hürriyet Newspaper. The desired data is converted into sentences and paragraphs using JSON tags. To perform the automatic summarization process, first the possible sentences were selected, and to find the roots of the words in these sentences, the morphological analysis method was used for each word. Then the closest roots of the words were evaluated together, and a word-sentence matrix was obtained, which in the next step produced summary sentences by reduction to this matrix. Thus, the evaluation showed that the similarity function measured among the experimenters had the highest similarity (41.5 %). It was found that the similarity rates between the summaries obtained from the program and the summaries obtained from the test device had the second highest similarity rate (36.5%). The results showed that the lowest similarity in the summaries was related to the level of difference between the randomly generated summaries and the summaries of the experimental devices (11%).

Özkan [45] in her dissertation tried to reach a sound conclusion by comparing Turkish words letter by letter based on matching letters and the position of these letters in the word. A total of 1,005 news items published between March and April 2018 on the “f5haber.com” website and the “hurriyet.com.tr” news portal were used for the dataset. After creating the dataset, sentence punctuation was normalized, word frequency index was created, sentence scoring was created, and three sentences were selected by humans and added to the database to create a reference summary. They were compared them with summaries obtained by a human to evaluate the results. Mathematical Method Rouge-N Measurement Results were obtained as 0.5078 for recall, 0.3951 for precision, and 0.4442 for F1 Score. In addition, Mathematical Method Rouge-N Full Score results of 23.66 for recall and 23 for precision were obtained.

Erhandi [18] developed a model in his study on Text Summarization using Deep Learning by employing a deep autoencoder structure and utilizing LSTM structures in the hidden layers. In this study, text summarization was performed using deep learning, and the summary text was obtained using the Keras library using the Tensorflow infrastructure. The system was run with 5, 20, 100, and 250 epoch values and nearly 5000 samples.

Afatsun [3] used over 50,000 texts and summaries from the Deutsche Welle news site with the help of a Python program to collect a Turkish news data (THV) set. In this study, a bidirectional LSTM model trained using attention-layered word embeddings is developed for abstractive text summarization. The performance of the model was evaluated separately using both the words in THV and the Wikipedia trained word vectors. According to the ROUGE-1 metric, the performance score of the model in THV is 40.90.

Karaca [30] selected suitable news texts from the Packaged Water Producers Association (SuDer) as a dataset in his study on "automatically generating headlines for Turkish news texts using the deep learning method" and used the Keras library as a library and the abstract summarization method with the

transformer architecture for training. It was found that the model is able to produce headlines with the ability to express the context in news texts, which reach 75% and 85% accuracy after 20 and 25 retraining, respectively. After examining 700,000 news texts from the SuDer news collection as a dataset, data containing approximately 50,000 news texts and headlines were obtained. The evaluation was done using BLEU and ROUGE metrics. The ROUGE metric calculates three values. In this study, the results obtained for 20 periods were 0.59, 0.54, and 0.55 for the ROUGE-1, BLEU, and f1 score, respectively, and 0.77, 0.70, and 0.73 for 25 epoch.

Çelik [44] used 421 scientific articles from Turkish library journals and Information World (https://bd.org.tr/index.php/bd) as a dataset in her study, and summaries were obtained by automated and structured methods. In the study, after the root-finding process using Zemberek, the sentences were given weights according to their word frequencies. According to this study, the automatic structural summary results provided better results than the classical system summary outputs. It is found that the calculated readability values of all outputs are significantly more readable than the readability values of author abstracts. The n-gram overlaps of the automatic structural abstracts and automatic classical abstracts with the original abstracts produced by the author were calculated using the ROUGE 2.0 package and the ROUGE-1, ROUGE-2, ROUGE-L, and ROUGE-SU4 metrics.

Ertam and Aydin [19] presented a deep learning approach for summarizing Turkish text in their study. The proposed approach consists of two steps. In the first stage, Turkish news content was collected from a news agency, and in the second stage, a Turkish text summarization system was developed using the collected data. In the study, the sequence-to-sequence (Seq2Seq) model was used as an approach. The goal is to use deep learning structures such as RNN or LSTM to work with a particular token and try to predict the next set of states from the previous set. In this study, news content and headline data were used to extract news headlines as news summaries using deep learning. Study performance values are presented by averaging the results for each sentence as well as for 50 randomly selected sentences. The F1 score values are 0.4317, 0.2194, and 0.4334 for Rouge-1, Rouge-2, and Rouge-L, respectively. According to the authors, the obtained results show that the proposed approach is effective in Turkish text summarization studies.

D. Comparison Between Techniques

Table 1 includes fourteen studies that show different techniques used in summarizing Turkish texts. The table shows the processing steps, characteristics, approach, dataset, measurement evaluation, result, and future work of each summarization technique.

Table 1. Different Text Summarization Techniques in Turkish

Reference	Processing steps	Features	Approach	Dataset	Measurement Evaluation	Result	Future Work
Kutlu et al. [32]	Sorting valuable sentences from original documents and extracting high-scoring sentences	Term frequency, title similarity, keywords, sentence position, sentence centrality	Natural Language Processing (NLP)	Milliyet, Hürriyet newspapers	Recall and precision with Rouge Non-Rouge recall and precision	0,54 0,809 0,324 0,354	Changing the keyword scoring function and the number of keywords, viewing keyword impact, and using Latent Semantic Analysis (LSA)
Özsoy et al. [46]	Selection of sentence with vector matrix	Sentence selection	LSA, fuzzy hierarchical analysis process, binary coding, genetic algorithm	Scientific article from the fields of medicine, sociology, and psychology	ROUGE-L F- score	Ds1 cross 0,320 Ds2 cross 0,263	Development and evaluation of proposed approaches in English texts. Using ideas used in other methods, such as graph-based approaches, with proposed approaches to improve summarization performance.
Pembe [47]	In the first step, each document is revealed with headings and subheadings related to its hierarchy, and then it is evaluated according to the success of the heading and hierarchy extraction processes. In the second stage, the	Sentence-based scoring and section-based scoring	Machine learning algorithms, support vector machines	On the documents on the Bogazici University website (50 documents in the boun.edu.tr domain)	Recall, precision, f-score	0.79 0.57 0.65	There may be two directions for future work: structural processing and summary extraction.

	document structures that emerged in the first stage were evaluated with sentence-based scoring and section-based scoring for the development of automatic summarization methods.						
Güran [25]	Transforming into term-sentence matrices, Sentence clustering Determining the importance of sentences	Distributional features of words	LSA	news sites	F-score for the first three datasets For the last dataset, ROUGE is based on the number of Ngrams.	For BET: 0,5048 VS1 0,552 VS2	contributing to other studies
Güran et al. [26]	Finding roots, Determining and scoring sentence features, summarizing	Sentence scoring	Fuzzy Analytical Hierarchy Process, Genetic algorithm	The first set (Turkish130) contains 130 documents related to different fields and a human-derived extractive summary set created with a summarization rate of 35%. The second set (Turkish20) contains 20 documents and 30 extractive summaries prepared by 30	F-score	For Turkish30 FAHP 0.562 Genetic algorithm 0565 For Turkish20 FAHP 0.552 and Genetic algorithm 0.556	Increasing the number of features and identifying the most useful ones can be extended by adding datasets to other languages.

				different evaluators (15 men, 15 women).			
Baydar [11]	Finding word roots, calculating sentence scores, and summarizing	finding word roots	Genetic algorithm	"Dataset 2" used in the study of Güran [25]	heuristic fixed integer scores, general interval random scoring	42,25 51.166 59.25	Using fuzzy logic methods to measure success
Aysu [10]	Selecting sentences, finding root words with the morphological analysis method, creating a word-sentence matrix, and obtaining summary sentences	Finding the roots of words and vocabulary using the morphological analysis method	Natural Language Processing (NLP)	Hürriyet newspaper	To test the performance of this study, 20 different people were asked to choose 3 sentences that they think are important from among 50 different news texts. Then the results obtained from individuals were compared with the results obtained from the study.	The results of the study show the summarization performance to be approximately 0.36.	obtaining a summary sentence from a combination of related sentences using proximity matrices, creating an algorithm that scans all text documents in the Hadoop cluster and finds each other's related documents and performs their automatic summarization (multi-document summarization)
Erhandi [18]	The embedding mechanism consists of finding a cell that represents a group of cells; this cell is fed to be embedded in the encoding, and then three outputs are obtained.	The system takes news articles as a dataset and summarizes them as single-sentence headlines.	Structure of deep autoencoder using deep learning, LSTM	Turkish and English datasets			Increasing the success rate with more datasets and the number of periodic tests, contribute to other studies

Karaca [30]	Tokenization is done and then given to the model and the summary is obtained.		Deep learning	Packaged Water Producers Association (SuDer)	F1 score Recall Precision	For 20 terms: 0.55 0.54 0.59 For 25 terms: 0.73 0.70 0.77	For context-independent tasks, a pre-trained corpus dictionary and language model trained with tokenizer and Bidirectional Encoder Representation from Transformer (BERT) suitable for the morphological structure of Turkish or with Wikipedia data by Google (Google BERT) can be used. More successful results can be achieved by using a tree structure.
Özkan [45]	sentence punctuation normalization was done, Word Frequency List was created, Sentence Scoring was done	Word Frequency List was created, Sentence Scoring was done	mathematical method	From "f5haber.com", "hürriyet.com.tr"	Recall, precision, F1 score	0,5078, 0,3951, 0,4442	It can be adapted to social media posts, forums, and site user comments, and data mining can be done to establish similarity relationships with this method.
Afatsun [3]	punctuation marks were removed, A file has been added to each line, and each file is like a paper.	Word Vector Extraction,	Experiments were conducted by developing a bidirectional LSTM model.	More than 50,000 texts and summaries from the Deutsche Welle news site were used.	Rouge-1	40,91	It is planned to work on developing models that can yield higher results in longer texts.
Çelik [44]	After the root-finding process, word frequency	Sentence position,		Turkish Librarianship	Rouge 1 Rouge 2	Rouge-1, 0.34250	Full text evaluation,

	determination, and sentence selection process	centrality, Inclusion of Noble Name Words		and Information World	Rouge L Rouge Su4	Rouge-2, 0.11493 Rouge-L, 0.06014 Rouge-su4, 0.15392 Rouge-, 0.31659 Rouge-2, 0.09967 Rouge-L, 0.05561 Rouge-SU4, 0.13733	work in librarianship, and other related fields
Ertam and Aydin [19]	Scan news titles, short news, news content, and keywords of the last 5 years.	Word Embedding	sequence to sequence	BeautifulSoup Library and Scrapy framework	Rouge-1 Rouge-2 Rouge-L	F1 0.4317 0.2194 0.4334 P 0.4973 0.2619 0.5034 R 0.3968 0.1998 0.3964	production of better and longer summaries of large documents in Turkish

The data in Table 1 are summarized in the following figures. Figure 1 shows the distribution of whether the studies were sentence-based or word-based. Figure 2 shows which metric is used more. Figure 3 shows which method is used more.

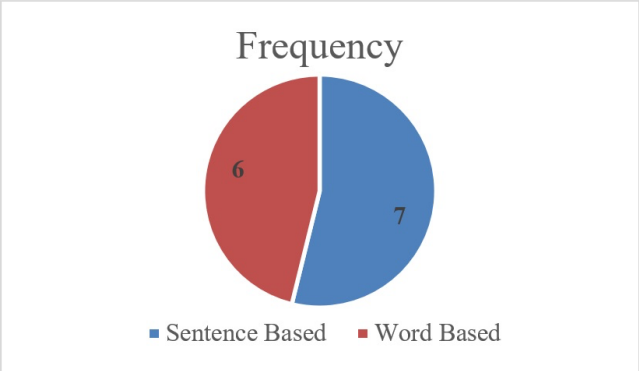

Figure 1. Frequency of Features

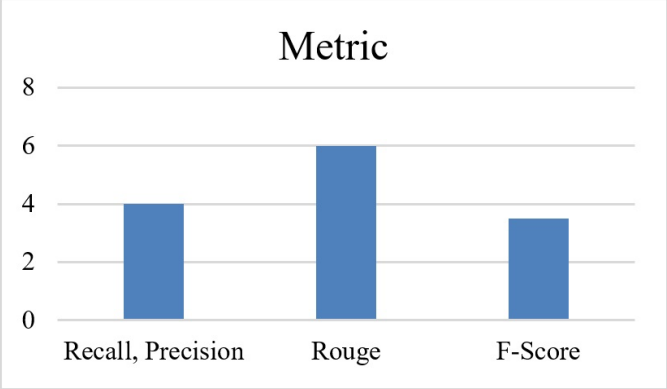

Figure 2. Frequency of Used Metrics

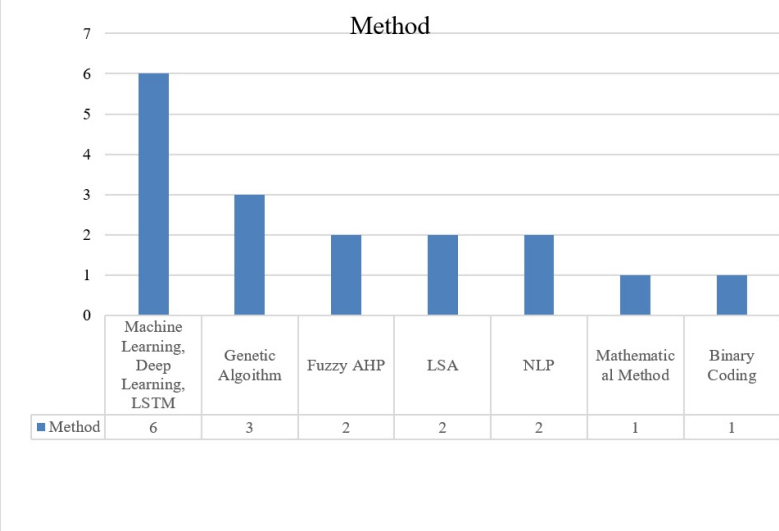

Figure 3. Frequency of Used Method

E. Conclusion and Future Research

Researchers rely on effective summaries of large text documents. The purpose of summarizing is to quickly review each topic and reach the closest and most meaningful summary. Text summarization is a demanded application for users to get the gist of the information in a short time. Text summarization for English began in the Document Understanding Conference (DUC) around 2001. However, in Turkey, text summarization research has been slow due to the lack of appropriate tools and resources and because Turkish is an agglutinative and difficult language. This study reviews the literature on text summarization techniques developed for the Turkish language. For this purpose, various parameters are used, such as processing steps, features, approach, dataset, and measurement evaluation. This study provides an idea to close the research gaps in the body of articles involved in the development of text summarization for Turkish. Several challenges and issues are highlighted for future studies in this area. 1) Development of resources including datasets, stop word lists, etc. for the Turkish language. 2) Developing multi-document summarization methods in Turkish and eliminating unnecessary sentences, creating coherence, and ordering summary sentences. 3) Identifying quality keywords for better summarization.

Dealing with bugs is important for developers on platforms because bug reports require the right response, so they need to pay attention to why these types of reports are generated and what decisions need to be made to deal with the bug to provide the optimal result. Therefore, an appropriate bug report summary should be available to resolve the issue. In this study, each technique was evaluated using a review of previous studies. Various metrics such as precision, recall, F-score and ROUGE score were calculated. Moreover, various summarization techniques were investigated in the Turkish language, which can be of significant assistance to developers in how to deal with domains. This study can be effective by clarifying the summarization techniques in revealing the selection of the appropriate method in summarization. In the future, different combinations of techniques can be studied to obtain better results.

F. References

[1] Abdel-Salam, S., & Rafea, A. (2022). Performance study on extractive text summarization using BERT models. *Information*, *13*(2), 67.

[2] Abdi, A., Shamsuddin, S. M., Hasan, S., & Piran, J. (2018). Machine learning-based multi-documents sentiment-oriented summarization using linguistic treatment. *Expert Systems with Applications*, *109*, 66-85.

[3] Afatsun, M.N. (2020). *Abstractive summarization from Turkish texts using deep learning methods* [Master's dissertation, Ankara University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.

[4] Agaskar, V., Babariya, M., Chandran, S., & Giri, N. (2017). Unsupervised learning for credit card fraud detection. *International Research Journal of Engineering and Technology (IRJET)*, *4*(3), 2343-2346.

[5] Alami, N., Meknassi, M., & Rais, N. (2015). Automatic texts summarization: Current state of the art. *Journal of Asian Scientific Research*, *5*(1), 1-15.

[6] Alimi, O. A., Ouahada, K., Abu-Mahfouz, A. M., Rimer, S., & Alimi, K. O. A. (2021). A review of research works on supervised learning algorithms for SCADA intrusion detection and classification. *Sustainability*, *13*(17), 9597.

[7] Altan, Z. (2004). February. A Turkish automatic text summarization system. *In IASTED International Conference on AIA*, 16-18.

[8] Aries, A., & Hidouci, W. K. (2019). Automatic text summarization: What has been done and what has to be done. *arXiv preprint arXiv:1904.00688*.

[9] Aswani, S., Choudhary, K., Shetty, S., & Nur, N. (2024). Automatic text summarization of scientific articles using transformers—A brief review. *Journal of Autonomous Intelligence*, *7*(5).

[10] Aysu, E. (2018). *Big data storage and automated text summarization in Turkish text* [Master's dissertation, Işık University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.

[11] Baydar, E. (2018). *Single document extractive automatic text summarization using genetic algorithms* [Master's dissertation, Van Yüzüncü Yıl University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.

[12] Chang, Z., Du, Z., Zhang, F., Huang, F., Chen, J., Li, W., & Guo, Z. (2020). Landslide susceptibility prediction based on remote sensing images and GIS: Comparisons of supervised and unsupervised machine learning models. *Remote Sensing*, *12*(3), 502.

[13] Cheung, J. C. (2008). *Comparing abstractive and extractive summarization of evaluative text* [Master's dissertation, University of British Columbia]. Semantic Scholar. https://www.semanticscholar.org/.

[14] Collingwood, L., & Wilkerson, J. (2012). Tradeoffs in accuracy and efficiency in supervised learning methods. *Journal of Information Technology & Politics*, *9*(3), 298-318.

[15] Deshpande, A. R., & Lobo, L. M. R. J. (2013). Text summarization using clustering technique. *International Journal of Engineering Trends and Technology*, *4*(8), 3348-3351.

[16] Donalek, C. (2011, April). Supervised and unsupervised learning. *In Astronomy Colloquia*. USA, 27.

[17] El-Kassas, W. S., Salama, C. R., Rafea, A. A., & Mohamed, H. K. (2021). Automatic text summarization: A comprehensive survey. *Expert systems with applications*, *165*, 113679.

[18] Erhandi, B. (2020). *Text summarization with deep learning* [Master's dissertation, Sakarya University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.

[19] Ertam, F., & Aydin, G. (2022). Abstractive text summarization using deep learning with a new Turkish summarization benchmark dataset. *Concurrency and Computation: Practice and Experience*, 34(9), e6482

[20] Fan, C., Xiao, F., Li, Z., & Wang, J. (2018). Unsupervised data analytics in mining big building operational data for energy efficiency enhancement: A review. *Energy and Buildings*, *159*, 296-308.

[21] Ferreira, R., de Souza Cabral, L., Lins, R. D., e Silva, G. P., Freitas, F., Cavalcanti, G. D., ... & Favaro, L. (2013). Assessing sentence scoring techniques for

extractive text summarization. *Expert systems with applications*, *40*(14), 5755-5764.

[22] Gambhir, M., & Gupta, V. (2017). Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review*, *47*(1), 1-66.

[23] Giannakopoulos, G., El-Haj, M., Favre, B., Litvak, M., Steinberger, J., & Varma, V. (2011). TAC 2011 MultiLing pilot overview.

[24] Gupta, V., & Lehal, G. S. (2010). A survey of text summarization extractive techniques. *Journal of emerging technologies in web intelligence*, 2(3), 258-268.

[25] Güran, A. (2013). *Automatic tex summarization system* [Doctoral dissertation, Yıldız Technical University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.

[26] Güran, A., Uysal, M., Ekinci, Y., & Güran, C. B. (2017). An additive FAHP based sentence score function for text summarization. *Information Technology and Control*, 46(1), 53-69.

[27] Han, M., Wu, H., Chen, Z., Li, M., & Zhang, X. (2023). A survey of multi-label classification based on supervised and semi-supervised learning. *International Journal of Machine Learning and Cybernetics*, *14*(3), 697-724.

[28] Haque, M., Pervin, S., & Begum, Z. (2013). Literature review of automatic multiple documents text summarization. *International Journal of Innovation and Applied Studies*, *3*(1), 121-129.

[29] Kanapala, A., Pal, S., & Pamula, R. (2019). Text summarization from legal documents: a survey. *Artificial Intelligence Review*, *51*, 371-402.

[30] Karaca, A., (2021). *Generating Turkish news headlines with deep learning* [Master's dissertation, Trakya University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.

[31] Kumar, Y., Kaur, K., & Kaur, S. (2021). Study of automatic text summarization approaches in different languages. *Artificial Intelligence Review*, *54*(8), 5897-5929.

[32] Kutlu, M., Cığır, C., & Cicekli, I. (2010). Generic text summarization for Turkish. *The Computer Journal*, 53(8), 1315-1323.

[33] Li, N., Shepperd, M., & Guo, Y. (2020). A systematic review of unsupervised learning techniques for software defect prediction. *Information and Software Technology*, *122*, 106287.

[34] Lu, L., Liu, Y., Xu, W., Li, H., & Sun, G. (2023). From task to evaluation: an automatic text summarization review. *Artificial Intelligence Review*, *56*(Suppl 2), 2477-2507.

[35] Lyaqini, S., Quafafou, M., Nachaoui, M., & Chakib, A. (2020). Supervised learning as an inverse problem based on non-smooth loss function. *Knowledge and Information Systems*, *62*(8), 3039-3058.

[36]Maybury, M. (1999). Advances in automatic text summarization. *MIT press*.

[37] Mehta, P., Bukov, M., Wang, C. H., Day, A. G., Richardson, C., Fisher, C. K., & Schwab, D. J. (2019). A high-bias, low-variance introduction to machine learning for physicists. *Physics reports*, *810*, 1-124.

[38] Mridha, M. F., Lima, A. A., Nur, K., Das, S. C., Hasan, M., & Kabir, M. M. (2021). A survey of automatic text summarization: Progress, process and challenges. *IEEE Access*, *9*, 156043-156070.
[39] Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An Unsupervised Machine Learning Algorithms: Comprehensive Review. *Int. J. Comput. Digit. Syst*.
[40] Nasar, Z., Jaffry, S. W., & Malik, M. K. (2019). Textual keyword extraction and summarization: State-of-the-art. *Information Processing & Management*, *56*(6), 102088.
[41] Nazari, N., & Mahdavi, M. A. (2019). A survey on automatic text summarization. *Journal of AI and Data Mining*, *7*(1), 121-135.
[42] Neto, J. L., Freitas, A. A., & Kaestner, C. A. (2002). Automatic text summarization using a machine learning approach. In *Advances in Artificial Intelligence: 16th Brazilian Symposium on Artificial Intelligence, SBIA 2002 Porto de Galinhas/Recife, Brazil, November 11–14, 2002 Proceedings 16* (pp. 205-215). Springer Berlin Heidelberg.
[43] Nolan, S. P., Pezzè, L., & Smerzi, A. (2021). Frequentist parameter estimation with supervised learning. *AVS Quantum Science*, *3*(3).
[44] Özkan Çelik, A. E. (2021). *Structured abstract extraction system for Turkish academic publications* [Doctoral dissertation, Hacettepe University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.
[45] Özkan, C. (2019). *Automatic summary technique for internet based Turkish texts* [Master's dissertation, Maltepe University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.
[46] Özsoy, M., Cicekli, I., & Alpaslan, F. (2010, August). Text summarization of turkish texts using latent semantic analysis. *In Proceedings of the 23rd international conference on computational linguistics Coling 2010*, pp. 869-876.
[47] Pembe, F. C. (2010). Automated query-biased and structure-preserving document summarization for web search tasks [Doctoral dissertation, Boğaziçi University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/.
[48] Prakash, J., Wang, V., Quinn III, R. E., & Mitchell, C. S. (2021). Unsupervised machine learning to identify separable clinical Alzheimer's disease sub-populations. *Brain Sciences*, *11*(8), 977.
[49] Pruneski, J. A., Pareek, A., Kunze, K. N., Martin, R. K., Karlsson, J., Oeding, J. F., ... & Williams III, R. J. (2023). Supervised machine learning and associated algorithms: applications in orthopedic surgery. *Knee Surgery, Sports Traumatology, Arthroscopy*, *31*(4), 1196-1202.
[50] Rodrigues, J. A., Martins, A., Mendes, M., Farinha, J. T., Mateus, R. J., & Cardoso, A. J. M. (2022). Automatic Risk Assessment for an Industrial Asset Using Unsupervised and Supervised Learning. *Energies*, *15*(24), 9387.
[51] Rodriguez-Nieva, J. F., & Scheurer, M. S. (2019). Identifying topological order through unsupervised machine learning. *Nature Physics*, *15*(8), 790-795.
[52] Saziyabegum, S., & Sajja, P. S. (2017). Review on text summarization evaluation methods. *Indian J. Comput. Sci. Eng*, *8*(4), 497500.

[53] Sethi, J. K., & Mittal, M. (2019). Ambient air quality estimation using supervised learning techniques. *EAI Endorsed Transactions on Scalable Information Systems*, *6*(22), e8-e8.
[54] Shafiq, M., Thakre, K., Krishna, K. R., Robert, N. J., Kuruppath, A., & Kumar, D. (2023). Continuous quality control evaluation during manufacturing using supervised learning algorithm for Industry 4.0. *The International Journal of Advanced Manufacturing Technology*, 1-10.
[55] Sharma, S., Singh, G., & Sharma, M. (2021). A comprehensive review and analysis of supervised-learning and soft computing techniques for stress diagnosis in humans. *Computers in Biology and Medicine*, *134*, 104450.
[56] Song, W., Choi, L. C., Park, S. C., & Ding, X. F. (2011). Fuzzy evolutionary optimization modeling and its applications to unsupervised categorization and extractive summarization. *Expert Systems with Applications*, *38*(8), 9112-9121.
[57] Sri, S. H. B., & Dutta, S. R. (2021, October). A Survey on Automatic Text Summarization Techniques*. In Journal of Physics: Conference Series Vol. 2040*, No. 1, p. 012044. IOP Publishing.
[58] Verma, P., & Verma, A. (2020). A review on text summarization techniques. *Journal of scientific research*, *64*(1), 251-257.
[59] Widyassari, A. P., Rustad, S., Shidik, G. F., Noersasongko, E., Syukur, A., & Affandy, A. (2022). Review of automatic text summarization techniques & methods. *Journal of King Saud University-Computer and Information Sciences*, *34*(4), 1029-1046.
[60] Yao, J. G., Wan, X., & Xiao, J. (2017). Recent advances in document summarization. *Knowledge and Information Systems*, *53*, 297-336.
[61] Yeh, J. Y., Ke, H. R., Yang, W. P., & Meng, I. H. (2005). Text summarization using a trainable summarizer and latent semantic analysis. *Information processing & management*, *41*(1), 75-95.
[62] Yogan, J. K., Goh, O. S., Halizah, B., Ngo, H. C., & Puspalata, C. (2016). A review on automatic text summarization approaches. *Journal of Computer Science*, *12*(4), 178-190.