

Article

A Novel Lexicon-Based Approach for Sentiment Analysis in Turkish

Harun Aksaya ^{1,*}  and Sevinç Gülseçen ² ¹ Department of Information Technology, Marmara University, 34100 Istanbul, Turkey² Department of Computer Science, Istanbul University, 34100 Istanbul, Turkey; gulsecen@istanbul.edu.tr

* Correspondence: harunaksaya@marmara.edu.tr; Tel.: +90-530-281-37-38

Abstract

This study investigates a target-based sentiment analysis approach on Turkish texts and examines how lexicon-based methods vary depending on language compatibility and translation strategies. The main objective is to accurately identify target-oriented expressions and to compare the performance of different sentiment lexicons within this context. For this purpose, Turkish user reviews obtained from the Turkish school review and evaluation platform were analysed using three lexicon configurations: SentiWordNet applied in its original English form with target-related term translation (SentiWordNet-EN), its fully Turkish-translated version (SentiWordNet-TR), and a native Turkish resource (SentiTurkNet). SentiTurkNet achieved the highest weighted average F1-score of 0.887 (positive-class F1: 0.926; negative-class F1: 0.760), followed by SentiWordNet-EN with a weighted average F1-score of 0.856 (positive-class F1: 0.898; negative-class F1: 0.720), and SentiWordNet-TR with a weighted average F1-score of 0.824 (positive-class F1: 0.868; negative-class F1: 0.679). One of the most significant findings is that using SentiWordNet in its original English form yields better results than the fully translated version, suggesting that the translation process leads to sentiment loss due to the incomplete preservation of sentiment intensity and contextual meaning. These findings carry important implications for sentiment analysis in low-resource languages: where comprehensive native lexicons are unavailable, translating only target-related terms into a language with richer sentiment resources can be more effective than directly translating the entire lexicon. Therefore, it is concluded that in target-based sentiment analysis, not only language compatibility but also the chosen translation strategy plays a critical role.

Keywords: aspect-based sentiment analysis; lexicon-based methods; target-based sentiment analysis; SentiTurkNet; Turkish natural language processing; cross-lingual sentiment analysis; sentiment lexicon



Academic Editors: Hiram Calvo and Pablo Gervás

Received: 14 May 2026

Revised: 20 June 2026

Accepted: 23 June 2026

Published: 2 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

With the widespread adoption of the Internet and social media platforms, users increasingly share their opinions about products, services, institutions, and a wide range of topics in online environments. This trend has led to the emergence of large volumes of user-generated textual data, and analyzing such data has become an important research topic in natural language processing and data mining [1]. Sentiment analysis studies, which aim to automatically identify users' opinions and emotions, have attracted substantial attention in both academic and industrial research in recent years [2]. Sentiment analysis is a text-mining approach that seeks to determine whether the opinions expressed in a text

are positive, negative, or neutral. In the literature, the methods used for sentiment analysis are generally examined under two main categories: machine-learning-based approaches and lexicon-based approaches [3].

In machine-learning approaches, sentiment classification is performed using models trained on previously labeled datasets. These approaches are commonly divided into supervised and unsupervised learning methods. Within supervised learning, various algorithms are employed, such as decision trees, linear classifiers, and rule-based classifiers. In recent years, the use of deep learning methods—a subfield of machine learning—has significantly improved performance in sentiment analysis tasks [3]. However, machine-learning-based approaches require large amounts of labeled data for model training, and preparing such datasets is time-consuming and costly [4].

An alternative to machine learning is the lexicon-based approach. In this approach, the sentiment polarities of words appearing in a text are determined using precompiled sentiment lexicons [5]. In lexicon-based methods, each word is associated with positive, negative, or neutral sentiment values in the lexicon, and the overall sentiment of a text is computed by aggregating these word-level scores. These methods do not require a labeled dataset or a pre-trained model [6]. Nevertheless, their performance depends strongly on the coverage and accuracy of the sentiment lexicon used. Therefore, the structure and scope of sentiment lexicons play a critical role in lexicon-based sentiment analysis studies [7].

Identifying semantic relationships among words is another important research area in natural language processing. One of the most influential lexical resources developed for this purpose is WordNet [8]. WordNet is a lexical database for English developed by the Cognitive Science Laboratory at Princeton University [9]. Its distinctive feature is that it organizes words into sets of synonyms and explicitly encodes semantic relations among them. The database contains hundreds of thousands of lexical items and systematically presents semantic relations between words [8]. One notable effort to develop multilingual lexical resources based on the WordNet structure is the BalkaNet project. Within BalkaNet, WordNet databases were created for several Balkan languages, including Greek, Bulgarian, Serbian, Romanian, Czech, and Turkish [10]. The Turkish WordNet developed in this project is an important lexical resource that captures semantic relations among Turkish words [11].

Domain-specific lexicons used in sentiment analysis are referred to as sentiment lexicons. Sentiment lexicons provide numerical sentiment values (e.g., positive, negative, and neutral/objective) associated with words [12]. They are typically constructed by compiling lists of sentiment-related terms and are widely used in sentiment analysis studies [13]. Lexicons that include positive, negative, and neutral scores for each word play a key role in determining sentiment in texts [5]. The literature includes several sentiment lexicons developed for English. One prominent resource is SentiWordNet, which is built upon WordNet and provides positive and negative sentiment scores for each entry, enabling the estimation of a word's sentiment polarity [14]. Another widely used resource is SenticNet, developed by the MIT Media Lab; it is a concept-based sentiment lexicon that contains a large-scale repository of words and concepts [15]. For Turkish, one of the best-known sentiment lexicons is SentiTurkNet [16]. SentiTurkNet was developed by researchers at Sabancı University by leveraging resources such as WordNet, SentiWordNet, and SenticNet. It provides positive, negative, and objective values for each term, which can be used to determine sentiment polarity [17]. Another sentiment lexicon developed for Turkish is HisNet. HisNet was built on the KeNet lexicon and introduced to the literature as a broad-coverage sentiment lexicon [18].

Sentiment analysis studies are also categorized according to the level of analysis [6]. In the literature, sentiment analysis is typically conducted at three levels: document level,

sentence level, and target (aspect) level [19]. In document-level studies, an entire text is assigned a single sentiment polarity, whereas sentence-level studies analyze each sentence separately. However, a single sentence may contain multiple sentiment expressions. To obtain more fine-grained results in such cases, target-based sentiment analysis methods are employed [20]. In target-based sentiment analysis, targets mentioned in the text are first identified, and then sentiment is inferred based on the words that describe these targets [3]. A substantial portion of sentiment lexicons used in the literature has been developed for English. To adapt these lexicons to other languages, direct translation methods are commonly used [21]. However, translating a sentiment lexicon from its source language into another language may lead to losses in sentiment intensity and contextual meaning. This issue can reduce the accuracy of lexicon-based sentiment analysis methods [22].

To mitigate this problem, this study proposes an alternative approach. In conventional methods, English sentiment lexicons are directly translated into the target language and then used. In contrast, this study uses SentiWordNet without translating the lexicon into Turkish; instead, only the related words that describe the target in the text are translated into English, and sentiment scores are computed accordingly. In this way, the approach aims to prevent sentiment information loss that may occur when translating the entire lexicon. Experimental results indicate that the proposed method outperforms conventional lexicon-translation approaches and yields an improvement of approximately 3 percentage points in weighted average F1-score.

Accordingly, the main contributions of this study can be summarized as follows: (i) an alternative lexicon-based sentiment analysis approach is proposed that relies on translating target-related words into English rather than directly translating the entire SentiWordNet lexicon into Turkish; (ii) the potential sentiment loss problem arising during cross-lingual lexicon translation is highlighted, and a method is developed to mitigate this issue; and (iii) the effectiveness of the proposed approach is evaluated through experiments, demonstrating an approximately 3-percentage-point improvement in weighted average F1-score compared to conventional methods.

2. Related Work

2.1. Sentiment Analysis and Lexicon-Based Methods

Sentiment analysis—the computational study of opinions, attitudes, and emotions expressed in natural language—has attracted sustained research interest since the foundational surveys of Pang and Lee [23] and Liu [13], which established core terminology, problem formulations, and evaluation frameworks. Lexicon-based methods estimate text polarity by aggregating the sentiment scores of constituent words against a predefined lexicon [5]. A practical advantage of this paradigm is that it requires neither labeled training data nor a pre-trained model, making it well suited to low-resource scenarios [13]. Taboada et al. [5] provided a systematic comparative evaluation of lexicon-based methods, demonstrating their competitiveness with supervised approaches on product reviews while identifying negation handling and domain coverage as persistent weaknesses. Hu and Liu [12] demonstrated the applicability of opinion lexicons to product feature extraction and sentiment association, laying important groundwork for aspect-level analysis.

2.2. Aspect-Based Sentiment Analysis

Aspect-based sentiment analysis (ABSA) extends document- and sentence-level approaches by identifying the sentiment expressed toward specific entities or attributes within a text [20]. The SemEval shared task series—spanning 2014 [24], 2015 [25], and 2016 [26]—standardised ABSA evaluation across restaurant and laptop domains, providing benchmark datasets that continue to guide the field. Early neural systems applied Long

Short-Term Memory networks [27] to capture sequential dependencies; Tang et al. [28] proposed target-dependent LSTMs that encode the queried target directly into the hidden representation, while Wang et al. [29] introduced attention mechanisms that allow the model to focus selectively on context words most relevant to the aspect. The gated recurrent unit architecture of Cho et al. [30] provided a complementary recurrent formulation widely adopted in sequence-to-sequence sentiment tasks. In applied educational contexts, Altrabsheh et al. [31] developed a sentiment analysis framework for classroom settings, and Wen et al. [32] examined sentiment signals in MOOC discussion forums—a motivation aligned with the educational platform examined in the present study.

2.3. Turkish NLP Resources and Sentiment Analysis

Turkish presents distinct challenges for NLP owing to its agglutinative morphology, whereby a single root may generate dozens of surface forms through suffix concatenation [33]. The ITU Turkish NLP Web Service [34] addressed this by providing an integrated pipeline covering tokenisation, normalisation, morphological tagging, named entity recognition, and dependency parsing. On the lexical resource side, Senti-TurkNet [17]—constructed by mapping Turkish WordNet synsets to SentiWordNet polarity scores via a semi-supervised classifier—provides positive, negative, and objective scores for 14,795 entries and remains the most widely used native sentiment lexicon for Turkish. HisNet [18], built upon the KeNet lexical database, extends coverage further. Among empirical studies, Kaya et al. [35] applied n-gram features and support vector machines to Turkish political news, reporting classification accuracies between 65% and 77%. More recently, Salur and Aydin [36] introduced an annotated ABSA corpus for the Turkish smart tourism domain, and Mutlu and Ozgur [37] released a targeted sentiment analysis dataset for Turkish and showed that fine-tuned transformer models substantially outperform lexicon-based baselines.

2.4. Cross-Lingual Sentiment Analysis

The scarcity of sentiment resources for non-English languages has motivated extensive work on cross-lingual transfer strategies. A common approach is to translate existing English lexicons—such as SentiWordNet [14]—directly into the target language; however, Demirtas and Pechenizkiy [38] demonstrated that machine-translated lexicons suffer from semantic drift and the loss of culturally grounded sentiment intensity. An alternative is to translate the source text into English rather than the lexicon, thereby preserving the original scoring environment at the cost of content fidelity. The present study operationalises precisely this distinction, comparing a native Turkish lexicon, a machine-translated version of SentiWordNet, and an approach in which only the target-adjacent terms are translated into English for scoring—a design that directly addresses the cross-lingual sentiment loss problem.

2.5. Deep Learning and Transformer-Based Approaches

The introduction of BERT [39]—a deeply bidirectional transformer pre-trained via masked language modelling on large English corpora—marked a watershed moment for NLP. Sun et al. [40] adapted BERT for ABSA by reformulating aspect-level classification as a sentence-pair task, constructing auxiliary sentences that encode the queried aspect; this approach yielded state-of-the-art results on the SemEval ABSA benchmarks. Multilingual BERT (mBERT), distributed alongside the original model [39], extends these capabilities to 104 languages through joint pre-training on multilingual Wikipedia, enabling zero-shot cross-lingual transfer without language-specific fine-tuning data. For Turkish specifically, Schweter [41] released BERTurk, a monolingual BERT variant pre-trained on a large Turkish corpus; Mutlu and Ozgur [37] subsequently confirmed that BERTurk-based

models outperform multilingual variants on Turkish targeted sentiment tasks. XLM-R, proposed by Conneau et al. [42], scales RoBERTa-style pre-training to 100 languages using 2.5 terabytes of filtered Common Crawl text, achieving substantial gains over mBERT on cross-lingual benchmarks. Despite these advances, transformer-based methods entail significant computational costs and require domain-specific labelled data for fine-tuning, rendering lexicon-based and hybrid approaches practically relevant—particularly for morphologically rich, resource-constrained languages.

2.6. Hybrid and Comparative Perspectives

Hybrid approaches combining lexicon-based scoring with statistical or neural classification have been explored as a means of leveraging lexical interpretability alongside representational capacity [3,5]. In the Turkish context, the limited availability of annotated corpora restricts the generalisation of purely data-driven models, and class imbalance further complicates supervised training. By systematically comparing three lexicon configurations—a native Turkish resource, a machine-translated English lexicon, and a cross-lingual scoring strategy—on a unified Turkish ABSA dataset, the present study contributes a transparent, reproducible reference point for future hybrid approaches that seek to integrate lexical priors into neural architectures for low-resource, morphologically complex languages.

3. Materials and Methods

3.1. Dataset Construction

The dataset used in this study was obtained from www.okul.com.tr (Istanbul, Turkey), an online platform where users share their opinions about educational institutions. The www.okul.com.tr platform provides information on private and public schools at the preschool, primary, middle, and high school levels. The website includes school images, educational facilities, contact information, and user comments. Users can visit the platform to obtain information about schools and share their own experiences in the form of reviews.

To enable the use of these user reviews for sentiment analysis, the platform administrators were contacted and the required permissions were obtained. Within the scope of this permission, a total of 1000 anonymized user reviews written about primary-level schools in Istanbul were collected. The collected reviews were de-identified to ensure that they contained no personal data and were transferred to a database created for this study.

Table 1 presents a statistical profile of the dataset. The 1000 collected reviews yielded 1929 aspect-level targets across 228 distinct aspect categories, of which 1925 are retained for evaluation after excluding four neutral-labelled instances. Review texts vary between 2 and 37 words (mean = 13.6, SD = 6.4, median = 12); approximately 65% of reviews fall within the 6–15-word range.

Each review in the corpus may express sentiment toward more than one school attribute; consequently, the 1000 collected reviews yield a total of 1929 annotated aspect-level targets. Four of these targets received a neutral (objective) label by annotator consensus and were excluded from the quantitative evaluation, leaving 1925 targets for analysis. The distribution of aspect counts per review is summarised in Table 1: 777 reviews (62.3%) contain a single aspect-level target, while the remaining 460 reviews (37.7%) contain two or more, with a maximum of eight targets observed in a single review.

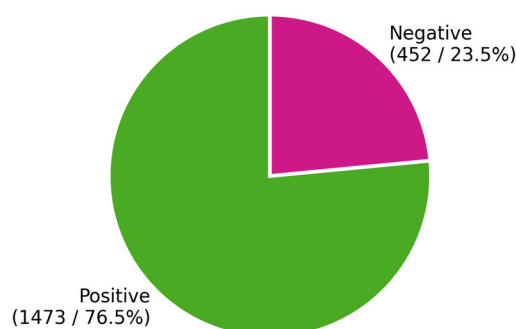
Of the 1929 annotated targets, four (0.21%) received a neutral label by annotator consensus and were excluded from the quantitative evaluation. The remaining 1925 targets comprise 1473 positive (76.5%) and 452 negative (23.5%) instances, as shown in Figure 1a. With respect to inclusion and exclusion criteria, only reviews concerning primary-level schools in Istanbul were retained. Texts consisting solely of numerical ratings with no

accompanying prose, and entries containing fewer than two meaningful tokens after normalization, were excluded. All collected texts were de-identified in accordance with the data-sharing agreement with the platform operator prior to annotation. These steps ensure that the corpus comprises exclusively content-rich, evaluative texts suitable for aspect-level sentiment analysis. Figure 1b presents the distribution of review lengths, which follows a right-skewed pattern consistent with informal user-generated content.

Table 1. Statistical Profile of the Dataset Used in This Study.

Property	Value
Total reviews collected	1000
Total aspect-level targets	1929 (1925 after excluding 4 neutral)
Distinct aspect categories	228
Avg. aspects per review	1.56 (range: 1–8)
Reviews with 1 aspect	777 (62.3%)
Reviews with 2+ aspects	460 (37.7%)
Positive targets	1473 (76.5%)
Negative targets	452 (23.5%)
Avg. words per review (mean)	13.6 (SD = 6.4, median = 12)
Min/Max words per review	2/37
Avg. characters per review	101.0
Avg. sentences per review	1.9 (range: 1–7)
Avg. related words per target	6.6 (range: 2–23)

(a) Sentiment Class Distribution



(b) Review Length Distribution

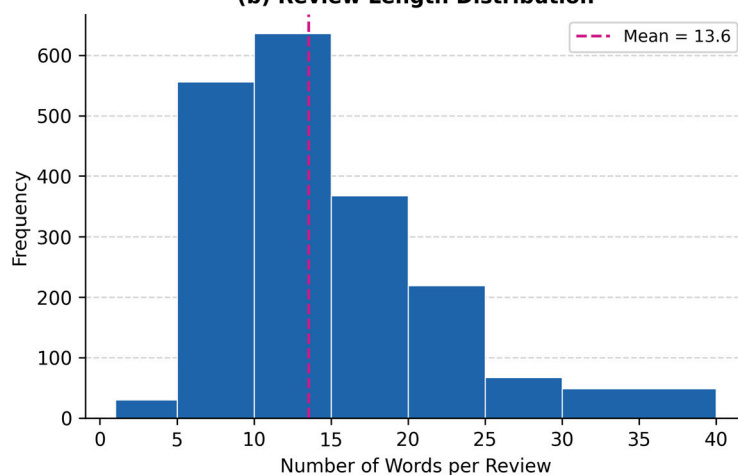


Figure 1. Dataset Overview: (a) Sentiment Class Distribution; (b) Review Length Distribution (words per review).

Figure 2 illustrates the distribution of the top 15 aspect categories alongside their sentiment breakdown. The most frequently annotated aspect is Okul (School, $n = 320$), followed by Eğitim (Education, $n = 176$) and Öğretmen (Teacher, $n = 171$). Notably, aspects such as Ulaşım (Transport), Servis (Shuttle), and Yönetim (Management) exhibit predominantly negative sentiment, whereas Hizmet (Service), Yemek (Food), and Sınıf (Classroom) are predominantly positive—a distribution consistent with realistic patterns in school evaluation discourse.

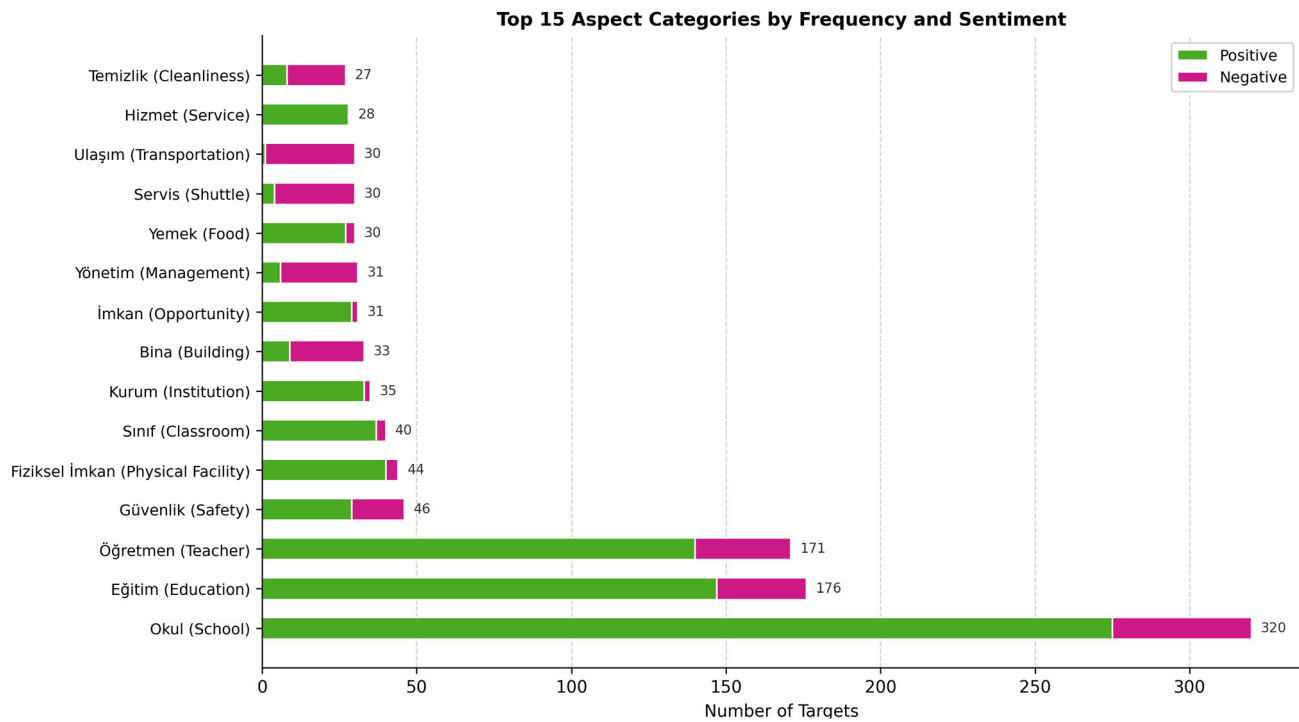


Figure 2. Distribution of the Top 15 Aspect Categories with Positive and Negative Sentiment Breakdown.

Figure 3a shows that 62.3% of reviews contain a single aspect, while 37.7% contain two or more. Figure 3b confirms that the related-word extraction step produces a sufficient number of context words per target (mean = 6.6, range 2–23) to support reliable lexicon-based scoring.

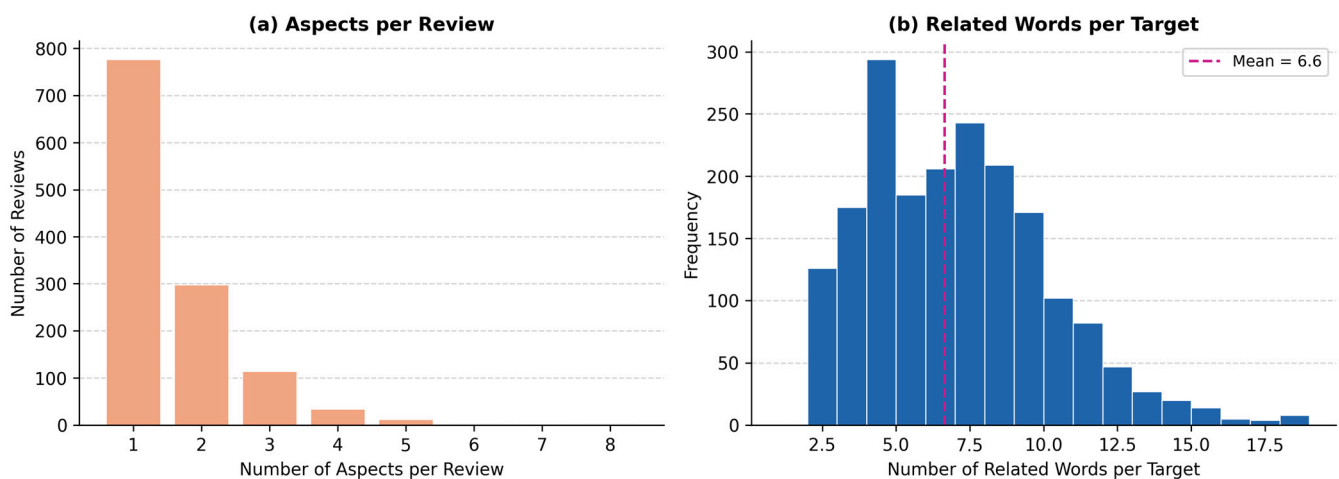


Figure 3. Structural Characteristics of the Dataset: (a) number of aspects per review; (b) distribution of related words per target.

3.2. Data Labeling Process

To evaluate the reviews under a target-based sentiment analysis setting, the review texts were annotated by domain experts. For this purpose, three independent experts in the field of education were consulted. Each expert reviewed the texts independently and identified the targets appearing in the reviews. In addition, for each target, the sentiment was labeled as positive, negative, or neutral. The annotations from the three experts were compared, and in cases of disagreement, a consensus decision was reached to determine the

final target and sentiment labels. As a result, a dataset suitable for target-based sentiment analysis was produced.

During the annotation process, a small number of targets ($n = 4$) received a neutral (objective) label by consensus—that is, all three annotators agreed that the text expressed no discernible positive or negative sentiment toward the target. As these neutral targets represent a negligible proportion of the dataset ($n = 4$ out of 1929; 0.21%), they were excluded from the performance evaluation. The final evaluation dataset therefore consists exclusively of targets carrying an explicit positive or negative sentiment label, yielding the 1473 positive and 452 negative targets reported in Table 1.

3.3. Natural Language Processing of the Dataset

The collected review dataset was processed using the ITU Turkish Natural Language Processing Web Service (ITU NLP) [34]. In this study, the Whole Pipeline class of ITU NLP (Istanbul Technical University, Istanbul, Turkey) was utilized. This pipeline includes the following NLP modules:

- Tokenizer;
- Normalizer;
- Morphological Tagger;
- Named Entity Recognizer;
- Dependency Parser;
- UD Mapper.

First, a text is passed through the Tokenizer module and segmented into tokens. Next, the Normalizer module is applied to correct spelling errors and character variations. The normalization process includes operations such as case conversion, lexicon-based corrections, proper-name detection, accent normalization, and spelling correction [43]. In the subsequent step, the Morphological Tagger is used to identify each word's stem, part of speech, and affixes. This step is particularly important for agglutinative languages such as Turkish, where accurate stem extraction is critical [33]. The Named Entity Recognizer module is employed to detect entities such as persons, organizations, locations, dates, and monetary expressions [44]. Finally, the Dependency Parser is used to determine syntactic relations among words. This is especially important in target-based sentiment analysis because it supports identifying the relationship between a target and the words that modify it [45].

3.4. Development for Dataset Processing

A dedicated software tool was developed to process the 1000 user reviews collected from www.okul.com.tr using the ITU NLP service. The tool submits review texts to the ITU NLP service and enables automated processing. The outputs returned by the service are then transferred into a MySQL database. This workflow makes it possible to store all NLP results obtained from the reviews in a database and reuse them for subsequent analyses.

Figure 4 shows the interface of the software developed to test the ITU NLP service. A text entered in the review input field is submitted to the service by clicking the “Send” button, and the processed output is returned in an array format.

Figure 5 presents an example output produced by the ITU NLP service for a processed review text. The output includes, for each word, its stemmed form, part of speech, the type of affix it takes, and the line number of the word it is syntactically linked to.

Figure 4. Interface of the Software Developed to Test the ITU NLP Web Service.

```

1 Array
2 (
3   [0] => Array
4     (
5       [0] => 1
6       [1] => Okulumuz (our school)
7       [2] => okul (school)
8       [3] => Noun
9       [4] => Noun
10      [5] => A3sg|Pl|Nom
11      [6] => 13
12      [7] => SUBJECT
13    )
14  [1] => Array
15    (
16      [0] => 2
17      [1] => ,
18      [2] => ,
19      [3] => Punc
20      [4] => Punc
21      [5] => 
22      [6] => 1
23      [7] => PUNCTUATION
24    )
25  [2] => Array
26    (
27      [0] => 3
28      [1] => harika (wonderful)
29      [2] => harika (wonderful)
30      [3] => Adj
31      [4] => Adj
32      [5] => 
33      [6] => 5
34      [7] => MODIFIER
35    )
36  [3] => Array
37    (
38      [0] => 4
39      [1] => bir (a / one)
40      [2] => bir (a / one)
41      [3] => Adj
42      [4] => Num
43      [5] => 
44      [6] => 5
45      [7] => DETERMINER
46    )
47  [4] => Array
48    (
49      [0] => 5
50      [1] => eğitim (education)
51      [2] => eğitim (education)
52      [3] => Noun
53      [4] => Noun
54      [5] => A3sg|Pnon|Nom
55      [6] => 13
56      [7] => OBJECT
57    )
58  )
59  )
60  )
61  )
62  )
63  )

```

Figure 5. Output of a Review Processed by the ITU NLP Web Service.

3.5. Identification of Related Words

In target-based sentiment analysis, accurately identifying the words associated with (i.e., describing) a target is essential for determining the target's sentiment. For this purpose, the Dependency Parser module provided in the ITU NLP Web Service was initially used. However, preliminary tests indicated that this module did not achieve sufficient performance in identifying targets and their related words. Therefore, an additional software component was developed to improve the accuracy of related-word extraction.

The processing steps of the software developed to identify related words are depicted as follows (Figure 6).

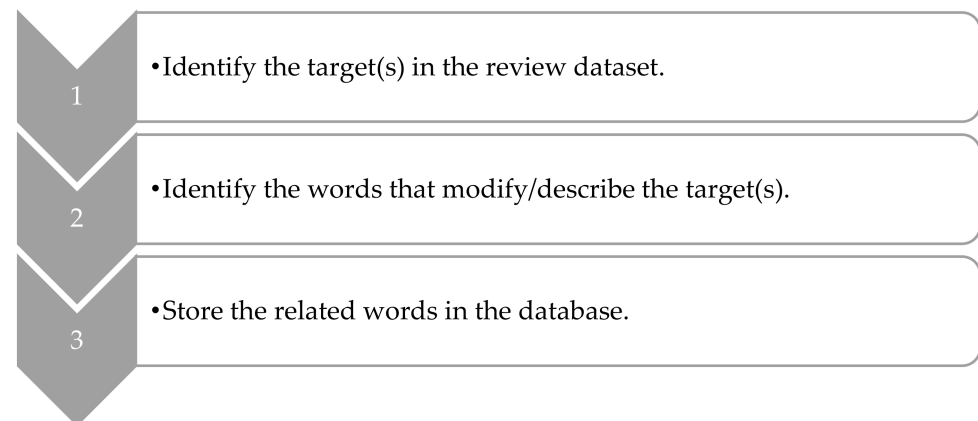


Figure 6. Steps for Identifying Related Words.

The following factors are considered when identifying related words:

1. How many times does the target occur within the sentence?
2. Is the target a single word or a multi-word expression?
3. Is the word modifying the target located in the same sentence as the target?

The script was executed on 1000 reviews, and the extracted related words were transferred to a MySQL database. Because related words are identified on a per-target basis, a new row was created in the database for each target occurrence within a sentence. As a result of the related-word extraction process, 1929 targets were obtained from 1000 reviews, since a single review may contain multiple targets. Of these, four neutral-labelled targets were subsequently excluded from evaluation, yielding the final set of 1925 targets. Using the developed software, the review associated with each target, the target sentiment, and the related words were stored in a relational format.

The following pseudocode (Algorithm 1) formally describes the related word extraction procedure. The algorithm traverses the dependency tree produced by the ITU NLP pipeline to collect, for each target occurrence, all content words (adjectives, verbs, adverbs, nouns) that directly or transitively modify that target within the same sentence boundary.

Figure 7 presents the corresponding flowchart of the algorithm. The flowchart details all branching conditions, including the distinction between single-word and multi-word aspects, the two-level dependency traversal, the sentence boundary check, and the conditions under which records are discarded. Negation markers (e.g., *değil*), intensifiers (e.g., *çok*), sarcasm, and cross-sentence modifier relations are not processed. These limitations are discussed in Section 5.

As shown in Table 2, the annotation process yielded 1929 targets in total: 1473 are labelled as positive, 452 as negative, and 4 as neutral. The four neutral instances were excluded from the quantitative evaluation, resulting in a final evaluation set of 1925 targets.

Table 2. Number of Targets by Sentiment Label.

Sentiment Label	Number of Targets
Negative	452
Positive	1473
Neutral	4
Total	1929

Algorithm 1. Related Word Extraction for Target-Based Sentiment Analysis

Input: Parsed sentence S (ITU NLP: tokens, stems, POS tags, dep. heads)

Aspect term A (single-word or multi-word)

Output: Set of stemmed related words R for aspect A in sentence S

Step 1—Preprocessing

1.1 Apply ITU NLP Whole Pipeline to S

→ tokens, stems, POS tags, dependency head indices

1.2 Normalize text (lowercase, spelling correction)

1.3 Determine: single-word or multi-word aspect A

Step 2—Locate target A in S

2.1 Find all token positions where A appears in S

(multi-word A: match consecutive token sequence)

2.2 IF A appears 0 times in S → discard record

2.3 IF A appears ≥ 1 times → proceed for each occurrence

Step 3—Collect related words (for each occurrence position p)

$R \leftarrow \emptyset$

FOR each token t_i in S:

// Depth-1: direct modifier of A

IF $\text{dep_head}(t_i) = p$

AND $\text{POS}(t_i) \in \{\text{ADJ}, \text{VERB}, \text{ADV}, \text{NOUN}\}$

AND $\text{sentence_id}(t_i) = \text{sentence_id}(A)$:

$R \leftarrow R \cup \{\text{stem}(t_i)\}$

// Depth-2: modifier of a direct modifier of A

IF $\text{dep_head}(t_i) \in \{t_j \mid \text{dep_head}(t_j) = p\}$

AND $\text{POS}(t_i) \in \{\text{ADJ}, \text{ADV}\}$

AND $\text{sentence_id}(t_i) = \text{sentence_id}(A)$:

$R \leftarrow R \cup \{\text{stem}(t_i)\}$

IF A is multi-word:

FOR each constituent word w in A:

apply Depth-1 and Depth-2 with head = position(w)

Step 4—Post-processing

4.1 Remove duplicate stems from R

4.2 POS filter (Steps 1–3) ensures only content words are retained

4.3 IF $R = \emptyset$ → discard record

Step 5—Store result

Save (A, sentence_id, R, sentiment_label) to database

3.6. Preparing Sentiment Lexicons for Use

A lexicon-based sentiment analysis approach was adopted in this study. Two sentiment lexicons were used: SentiWordNet and SentiTurkNet. SentiWordNet is an English sentiment lexicon built on the WordNet structure; for each entry, positive and negative sentiment scores are defined. SentiWordNet contains 117,659 entries [14]. SentiTurkNet, developed by researchers at Sabancı University, was constructed using resources such as WordNet, SentiWordNet, and SenticNet. For each term, it provides positive, negative, and objective sentiment values, and it contains 14,795 entries in total [17]. To use SentiWordNet in this study, a three-stage procedure was followed (Figure 8).

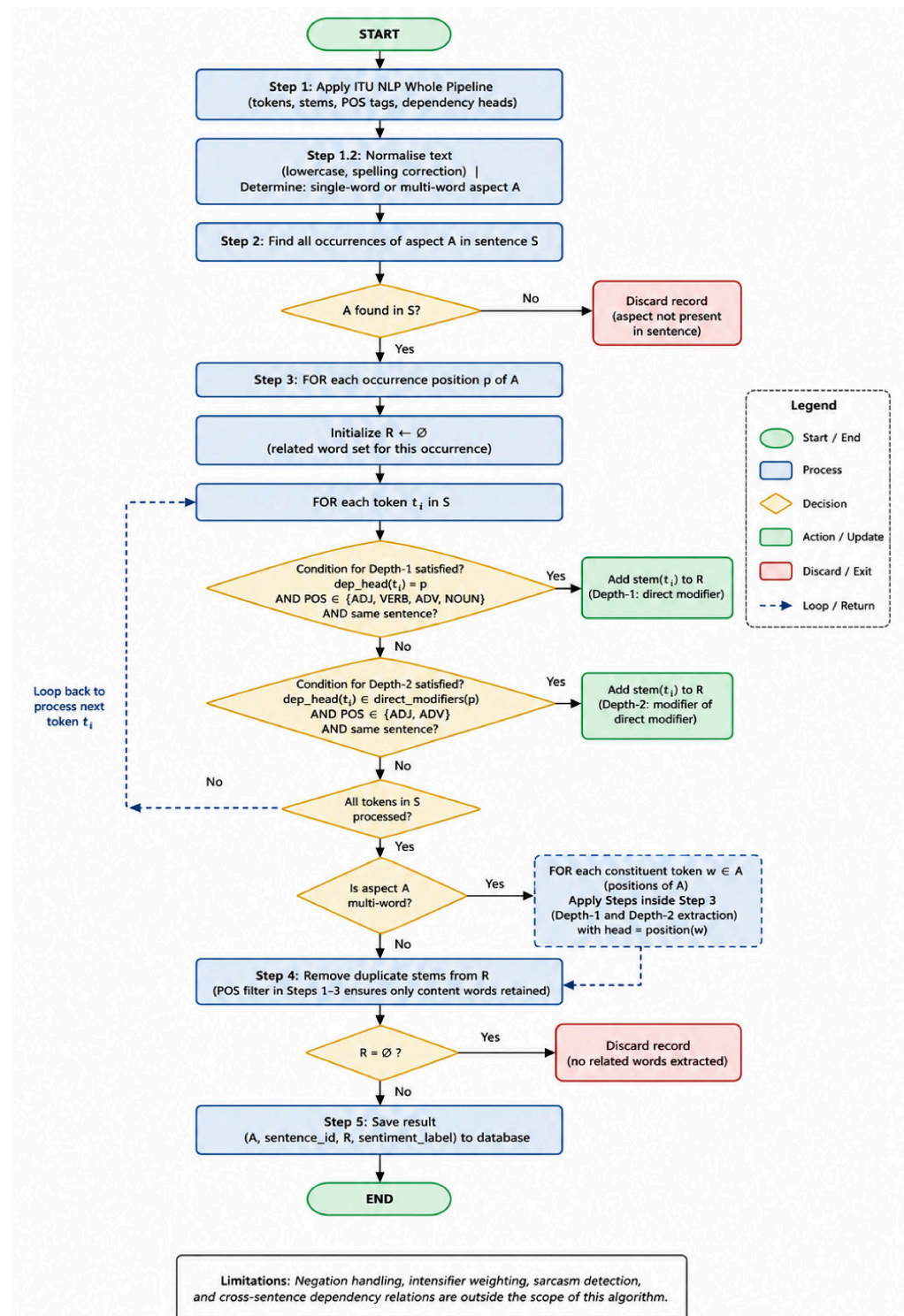


Figure 7. Flowchart of the Related Word Extraction Algorithm (Algorithm 1). Blue boxes: process steps; yellow diamonds: decision points; green boxes: output additions; red boxes: discard conditions.

To use SentiWordNet in this study, a three-stage procedure was followed (Figure 8). First, the lexicon file (in *.txt* format) was imported into a MySQL database. During the import, each line in the file was converted into SQL statements and stored in the database, enabling fast querying of words and sentiment scores. Second, a software tool was developed using Google Translate API to translate SentiWordNet into Turkish. The software was implemented in PHP, and each translated word was automatically inserted into a new database. Third, the translated entries were stored in a MySQL database.

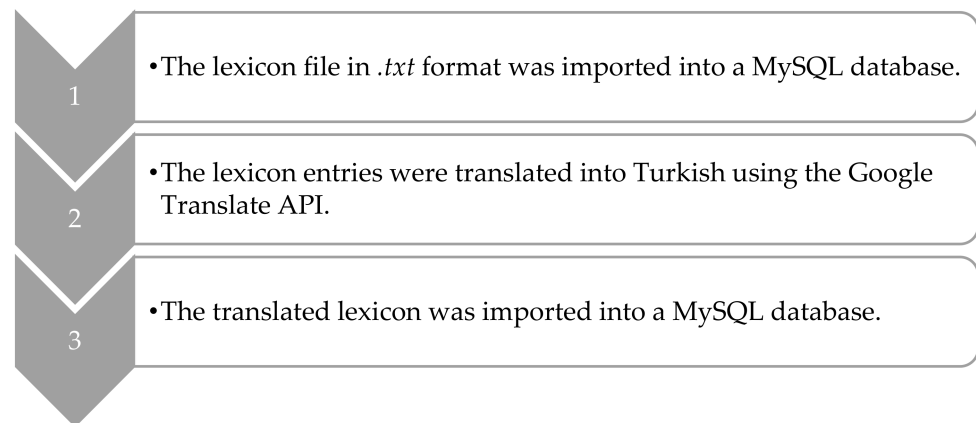


Figure 8. Preparation Steps for the SentiWordNet Lexicon.

Table 3 shows that the lexicon consists of the above columns. The Type field indicates the part of speech: “a” for adjective, “n” for noun, and “v” for verb. Ngstv_Score and Pstv_Score store the negative and positive scores, respectively. Word_Explains contains the Turkish explanation of the word.

Table 3. Data Schema of the SentiWordNet Lexicon Translated into Turkish.

Column Name	Description
Word_Id	unique identifier assigned to the word
Type	part-of-speech tag.
Ngstv_Score	negative score.
Pstv_Score	positive score.
Word_Tr	Turkish translation of the lexicon entry
Word_En	original English lexicon entry
Word_Explains	Turkish description of the word

The SentiTurkNet lexicon file was downloaded from the research group’s website and imported into a MySQL database using an SQL script. Its database schema includes 12 columns.

As shown in Table 4, the lexicon schema consists of 12 columns. Sentitürk_ID contains the unique identifier assigned to each entry. The Synonyms field stores the term itself, while TurkishGloss provides its Turkish description. PolarityLabel indicates the sentiment label, where “p” denotes positive, “n” denotes negative, and “o” denotes objective. Postag specifies the part-of-speech tag: “a” for adjective, “n” for noun, and “v” for verb. Neg_Value, Pos_Value, and Obj_Value represent the negative, positive, and objective scores, respectively. The sum of these three scores equals 1. For each term, the sentiment is assumed to be closer to the class whose score is closest to 1.

Table 4. Data Schema of the SentiTurkNet Lexicon.

Column Name	Description
Sentitürk_Id	unique identifier assigned to the term
Synonyms	term in the lexicon
TurkishGloss	Turkish description
PolarityLabel	polarity label (“p” positive, “n” negative, “o” objective)
Postag	part-of-speech tag (“a” adjective, “n” noun, “v” verb)
Neg_Value	negative score
Obj_Value	objective score
Pos_Value	positive score
Eng_Synonyms	English equivalent
EnglishGloss	English description
Swnpos	SentiWordNet positive score
Swnneg	SentiWordNet negative score

When querying the lexicon for a given word stem, a term may correspond to multiple synset entries carrying different polarity scores, a situation reflecting lexical polysemy. In such cases, the first matching entry in the database is retrieved and used for scoring; no explicit word sense disambiguation is applied.

An examination of the SentiTurkNet lexicon reveals that 308 of its 12,024 unique terms (approximately 2.6%) occur in more than one synset with distinct polarity values. The limited prevalence of such polysemous terms suggests that the practical impact of this simplification on overall classification accuracy remains restricted.

3.7. Model Development

For the lexicon-based approach, a three-stage workflow was followed:

1. Preparing the review dataset for processing;
2. Developing the model;
3. Testing with different lexicons and determining performance metrics.

The stem forms of related words were stored in the database in association with each target. ITU NLP can provide the stemmed form of each word when processing a review, and this functionality was used to obtain stemmed related words. These stem forms were then transferred into the database.

The model construction comprises five steps:

1. A review sentence is processed.
2. Targets within the review sentence are identified.
3. Stemmed word chains associated with each target are generated.
4. For each target, the corresponding word chains are sent to the lexicon for scoring.
5. The sentiment of the target is determined.

Figure 9 illustrates the workflow of the model developed for the lexicon-based approach.

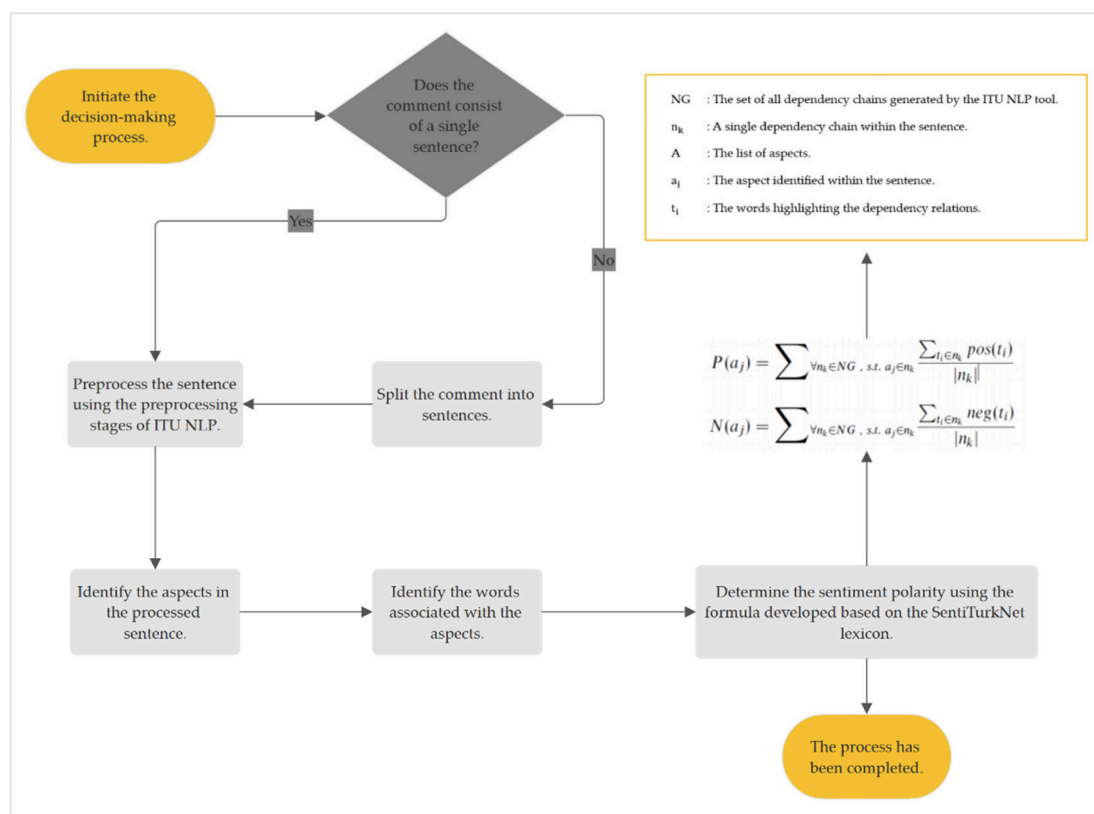


Figure 9. Model Developed for the Lexicon-Based Approach.

Figure 10 presents the sentiment classification function. To determine the positive sentiment score of each target, the following steps are applied:

1. Each related word is submitted to the sentiment lexicon.
2. The positive scores obtained from the lexicon are summed.
3. The total score is divided by the number of related words.
4. The average positive sentiment score is calculated.

$P(a_i) = \frac{\sum_{k_j \in t_y} pos(k_j)}{ t_y }$ $N(a_i) = \frac{\sum_{k_j \in t_y} neg(k_j)}{ t_y }$	$P(a_i) > N(a_i) \Rightarrow \text{Positive}$ $P(a_i) < N(a_i) \Rightarrow \text{Negative}$
---	---

Figure 10. Sentiment Classification Function [46].

To determine the negative sentiment score of a target, the following steps are applied:

1. Each related word is submitted to the sentiment lexicon.
2. The negative scores obtained from the lexicon are summed.
3. The total score is divided by the number of related words.
4. The average negative sentiment score is calculated.

If the average positive score is higher than the average negative score, the sentiment polarity of the target is classified as positive. Conversely, if the average negative score is higher than the average positive score, the target is classified as negative.

In the present implementation, all related words are assigned equal weight in the scoring process; no distance-based or word-order-based weighting is applied. Each related word is independently queried against the sentiment lexicon, and the resulting scores are summed and divided by the total number of related words to yield the average positive and negative sentiment scores for the given target.

Negation handling is not incorporated in the current model. In Turkish, negation may be expressed morphologically—through verb suffixes such as *-me/-ma*—or lexically, as with the postposition “*değil*” (not). In the present approach, related words are scored as they appear in the text, and no score inversion is applied when a negation marker is present in the surrounding context. This constitutes a known limitation of the model, the implications of which are discussed in Section 5.

3.8. Implementation Details

The system was developed and executed on a local Windows workstation. The backend was implemented in PHP 7 using a MySQL 8.0 database to store all intermediate and final outputs, including parsed NLP data, lexicon entries, related word chains, and sentiment scores. Text processing was performed via the ITU Turkish NLP Web Service [34], queried using the Whole Pipeline class with JSON-formatted responses; each review was submitted as a separate API call. The Google Translate API (v2 Basic) was used for batch translation of SentiWordNet entries into Turkish and for translating target-related terms into English in the cross-lingual scoring scenario. The SentiWordNet lexicon is distributed under a Creative Commons Attribution-ShareAlike 4.0 International licence (CC BY-SA 4.0). SentiTurkNet was released as part of an academic study [17]; no explicit open data licence is stated by the authors, and users are advised to contact the data providers prior to any commercial application. Both lexicons were used strictly within an academic research

context in the present study. The source code, lexicon loading utilities, and scoring scripts used in this study are publicly available at: <https://github.com/haksaya/turkish-lexicon-absa> (accessed on 19 June 2026). The full annotated dataset is available upon request from the corresponding author, subject to the de-identification agreement with the data provider.

4. Results

To evaluate the model developed within the scope of the lexicon-based approach, the SentiTurkNet and SentiWordNet sentiment lexicons were utilized. The following formulas were used to calculate the performance metrics of the model, including accuracy, precision, recall, and F1-score.

Figure 11a presents the formula used to calculate the precision of the lexicon-based approach. Precision measures the proportion of correctly identified positive instances out of all instances predicted as positive, reflecting the model's ability to avoid false positives.

$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$	
(a)	
$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$	
(b)	
$\text{F1-Score} = \frac{2 \times \text{True Positive}}{2 \times \text{True Positive} + \text{False Positive} + \text{False Negative}}$	
(c)	

Figure 11. (a) Formula Used for Precision Calculation. (b) Formula Used for Recall Calculation. (c) Formula Used for F1-Score Calculation.

Figure 11b presents the formula used to calculate the recall of the lexicon-based approach. Recall measures the proportion of actual positive instances that were correctly identified by the model, reflecting its ability to avoid false negatives.

Figure 11c presents the formula used to calculate the F1-Score of the lexicon-based approach. The F1-Score is the harmonic mean of precision and recall, providing a balanced evaluation metric that accounts for both false positives and false negatives.

To provide a more detailed evaluation of model performance under class imbalance, Table 5 reports precision, recall, and F1-score for each class across all three lexicons. The dataset contains 1473 positive and 452 negative targets, yielding a positive-to-negative ratio of approximately 3.3:1.

Table 5. Class-Wise Precision, Recall, and F1-Score for Each Lexicon (dataset: 1473 positive, 452 negative targets).

Lexicon	Class	Precision	Recall	F1-Score	Support
SentiTurkNet	Positive	0.9277	0.9240	0.9259	1473
SentiTurkNet	Negative	0.7555	0.7655	0.7604	452
SentiTurkNet	Weighted avg	-	-	0.8870	1925
SentiWordNet-TR	Positive	0.9434	0.8038	0.8680	1473
SentiWordNet-TR	Negative	0.5687	0.8429	0.6791	452
SentiWordNet-TR	Weighted avg	-	-	0.8237	1925
SentiWordNet-EN	Positive	0.9399	0.8595	0.8979	1473
SentiWordNet-EN	Negative	0.6419	0.8208	0.7204	452
SentiWordNet-EN	Weighted avg	-	-	0.8562	1925

Figure 12 presents the class-wise evaluation metrics for each lexicon. Across all three lexicons, the positive class consistently achieves higher scores than the negative class, reflecting the impact of class imbalance on model performance.

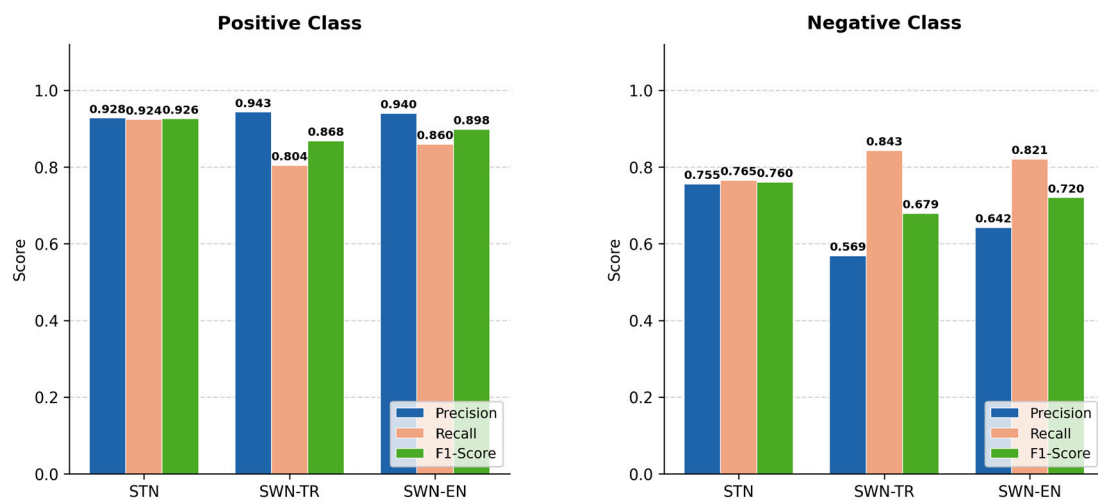


Figure 12. Class-Wise Evaluation Metrics (Precision, Recall, F1-Score) for Each Lexicon.

5. Discussion

Within the scope of the lexicon-based approach, three different sentiment lexicons were evaluated: SentiWordNet-TR, SentiWordNet (used in its original English form; hereafter SentiWordNet-EN), and SentiTurkNet. SentiWordNet-TR was obtained by translating the original SentiWordNet lexicon into Turkish and contains 73,386 Turkish word entries; the model developed using this lexicon achieved a weighted average F1-score of 0.824 (positive-class F1: 0.868; negative-class F1: 0.679). SentiWordNet-EN is a sentiment lexicon originally developed using the WordNet database and contains 117,659 English entries; the model in which only target-related terms were translated into English for scoring against this lexicon achieved a weighted average F1-score of 0.856 (positive-class F1: 0.898; negative-class F1: 0.720). SentiTurkNet is a sentiment lexicon specifically developed for the Turkish language and contains 14,795 entries; the model based on this lexicon achieved the highest overall performance, with a weighted average F1-score of 0.887 (positive-class F1: 0.926; negative-class F1: 0.760). Detailed per-class precision, recall, and F1 values for all three lexicons are reported in Table 5.

One of the most notable findings is that using SentiWordNet in its original English form—rather than translating it fully into Turkish—yields higher classification performance than the fully translated version. SentiWordNet-TR achieved a weighted average F1-score of 0.824, whereas the configuration in which only target-adjacent terms were translated into English for scoring against the original SentiWordNet produced a higher weighted average F1-score of 0.856. While this pattern suggests that sentiment loss may occur during full lexicon translation due to incomplete preservation of sentiment intensity, contextual information, and semantic nuances, it should be noted that the observed performance difference may also reflect other implementation-specific factors—including the quality of machine translation used for SentiWordNet-TR construction, the handling of lexical ambiguity during translation, and the absence of word sense disambiguation in both configurations. These factors may jointly contribute to the observed gap and warrant further investigation.

The superiority of SentiTurkNet (weighted average F1: 0.887) is largely attributable to its native design for the Turkish language, which avoids the translation losses inherent in cross-lingual lexicon adaptation. Nevertheless, the strong performance of SentiWordNet-

EN (weighted average F1: 0.856) suggests a viable alternative strategy: rather than constructing or translating large-scale lexicons, using high-quality resources in their original language and translating only the target text can be both effective and resource-efficient.

These results provide an important implication for sentiment analysis studies, particularly for languages with limited lexical resources other than well-resourced languages such as English and Spanish. In cases where high-quality native lexicons are not available, translating texts into a language with richer sentiment resources and performing analysis on that language can be more effective than constructing or translating lexicons directly. Therefore, rather than focusing solely on lexicon construction, hybrid approaches that effectively utilize existing high-quality lexicons across different languages should be considered in sentiment analysis studies.

A further limitation concerns the absence of negation handling and contextual weighting in the scoring procedure. In the current implementation, all related words receive equal weights regardless of their syntactic proximity to the target, and no score inversion is applied for negated expressions. For instance, a phrase such as “okul iyi değil” (the school is not good) would still yield a positive contribution from iyi (good), since the negation marker “değil” is not processed. Although negated constructions represent a limited proportion of the dataset, integrating negation detection—for example, by exploiting dependency relations identified by the ITU NLP pipeline—is expected to improve classification accuracy for negative targets. Distance-weighted scoring is likewise a direction for future investigation, as modifiers in closer syntactic proximity to a target may provide stronger evidence of its sentiment polarity. These limitations represent potential sources of classification error, particularly for targets expressed through negated or intensified constructions, and remain priority areas for future work.

Polysemy disambiguation represents a related limitation. When a queried term occurs in multiple lexicon entries with differing polarity scores, the current implementation retrieves the first matching database entry without applying word sense disambiguation. As noted in Section 3.6, however, such cases account for only approximately 2.6% of the unique terms in SentiTurkNet—308 out of 12,024—which constrains the overall effect on classification performance. This simplification likewise represents a potential source of classification error when polysemous terms carry context-dependent sentiment; incorporating context-aware or dependency-based disambiguation strategies in future work would substantially mitigate this issue.

As a reference point, a majority-class classifier that assigns all predictions to the positive class would achieve an accuracy of approximately 76.5% (1473 of 1925 targets). All three lexicon-based configurations exceed this trivial baseline in weighted average F1-score—0.887 for SentiTurkNet, 0.856 for SentiWordNet-EN, and 0.824 for SentiWordNet-TR—demonstrating that the lexicon-based scoring adds genuine predictive value beyond the majority classifier. Nevertheless, the class-wise evaluation results presented in Table 5 and Figure 12 reveal a consistent performance gap between the positive and negative classes across all three lexicons. SentiTurkNet achieves the highest negative-class F1-score (0.760), outperforming SentiWordNet-EN (0.720) and SentiWordNet-TR (0.679). This advantage is largely attributable to the native Turkish design of SentiTurkNet, which captures language-specific sentiment nuances more accurately than translated resources. The lower precision scores observed for the negative class suggest that the lexicon-based scoring function tends to predict positive sentiment by default when matched word scores are low or balanced, which disproportionately penalises the minority class. Future work should investigate threshold adjustment and class-weighting strategies to improve minority-class performance in imbalanced settings.

The primary objective of this study is to evaluate and compare distinct lexicon-based sentiment analysis strategies for Turkish, with particular focus on the effect of cross-lingual lexicon translation on classification performance. The experimental comparisons are therefore intentionally scoped to three lexicon-based configurations: a native Turkish lexicon (SentiTurkNet), a machine-translated English lexicon (SentiWordNet-TR), and an approach in which only target-adjacent terms are translated into English for scoring against the original SentiWordNet. A systematic comparison with supervised machine learning methods or transformer-based architectures such as BERTurk [41] or XLM-R [42], while valuable as a broader benchmark, falls outside the scope of the present study and constitutes a clear direction for future work. The lexicon-based framework examined here is especially pertinent in scenarios where labelled training data are scarce, where low computational overhead is required, or where the transparency and interpretability of the classification process are priorities.

Table 6 presents the SWOT analysis of the lexicon-based approach. The table consists of four main dimensions: strengths, weaknesses, opportunities, and threats.

Table 6. SWOT Analysis of the Lexicon-Based Approach.

Strengths		Weaknesses	
		1.	Dependence on the capability and coverage of the lexicon.
1.	Relatively simple technical process.	2.	Lower accuracy rates compared to machine learning approaches.
2.	No training process is required for the model.	3.	Insufficient number and limited coverage of Turkish sentiment lexicons compared to other languages.
3.	Domain-independent.		
Opportunities		Threats	
		1.	Due to the fact that language is a living and evolving system, lexicons may lose their relevance over time.
1.	The limited number of Turkish sentiment lexicons indicates a clear need for further research and development in this area.	2.	Risk of discontinuation of development teams, leading to the absence of new versions or updates of the lexicons.

As shown in Table 6, the strengths of the lexicon-based approach include its relatively simple technical process, the absence of a training phase for the model, and its domain-independent nature. However, dependence on the coverage and capability of the lexicon, lower accuracy rates compared to machine learning approaches, and the limited number and coverage of Turkish sentiment lexicons are considered key weaknesses of the approach.

The limited availability and insufficient capacity of Turkish sentiment lexicons indicate a clear need for further research in this area. In addition, since language is a living and evolving system, lexicons may lose their relevance over time. There is also a risk that development teams may disband, resulting in the discontinuation of updates and the absence of new versions.

6. Conclusions

This study investigated target-based sentiment analysis on Turkish educational reviews by systematically comparing three lexicon-based configurations: a native Turkish lexicon (SentiTurkNet), a machine-translated version of SentiWordNet (SentiWordNet-TR), and an approach in which only target-adjacent terms are translated into English for scoring

against the original SentiWordNet (SentiWordNet-EN). A dataset of 1000 anonymized user reviews collected from www.okul.com.tr was annotated by three domain experts, yielding 1925 aspect-level targets across 228 categories. Related words were extracted using a two-level dependency-based algorithm built on the ITU Turkish NLP pipeline, and sentiment scores were aggregated via an averaging function applied to lexicon-matched stems. SentiTurkNet achieved the highest weighted average F1-score (0.887) and the strongest negative-class F1-score (0.760), followed by SentiWordNet-EN (weighted average F1: 0.856; negative-class F1-score: 0.720) and SentiWordNet-TR (weighted average F1: 0.824; negative-class F1-score: 0.679). A key finding is that translating only target-related terms into English and scoring them against the original SentiWordNet consistently outperforms full lexicon translation into Turkish, indicating that sentiment intensity and contextual meaning are partially lost during lexicon-level translation. These results carry practical implications for sentiment analysis in low-resource languages: when a high-quality native lexicon is unavailable, translating the target text rather than the lexicon represents a more effective strategy. The lexicon-based framework requires no labelled training data, is domain-independent, and offers full interpretability—properties that are particularly valuable in educational and institutional settings. Limitations include the absence of negation and intensifier handling, a first-match word sense disambiguation strategy, class imbalance in the evaluation dataset, and a single-domain scope. Future work should incorporate dependency-based negation detection, context-aware disambiguation, and comparative evaluation against transformer-based models such as BERTurk [41] and XLM-R [42] across multiple domains.

Author Contributions: Conceptualization, H.A. and S.G.; methodology, H.A.; software, H.A.; validation, H.A. and S.G.; formal analysis, H.A.; writing—original draft preparation, H.A.; writing—review and editing, S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. The APC was not funded by any external organization.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to ongoing and future research activities and data usage restrictions.

Acknowledgments: This work is based on the Ph.D. dissertation of Harun Aksaya, titled “Turkish Natural Language Processing Using Dictionary Based Approach and Machine Learning: Sentiment Analysis in Educational Institutions,” completed in 2021 at the Institute of Graduate Studies in Science, Istanbul University, Turkey. The authors thank all the anonymous reviewers for their insightful comments.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviation

The following abbreviation is used in this manuscript:

ITU NLP Istanbul Technical University Natural Language Processing Software Pipeline

References

1. Hu, J.J.H.; Ahmad, F.; Bader-El-Den, M. Decoding brand sentiments: Leveraging customer reviews for insightful brand perception analysis using natural language processing and Tableau. *PLoS ONE* **2025**, *20*, e0334330. [[CrossRef](#)] [[PubMed](#)]
2. Cui, J.; Wang, Z.; Ho, S.B.; Cambria, E. Survey on sentiment analysis: Evolution of research methods and topics. *Artif. Intell. Rev.* **2023**, *56*, 8469–8510. [[CrossRef](#)] [[PubMed](#)]

3. Zhang, L.; Wang, S.; Liu, B. Deep learning for sentiment analysis: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1253. [[CrossRef](#)]
4. Kearney, C.; Liu, S. Textual sentiment in finance: A survey of methods and models. *Int. Rev. Financ. Anal.* **2014**, *33*, 171–185. [[CrossRef](#)]
5. Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K.; Stede, M. Lexicon-based methods for sentiment analysis. *Comput. Linguist.* **2011**, *37*, 267–307. [[CrossRef](#)]
6. Feldman, R. Techniques and applications for sentiment analysis. *Commun. ACM* **2013**, *56*, 82–89. [[CrossRef](#)]
7. Cambria, E.; Schuller, B.; Xia, Y.; Havasi, C. New avenues in opinion mining and sentiment analysis. *IEEE Intell. Syst.* **2013**, *28*, 15–21. [[CrossRef](#)]
8. Miller, G.A. WordNet: A lexical database for English. *Commun. ACM* **1995**, *38*, 39–41. [[CrossRef](#)]
9. Fellbaum, C. (Ed.) *WordNet: An Electronic Lexical Database*; MIT Press: Cambridge, MA, USA, 1998. [[CrossRef](#)]
10. Tufiş, D.; Cristea, D.; Stamou, S. BalkaNet: Aims, methods, results and perspectives. A general overview. *Rom. J. Inf. Sci. Technol.* **2004**, *7*, 9–43.
11. Bilgin, O.; Çetinoğlu, Ö.; Oflazer, K. Building a WordNet for Turkish. *Rom. J. Inf. Sci. Technol.* **2004**, *7*, 163–172.
12. Hu, M.; Liu, B. Mining and summarizing customer reviews. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 22–25 August 2004; pp. 168–177. [[CrossRef](#)]
13. Liu, B. Sentiment analysis and opinion mining. *Commun. ACM* **2012**, *55*, 76–84.
14. Baccianella, S.; Esuli, A.; Sebastiani, F. SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Proceedings of the International Conference on Language Resources and Evaluation (LREC), Valletta, Malta, 17–23 May 2010; pp. 2200–2204.
15. Cambria, E.; Speer, R.; Havasi, C.; Hussain, A. SenticNet: A publicly available semantic resource for opinion mining. In Proceedings of the AAAI Conference on Artificial Intelligence, Atlanta GA, USA, 11–15 July 2010.
16. Vural, A.G.; Cambazoglu, B.B.; Sen, S.; Varol, C.; Ulusay, M. A framework for sentiment analysis in Turkish: SentiTurkNet. In Proceedings of the ACL Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA 2013), Atlanta, GA, USA, 14 June 2013.
17. Dehkharghani, R.; Saygin, Y.; Yanikoglu, B.; Oflazer, K. SentiTurkNet: A Turkish polarity lexicon for sentiment analysis. *Lang. Resour. Eval.* **2016**, *50*, 667–685. [[CrossRef](#)]
18. Bakay, Ö.; Ergelen, Ö.; Sarıms, E.; Yıldırım, S.; Arıcan, B.N.; Kocabalcıoğlu, A.; Özçelik, M.; Saniyar, E.; Kuyrukçu, O.; Avar, B.; et al. Turkish WordNet KeNet. In Proceedings of the Global WordNet Conference (GWC), Pretoria, South Africa, 18–21 January 2021.
19. Wang, J.; Xu, B.; Zu, Y. Deep learning for aspect-based sentiment analysis. In Proceedings of the 2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Chongqing, China, 9–11 July 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 267–271. [[CrossRef](#)]
20. Liu, H.; Chatterjee, I.; Zhou, M.; Lu, X.S.; Abusorrah, A. Aspect-based sentiment analysis: A survey of deep learning methods. *IEEE Trans. Comput. Soc. Syst.* **2020**, *7*, 1358–1375. [[CrossRef](#)]
21. Mohammad, S.M.; Turney, P.D. Crowdsourcing a word–emotion association lexicon. *Comput. Intell.* **2013**, *29*, 436–465. [[CrossRef](#)]
22. Vilares, D.; Alonso, M.A.; Gómez-Rodríguez, C. On the usefulness of lexical and syntactic processing in polarity classification of Twitter messages. *J. Assoc. Inf. Sci. Technol.* **2015**, *66*, 1799–1816. [[CrossRef](#)]
23. Pang, B.; Lee, L. Opinion mining and sentiment analysis. *Found. Trends Inf. Retr.* **2008**, *2*, 1–135. [[CrossRef](#)]
24. Pontiki, M.; Galanis, D.; Pavlopoulos, J.; Papageorgiou, H.; Androutsopoulos, I.; Manandhar, S. SemEval-2014 Task 4: Aspect based sentiment analysis. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval-2014), Dublin, Ireland, 23–24 August 2014; pp. 27–35. [[CrossRef](#)]
25. Pontiki, M.; Galanis, D.; Papageorgiou, H.; Manandhar, S.; Androutsopoulos, I. SemEval-2015 Task 12: Aspect based sentiment analysis. In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval-2015), Denver, CO, USA, 4–5 June 2015; pp. 486–495. [[CrossRef](#)]
26. Pontiki, M.; Galanis, D.; Papageorgiou, H.; Androutsopoulos, I.; Manandhar, S.; Al-Smadi, M.; Al-Ayyoub, M.; Zhao, Y.; Qin, B.; De Clercq, O.; et al. SemEval-2016 Task 5: Aspect based sentiment analysis. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), San Diego, CA, USA, 16–17 June 2016; pp. 19–30. [[CrossRef](#)]
27. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
28. Tang, D.; Qin, B.; Feng, X.; Liu, T. Effective LSTMs for target-dependent sentiment classification. In Proceedings of the 26th International Conference on Computational Linguistics (COLING 2016), Osaka, Japan, 11–16 December 2016; pp. 3298–3307. Available online: <https://aclanthology.org/C16-1311> (accessed on 1 May 2026).
29. Wang, Y.; Huang, M.; Zhu, X.; Zhao, L. Attention-based LSTM for aspect-level sentiment classification. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP 2016), Austin, TX, USA, 1–4 November 2016; pp. 606–615. [[CrossRef](#)]

30. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014), Doha, Qatar, 25–29 October 2014; pp. 1724–1734. [\[CrossRef\]](#)
31. Altrabsheh, N.; Gaber, M.M.; Cocea, M. SA-E: Sentiment analysis for education. *Front. AI Appl.* **2013**, *255*, 353–362. [\[CrossRef\]](#)
32. Wen, M.; Yang, D.; Rose, C.P. Sentiment analysis in MOOC discussion forums: What does it tell us? In Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014), London, UK, 4–7 July 2014; pp. 130–137.
33. Şahin, M.; Sulubacak, U.; Eryiğit, G. Redefinition of Turkish morphology using flag diacritics. In Proceedings of the Tenth Symposium on Natural Language Processing (SNLP-2013), Phuket, Thailand, 28–30 October 2013.
34. Eryiğit, G. ITU Turkish NLP web service. In Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2014), Gothenburg, Sweden, 26–30 April 2014; pp. 1–4.
35. Kaya, M.; Fidan, G.; Toroslu, I.H. Sentiment analysis of Turkish political news. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2012), Macau, China, 4–7 December 2012. [\[CrossRef\]](#)
36. Salur, M.U.; Aydin, I. An annotated Turkish aspect based sentiment analysis corpus for smart tourism. In Proceedings of the 2021 Innovations in Intelligent Systems and Applications Conference (ASYU), Elazığ, Turkey, 6–8 October 2021. [\[CrossRef\]](#)
37. Mutlu, M.M.; Ozgur, A. A dataset and BERT-based models for targeted sentiment analysis on Turkish texts. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop (ACL SRW 2022), Dublin, Ireland, 22–27 May 2022; pp. 467–472. [\[CrossRef\]](#)
38. Demirtas, E.; Pechenizkiy, M. Cross-lingual polarity detection with machine translation. In Proceedings of the Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM 2013), Chicago, IL, USA, 11 August 2013. Available online: <https://dl.acm.org/doi/10.1145/2502069.2502072> (accessed on 22 June 2026).
39. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019), Minneapolis, MN, USA, 2–7 June 2019; pp. 4171–4186. [\[CrossRef\]](#)
40. Sun, C.; Huang, L.; Qiu, X. Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019), Minneapolis, MN, USA, 2–7 June 2019; pp. 380–385. [\[CrossRef\]](#)
41. Schweter, S. BERTurk—A BERT language model for Turkish. *Zenodo* **2020**. [\[CrossRef\]](#)
42. Conneau, A.; Khandelwal, K.; Goyal, N.; Chaudhary, V.; Wenzek, G.; Guzmán, F.; Grave, E.; Ott, M.; Zettlemoyer, L.; Stoyanov, V. Unsupervised cross-lingual representation learning at scale. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020), Online, 5–10 July 2020; pp. 8440–8451. [\[CrossRef\]](#)
43. Eryiğit, G.; Selamet, D.T. Social media text normalization for Turkish. *Nat. Lang. Eng.* **2017**, *23*, 835–875. [\[CrossRef\]](#)
44. Şeker, G.A.; Eryiğit, G. Extending a CRF-based named entity recognition model for Turkish well-formed text and user-generated content. *Semant. Web* **2017**, *8*, 625–642. [\[CrossRef\]](#)
45. Eryiğit, G.; Nivre, J.; Oflazer, K. Dependency parsing of Turkish. *Comput. Linguist.* **2008**, *34*, 357–389. [\[CrossRef\]](#)
46. Dehkharghani, R.; Yanikoglu, B.; Saygin, Y.; Oflazer, K. Sentiment analysis in Turkish at different granularity levels. *Nat. Lang. Eng.* **2017**, *23*, 535–559. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.