

**DEVELOPING A COMPREHENSIVE EMOTION LEXICON  
FOR TURKISH**

**TÜRKÇE İÇİN KAPSAMLI BİR DUYGU SÖZLÜĞÜ  
GELİŞTİRİLMESİ**

**ELİF ÜNAL**

**ASSOC. PROF. DR. BURKAY GENÇ**

**Supervisor**

Submitted to

Graduate School of Science and Engineering of Hacettepe University

as a Partial Fulfillment to the Requirements

for the Award of the Degree of Master of Science

in Computer Engineering

September 2023

## **ABSTRACT**

### **DEVELOPING A COMPREHENSIVE EMOTION LEXICON FOR TURKISH**

**Elif Ünal**

**Master of Science, Computer Engineering**

**Supervisor: Assoc. Prof. Dr. Burkay Genç**

**September 2023, 103 pages**

Social media and mood meters on websites have made expressing ideas easy and clear. These communication channels have produced useful data that can be used across disciplines. These data have helped researchers in several ways. Such investigations, also called text analysis, can yield materials from politics to psychology.

Emotion analysis, focuses on the extraction of emotions from textual data. Extraction from texts enables the conduct studies that have interdisciplinary applications. For one, emotional analysis can provide insights into the sentiments and emotions experienced by customers towards a particular product, service, or brand. This data has the potential to enhance customer satisfaction and foster customer loyalty. Emotional analysis also holds potential in fraud detection as it enables the identification of distinctive patterns of emotional language that are closely linked to fraudulent behavior.

When the recent studies on emotion analysis in Turkish are analysed, it is seen that mostly English sources are translated into Turkish. However, it is inevitable that there are structures that cannot be translated and suffixed languages such as Turkish make this structure even more complex. For this purpose, our main hypothesis in this study is that the data used in a

language analysis should be in its own language. In this way, an accurate and more reliable analysis will be possible. For this, it was decided to create a emotion analysis data containing many emotions in Turkish.

While this thesis was being prepared, studies were carried out in three main phases. In the first phase, it was aimed to create a data that can be used with emotions in Turkish by using 100 literary works. For this purpose, 213 different emotions in Turkish were searched in the sentences taken from the books and the number of times the words were used with the related emotion was recorded. Thus, a word for each emotion and an emotion-word vector containing the frequency of occurrences of the term with that emotion were created.

In the next phase, it was aimed to cluster these emotions by using clustering algorithms on the data to form groups with each other. For this, firstly, words that do not make sense in the data were removed and the data was scaled. Since the data is very large and sparse on these processes, valuable words were used by putting this data into PCA structure. Then, these data were tested on different clustering algorithms and results were obtained.

Finally, an interface was created to test the behaviour of text input. Afterwards, the distance between the input text and the emotion vectors was calculated with the cosine distance formula to extract emotions from input data. The results of the model created using texts of different lengths and with different emotions were evaluated. A cross-validation study was conducted to compare the emotion assignments from the study and ChatGPT. A consensus was reached by four individuals regarding the experience of at least one emotion among the five identified emotions, thus indicating successful recognition of the predominant emotion conveyed in the text.

**Keywords:** Turkish Emotion Analysis, Emotion Detection, Emotion Extraction, Emotion Analysis, Natural Language Processing

## ÖZET

# TÜRKÇE İÇİN KAPSAMLI BİR DUYGU SÖZLÜĞÜ GELİŞTİRİLMESİ

**Elif Ünal**

**Yüksek Lisans, Bilgisayar Mühendisliği**

**Danışman: Assoc. Prof. Dr. Burkay Genç**

**Eylül 2023, 103 sayfa**

Sosyal medya platformlarının büyümesi ve mod ölçerlerin çeşitli web sitelerine entegre edilmesi, bireylerin görüşlerini rahat ve açık bir şekilde ifade etmelerini kolaylaştırmıştır. Bu iletişim kanalları tarafından kullanılan etkileşimler, farklı çalışma alanlarında uygulanabilecek verilerin ortaya çıkmasını sağlamıştır. Bu verileri kullanan bilim dalı olan metin analizi, siyasetten psikolojiye kadar birçok bağlamda kullanılacak materyaller üretilebilmektedir.

Duygu analizi, metinsel verilerden duyguların çıkarılmasına odaklanmaktadır. Bu, müşterilerin belirli bir ürün, hizmet veya markaya yönelik duygu ve hisleri hakkında bilgi sağlayabilir. Bu veriler müşteri memnuniyetini artırma ve müşteri sadakatini teşvik etme potansiyeline sahiptir. Ayrıca, hedeflenen alıcıların duygusal durumundan yararlanarak pazarlama kampanyalarını özelleştirme potansiyeline de sahiptir. Duygu analizi, dolandırıcılık davranışıyla yakından bağlantılı olan ayırt edici duygusal dil kalıplarının tanımlanmasını sağladığından dolandırıcılık tespiti alanında potansiyel taşımaktadır.

Bir metin örneğini analiz ederken göz önünde bulundurulması gereken en önemli faktörler, dile özgü kaynakların mevcudiyeti ve yeterliliğidir. Türkçe de dahil olmak üzere İngilizce

dışındaki dillerde yeterli kaynak bulunmamaktadır. Dillerin belirli yapısal unsurları paylaşmasına rağmen, her birinin kendine özgü özellikleri, diller arasında standart bir kavramsallaştırma yolunun geliştirilmesini engellemiştir.

Türkçe duygu analizi üzerine son zamanlarda yapılan çalışmalara bakıldığında çoğunlukla İngilizce kaynakların çevrilerek kullanıldığı görülmüştür. Ancak çeviri yapılırken tam karşılık alınamayacak yapıların bulunması ve Türkçe gibi sondan eklemeli dillerin bu yapıyı daha da karmaşık hale getirecek olması kaçınılmazdır. Bu amaçla bu çalışmadaki hipotezimiz bir dille ilgili bir analiz yaparken kullanılan verinin kendi dilinde olması gerektiğidir. Bu amaçla Türkçe’de bulunan birçok duyguyu içeren bir duygu analiz verisi oluşturmaya karar verilmiştir. Bu veri, duygu analizi için bir kaynak olacaktır.

Bu tez oluşturulurken üç ana faz üzerinde işlemler yapılmıştır. İlk fazda, 100 edebi eser kullanılarak Türkçe’deki duygularla kullanılabilir bir veri oluşturmak amaçlanmıştır. Bunun için kitaplardan alınan cümleler içerisinde Türkçe’de bulunan 213 farklı duygu aranmış ve cümleler içerisindeki diğer kelimelerin ilgili duyguyla eşleştirebilmek frekansları kaydedilmiştir. Böylelikle her duyguya ait bir kelime ve kelimenin o duygu ile frekansını içeren bir duygu-kelime vektörü oluşmuştur. Bu yapı içerisinde 167 farklı duygu ve 179,854 adet kelime bulunmaktadır.

İkinci fazda, bu duyguların birbirleri ile grup oluşturması için veri üzerinde kümeleme algoritmalarına sokularak kümeleme yapılması hedeflenmiştir. Bunun için önce veri üzerinde bir anlam ifade etmeyen kelimelerin kaldırılma ve veriyi ölçeklendirme işlemleri yapılmıştır. Bu işlemlerin üzerine verinin çok büyük ve seyrek olması sebebiyle bu veri PCA yapısına sokularak değerli kelimelerin kullanılması sağlanmıştır. Daha sonra bu veri farklı makine öğrenimi algoritmaları üzerinde denenerek gruplama sonuçları elde edilmiştir.

Son olarak, anlık veri girişinde verinin davranışını test etmek için bir arayüz oluşturulmuştur. Bu arayüz üzerinden girilen verideki duyguları çıkarmak için kosinüs uzaklık formülü ile metin ile duygu vektörleri arasındaki uzaklık hesaplanmıştır. Farklı uzunluklarda ve farklı duygulara sahip metinler kullanılarak oluşturulan modelin sonuçları değerlendirilmiştir. Çalışmadan elde edilen duygu atamalarını değerlendirmek ve aynı metin girdi olarak verilen

ChatGPT'den çıkan sonuçları karşılaştırmak için çapraz doğrulama çalışması yapılmıştır. Çalışmanın sonuçları, duygu sözlüğünün metinsel verilere uygulandığında yüksek bir doğruluk sergilediğini göstermiştir. Belirlenen beş duygu arasından en az bir duygunun deneyimlenmesine ilişkin dört kişi tarafından fikir birliğine varılmış, böylece metinde aktarılan baskın duygunun başarılı bir şekilde tanındığı gösterilmiştir.

**Anahtar Kelimeler:** Türkçe Duygu Analizi, Duygu Tespiti, Duygu Çıkarımı, Doğal Dil İşleme

## **ACKNOWLEDGEMENTS**

I express my gratitude to Burkay Genç, my supervisor, with whom I became acquainted during the beginning of my master's degree. His guidance and support throughout my thesis journey, coupled with his insightful ideas and opinions, have been instrumental in achieving enhanced outcomes. Additionally, I extend my appreciation to Fatih Sağlam for his invaluable contributions and unwavering assistance during the development of my thesis.

I would also like to express my gratitude to my parents, Tefika Ünal and Hasan Ünal, for their constant encouragement, tolerance, and patience throughout the entirety of my graduate studies.

# CONTENTS

	<u>Page</u>
ABSTRACT .....	i
ÖZET .....	iii
ACKNOWLEDGEMENTS .....	vi
CONTENTS .....	vii
TABLES .....	ix
FIGURES .....	x
ABBREVIATIONS.....	xi
1. INTRODUCTION .....	1
1.1. Scope Of The Thesis .....	3
1.2. Contributions .....	4
1.3. Organization .....	4
1.4. Background and Definitions .....	5
1.4.1. Clustering Algorithms .....	5
1.4.2. Other Components .....	6
2. BACKGROUND OVERVIEW .....	8
2.1. Natural Language Processing .....	8
2.2. Emotion Analysis.....	11
2.2.1. Approaches on Emotion Analysis.....	12
2.2.1.1. Keyword-Based Approaches.....	13
2.2.1.2. Rule-Based Approaches.....	14
2.2.1.3. Machine Learning-Based Approaches.....	14
2.2.1.4. Deep Learning-Based Approaches .....	15
2.2.1.5. Hybrid Approaches .....	15
2.2.2. Data for Emotion Analysis .....	16
3. RELATED WORK.....	18
4. PROPOSED METHOD.....	23
4.1. Creating a Dataset .....	23



4.2. Correction of Generated Data .....	29
4.3. Clustering Emotion Groups .....	31
4.4. Clustering with Cosine Distance .....	34
4.5. Embedding with Generated Data .....	37
5. EXPERIMENTAL RESULTS .....	40
5.1. Performance Evaluation .....	40
6. CONCLUSION .....	50
APPENDIX .....	68
CV .....	88

## TABLES

	<u>Page</u>
Table 4.1 Example sentences from data. ....	24
Table 4.2 Emotion list used in this study. ....	25
Table 4.3 A snippet form the data from the first try. ....	28
Table 4.4 A snippet form the data after graph clustering. ....	30
Table 4.5 First four iterations of cosine distance clustering. ....	35
Table 4.6 The names and features of datasets. ....	36
Table 5.1 The result of emotion list checking. ....	40
Table 5.2 The result of scale methods. ....	40
Table 5.4 The results from participants' votes for every text. ....	46
Table 5.5 Emotion to number of participants agreement. ....	49

## FIGURES

	<u>Page</u>
Figure 4.1 Flowchart of data preparation.....	26
Figure 4.2 Graph for 'kuşku' emotion.....	29
Figure 4.3 Graph for PCA output.....	32
Figure 4.4 Silhouette score graph for PCA 0.94. ....	33
Figure 4.5 Interface for finding common words. ....	37
Figure 4.6 GUI process.....	38
Figure 4.7 GUI process.....	39
Figure 5.1 GUI output for the paragraph. ....	45
Figure 5.2 Graph showing the number of emotions agreed with the number of participants.....	49

## ABBREVIATIONS

<b>ePub</b>	:	<b>e</b> lectronic <b>P</b> ublication
<b>GAAC</b>	:	<b>G</b> roup <b>A</b> verage <b>A</b> gglomerative <b>C</b> lusterer
<b>k-NN</b>	:	<b>k</b> eep <b>N</b> earest <b>N</b> eighbors
<b>LSTM</b>	:	<b>L</b> ong <b>S</b> hort <b>T</b> erm <b>M</b> emory
<b>NLP</b>	:	<b>N</b> atural <b>L</b> anguage <b>P</b> rocessing
<b>PoS</b>	:	<b>P</b> art of <b>S</b> peech
<b>PCA</b>	:	<b>P</b> rincipal <b>C</b> omponent <b>A</b> nalysis
<b>TF-IDF</b>	:	<b>T</b> erm <b>F</b> requency - <b>I</b> nverse <b>D</b> ocument <b>F</b> requency

# 1. INTRODUCTION

The growth of social media platforms and the integration of mood meters into various websites have facilitated the comfortable and open expression of individuals' views. The interactions utilised by these communication channels have generated valuable data that can be applied in different fields of study. The use of textual data as a source is a common practice in the NLP sub-field known as textual analysis. Emotion analysis is a field of study that focuses on the study of emotions in textual data. The field of emotion analysis requires a language-specific emotion lexicon that takes into account differences in the expression and contextualisation of emotions.

Emotion analysis, focuses mainly on the identification and extraction of emotions from textual data. It allows the development of interdisciplinary studies with many applications. Firstly, it is important to note that emotional analysis has the potential to offer valuable in-depth look of the emotions conveyed by consumers in relation to a specific product, service, or brand. This information can boost client satisfaction and cultivate customer loyalty. Moreover, there is the ability to tailor marketing messages by utilizing the emotional state of the target consumers. This has a chance to optimize the efficacy of advertising campaigns. The use of emotional analysis exhibits promise within the field of fraud detection as it helps the discernment of unique emotional language patterns that are strongly associated with fraudulent activities. In general, the utilization of emotional analysis exhibits considerable promise in augmenting a wide range of company operations. Through the understanding and interpretation of the emotional states exhibited by consumers, employees, and other stakeholders, businesses possess the ability to optimize their decision-making procedures and improve their financial outcomes.

Although using of this type of analysis provides certain benefits, it also raises different challenges. For example, considering the presence of many levels of emotion recognition, spanning from document-level to word-level, it is crucial to effectively categorize textual material. An additional concern related to these analyses involves the presence of metaphor,

sarcasm, and irony in the text. In order to effectively interpret the intended meaning of the text, it is important to possess the appropriate contextual knowledge for the analysis of these words.

There are many different languages in the world and each has its own characteristics. Languages can be categorised into families, but even within these families languages can differ structurally. For this reason, language studies are usually conducted in the native language of the researcher. However, a significant number of researchers are also involved in the study of the English language, which has a widespread use worldwide. Many methodologies, libraries and datasets have been developed for English due to various studies on the language with different perspectives from different cultures. Unlike English, there are not enough libraries or data sources for Turkish. A significant part of the research in Turkish is covered by the use of English sources that have been translated into Turkish. The loss of emotion in translation and the need to adequately explain the meaning of the word are the result of linguistic differences. This has a detrimental effect on emotion analysis. Apart from translation resources, there are also resources produced from scratch, but this number is quite small compared to translation resources.

It has been observed that Turkish's source and library of emotion data lag behind English and that a comprehensive collection is required. While it is encouraging that Turkish has recently produced a number of studies on polarity and emotion analysis, it has been noted that the demand for a Turkish emotion data that encompasses all emotions in Turkish has not come across. This thesis's starting point has been this shortcoming.

One study conducted to investigate these deficiencies involved the development of a website where participants were asked to vote on the emotions elicited by a given sentence. The website sourced sentences of varying emotional valence from a variety of literary works. Nevertheless, the emotion categories that were surveyed could not be clustered together, and the emotions themselves could not be determined. Consequently, the development of this website has served as a source of inspiration for the present work, which encompasses

effective elements. Similarly, there was a significant demand for additional data in this context.

When discussing a text's use of emotion analysis, polarity—a text's analysis of its negative, positive, or neutral aspects—is the first concept that comes to mind when analyzing the language on its own terms. In fact, it has been observed that numerous studies has been conducted on the subject of sentiment analysis while more in-depth emotion analysis has not. It has been seen that, most often, just a small number of feelings are covered by the emotion analysis. This thesis also addresses this crucial issue.

## **1.1. Scope Of The Thesis**

The main aim of this thesis, which builds on the previously mentioned topics, is to compile a comprehensive data for emotions in Turkish. In this way, it is aimed to increase the accuracy of emotion analysis. Because it is useful to be able to evaluate the emotions in a text in many fields from politics to psychology, it has been determined that each language having its own data sources will provide better analysis results. With the development emotion data for Turkish, it will become one of the primary sources for researchers who study and research in Turkish.

The clustering of a large number of emotions in Turkish is another subject addressed in this thesis. In this study, the clustering algorithms used with the data generated within the scope of this study were investigated to see how closely the emotions are clustered with each other. Here, the emotion data has been modified to work with existing clustering algorithms.

Furthermore, this thesis investigated the behaviour of the data when extracting emotions from a text. This was achieved by creating an interface to extract emotions from text by calculating the cosine distance between text and emotion data entered during runtime.

In addition to providing solutions to the aforementioned challenges, this thesis is expected to serve as a foundation for future work in the field of emotion analysis.

## **1.2. Contributions**

In this research, we aim to address these shortcomings by proposing a novel, simple and efficient approach. The main contributions of this paper can be summarised as follows:

- It provides researchers working on Turkish emotion analysis with the most comprehensive and accessible data for Turkish emotions.
- Since the techniques used to develop emotion data are universal, it should be possible to adapt this research to other languages by modifying the book and the emotions used.
- It is feasible to refine and apply emotion data initially generated in a broad context in a domain-specific way.
- This study represents a new approach in Turkish language research, examining the complex processing of emotions and offering a new perspective for future research.

## **1.3. Organization**

The organization of the thesis is as follows:

- Chapter 1 of the thesis describes the mission to fill the gap created by scarcity of available resources for conducting emotion analysis in Turkish and the contribution of the work carried out specifically for this purpose to existing academic research.
- Chapter 2 provides a description of existing research on emotions, NLP and emotion analysis. This study is based on the findings of this research.
- Chapter 3 reviews research papers on emotion analysis. This paper provides a comprehensive discussion on the data used in emotion analysis, the methods used to generate the data, clustering algorithms and studies on distance algorithms.



- Chapter 4 summarises the process of generating a comprehensive set of emotion data, followed by clustering studies using this data to determine the proximity of emotions to each other.
- Chapter 5 discusses the procedures involved in extracting emotions from textual data entered into a testing interface. Purpose of this interface is to evaluate the quality of the data produced and to measure its effectiveness.
- Chapter 6 of this study highlights the limitations of the research and suggests avenues for future research. In particular, this chapter discusses potential areas of enquiry that could be explored to complement the findings of the current study and the potential contribution this study could make to the wider research field.

## **1.4. Background and Definitions**

One of the studies carried out in this thesis was to cluster emotions. For this purpose, the clustering algorithms and libraries used here are briefly summarised below. All experiments were performed with Python version 3.10.8.

### **1.4.1. Clustering Algorithms**

K-Means<sup>1</sup>: The library used is “sklearn.cluster.KMeans”. The K-Means algorithm employs a clustering technique by aiming to divide samples into n groups. The objective is to ensure that the clusters have similar variances, while simultaneously minimizing the inertia or within-cluster sum-of-squares. The number of clusters must be given for this algorithm. A set of N samples X is divided into K distinct clusters C by the procedure. The cluster “centroids” are another name for the means. In this study, “k-means++” for the init function and “lloyd” chosen as algorithm. Between 2 and 100 tried as number of clusters.

MiniBatchKMeans<sup>2</sup>: MiniBatchKMeans is an approach involves incorporating mini-batches into the K-Means algorithm in order to reduce computational overhead, while yet striving to

optimize the identical objective function. Mini-batches refer to randomly sampled chunks of the input data that are utilized during each training process. The use of mini-batches results in a notable reduction in computational requirements for achieving a local solution. The library for this clustering is “sklearn.cluster.MinibatchKMeans”. The variables used were “batch\_size=2048”, “init\_size=1024”, “random\_state=20”.

Hierarchical Clustering<sup>3</sup>: The library “sklearn.cluster.AgglomerativeClustering” is used for this clustering algorithm. Hierarchical clustering is a broad class of clustering methods that may successfully merge or break nested clusters. A tree is used to represent the cluster hierarchy. The tree’s root is the one-of-a-kind cluster that collects all the samples, while the leaves are clusters with only one instance. The variables “n\_clusters=40”, “metric=’braycurtis’”, “linkage=’average’” were given.

Spectral Clustering<sup>4</sup>: Spectral clustering applies clustering on a normalized Laplacian projection. This approach is especially beneficial when the structure of the individual clusters is significantly non-convex or, more broadly, when a measure of the cluster’s center and spread is insufficient to describe the entire cluster. Variables “n\_clusters = 40”, “affinity =’laplacian’” were given to the library “sklearn.cluster.SpectralClustering”.

GAAC<sup>5</sup>: The library “nltk.cluster.gaac.GAACClusterer” is used in the part. The GAAC method begins by treating each of the N vectors as a singleton cluster. It proceeds by iteratively merging pairs of clusters that have the closest centroids. This process is repeated until only one cluster remains. It closely resembles hierarchical clustering. The variables “num\_clusters=40”, “normalise=True” were given in the trials.

#### **1.4.2. Other Components**

PCA<sup>6</sup>: PCA is utilized to break down a multivariate dataset into a series of orthogonal components that account for the highest possible variance. In “sklearn.decomposition.PCA” library, PCA is implemented as a transformer object that learns a specified number of components during the fit process. These learned components can then be applied to new

data to project it onto the identified components. In this project for “n\_components” variable the values between 0.70 to 1.0 were tried.

MinMaxScaler<sup>7</sup>: The “sklearn.preprocessing.MinMaxScaler”: library, if different values are not given, matches the smallest value to 0 and the largest value to 1, and then adjusts the values in between accordingly.

Silhouette Score<sup>8</sup>: The silhouette score is a metric used to assess the validity of clustering outcomes by evaluating the degree to which each sample aligns with its assigned cluster in relation to adjacent clusters. It ranges from -1 to 1, with higher values indicating better clustering. The library for this is “sklearn.metrics.silhouette\_score”. In this part, while clusters are given for silhouette score calculation, a dictionary was created with K-Means clusters created with PCA component values between 0.7 and 1, and all values in this dictionary were tried.

Davies-Bouldin Score<sup>9</sup>: The score is the mean similarity measure between clusters and their most comparable cluster. It involves the comparison of distances within clusters to distances between clusters. If the algorithm gives lower scores than the clustering is more successful, with zero being the minimum score. For the library “sklearn.metrics.davies\_bouldin\_score” the same dictionary structure was used while experimenting here in the silhouette score.

Calinski and Harabasz Score<sup>10</sup>: The Calinski and Harabasz score is a measure used to assess clustering quality. It takes into account the ratio of dispersion between clusters to dispersion within clusters. Stronger clustering is indicated by higher scores., with well-separated and compact clusters being desirable. For the library “sklearn.metrics.calinski\_harabasz\_score” the same dictionary structure was used while experimenting here in the silhouette score.

## **2. BACKGROUND OVERVIEW**

### **2.1. Natural Language Processing**

One of the most important features that distinguish countries from each other is their language. Although language serves a vital purpose, its structure needs to be understood and explored. There are numerous philosophical studies on the connection of language with truth and reality. There are various perspectives and experiences with language, so it can be difficult to understand the text when others are talking about a topic [1].

The field of NLP has emerged as a sub-field within computer science that focuses on the subtleties of language, which is a complex and not yet fully understood structure. It involves analysing human-generated text and audio data using various methods. NLP, is a systematic and methodical approach to language. It is a special field in the field of Artificial Intelligence, concerned with the use of computational techniques to analyse, understand and produce human language. Language processing is considered a very important aspect of human intelligence, which is why it is included in the field of artificial intelligence [2–4].

NLP is a concept that has been studied since the 1950s. Its origins can be traced back to Alan Turing's pioneering work on the Turing test, followed by Chomsky's contributions in the early 1950s. In the 1950s, the first methodologies used in the field of NLP were based on the use of predetermined rules and templates programmed into various systems. As an example of this concept, there were established patterns that determined the appropriate translation of a particular word or phrase from one language to another. Another approach that builds on these approaches is an expert-based approach. A notable advantage of this approach is that it relies directly on expert knowledge, in contrast to rule and template-based methods. There are cases where this approach is effective. After the 1980s, it was recognised that rule-based methodologies were inadequate in light of the diversity, ambiguity and evolution of language. Since the 1980s, significant progress has been made in NLP. This was facilitated by the adoption of statistical methodologies based on language data observations, statistical analysis and machine learning algorithms [5–7].

NLP covers a large number of areas of interest. These areas can be listed as follows.

- Extracting information,
- Sentiment or emotion analysis,
- Correction of typos,
- Text summarization,
- Inter-lingual translation,
- Speech recognition,
- Speech production,
- Text voice over,
- Question and answer machines

Natural language problems are addressed at different levels of analysis, including morphology, lexical, grammatical, syntactic, part of speech (PoS), semantics, discourse, and phonetics.

Finding the right option among the various findings returned by the morphological analyzer for each word in a sentence can be called morphological analogy. Agglutinative languages such as Turkish, Finnish and Hungarian are challenging for NLP because of their complex morphemes. The morphemic analyzer may produce several parse outputs with various root and morpheme orderings due to the uncertainty introduced by the complex morphemic structure [8–10].

The process of automatically assigning part-of-speech tags to words in a sentence is referred to as part-of-speech (PoS) analysis. PoS analysis pertains to the examination of word categories, including but not limited to nouns, adjectives, verbs, adverbs, and prepositions. This is a commonly used technique in research studies that utilize machine learning [8, 11].

The integration of semantic analysis forms an important part of NLP methodology. In this analysis, an appropriate format is used to indicate the contextual framework of a sentence or paragraph. Semantics is a field of study focussed on examining the meaning and significance of language. The skilful use of word choices indicates the importance of the subject due to the interdependence of linguistic classifications. This subject, which is already a major topic of interest in academic studies, is concerned with the process of ambiguity of word meanings and is often referred to as word-meaning ambiguity [12].

NLP has recently experienced an increase in work on this topic with the discovery of new challenges. Despite the significant progress made in the field, these challenges persist. One of the common challenges in language processing is the presence of contextualised words and phrases. These linguistic elements can have more than one meaning within a sentence, and although they are understandable to human readers, they make comprehension difficult for the computer. The use of informal expressions, idioms and culture-specific terminology poses a challenge in the development of models for general use. Similar challenges arise in addressing a variety of domains, including the interpretation of terminologies or expressions hold potential for use within the realm of education but have different connotations in law, health, national security and so on. The effectiveness of NLP models can be optimised for a particular domain or geographical location, but it is imperative that such barriers are addressed when considering wider applicability. However, the availability of abundant data for regular training and updating of the models can improve the models [13, 14].

Apart from the previously mentioned challenges, the incorrect spelling or usage of words can pose a significant predicament. The development of auto-correction and grammar-correction applications has been a significant advancement due to ongoing progress in this field. Nevertheless, the task of precisely forecasting the author's purpose based on a specific region or geographical location, while considering variables such as sarcasm and colloquial language, continues to be a significant obstacle. NLP models, have demonstrated exceptional efficacy and ongoing advancement in their ability to process the most commonly utilized languages. However, there is still a requirement for models that cater to the language and technological proficiencies of individuals [15, 16].

## **2.2. Emotion Analysis**

Emotions are psychological states that have an impact on an individual's behaviour, mood and social interactions. Verbal and non-verbal communication are both essential components of human interaction. Verbal emotions can be understood through written communication, while non-verbal emotions can be distinguished through bodily gestures, tone of voice and facial demeanour. The task of understanding and assessing verbal emotions is of great importance in the field of NLP. This is largely due to the increasing amount of textual data, such as the increasing use of social media [17].

Sentiment analysis is an area of NLP developed to identify binary (positive, negative), ternary (positive, negative or neutral) or quaternary (positive, negative, neutral and mixed) viewpoints in textual sources. Text analysis also includes opinion mining, opinion inference, sentiment mining, subjectivity analysis, impact analysis and comment mining. Emotion analysis is a special field of sentiment analysis that involves the direct identification and labelling of sentiments within textual data, as opposed to the traditional approach of categorising opinions as positive, negative or others. In addition to the analysis of sentiments, emotion analysis is important as it engages in a more complex examination of emotions [18–24].

The analysis of emotions and sentiments within textual data holds considerable importance. This includes the ability to gauge the welfare of a community, prevent instances of suicide, and provide valuable insights for organizations seeking to measure customer satisfaction through the analysis of feedback and comments. Emotion and sentiment analysis can facilitate opinion mining for business organizations. Specifically, the text derived from e-learning and e-commerce settings can be examined for emotion analysis purposes [25]. It can also be used for handling of customer relations involves identifying the most precise response in accordance with assessments of products and services [26], as well as furnishing precise and sound feedback to departments responsible for product development. Also, it can be useful for the act of observing the temporal dimension of effective experiences across

various digital platforms pertaining to political figures, films, commodities, nations, and other related entities has been documented in literature [27].

Emotion analysis, which focuses on analyzing the emotions and thoughts of individuals, is commonly regarded as a sub-branch of NLP and sentiment analysis. However, it has also garnered attention from various other disciplines from politics to sociology to economics [18].

The development of an emotion model is deemed imperative for the purpose of accurately identifying and discerning emotions. The three primary models of emotion are the categorical approach, the dimensional approach, and the appraisal-based approach. The categorical perspective posits that there exists a finite set of fundamental emotions that are universally acknowledged. The dimensional approach posits that emotional states exhibit inter-relatedness across three dimensions, namely valence (positive or negative), arousal (excited or apathetic), and power. The appraisal-based approach expands upon the dimensional approach by integrating appraisal theory, which posits that emotions stem from an individual's assessment of events and their subjective experiences, objectives, and potential for action. Emotions are commonly conceptualized as alterations in various domains, including cognition, physiology, motivation, motor responses, effective experiences, and outward expressions [28–30].

### **2.2.1. Approaches on Emotion Analysis**

Normally, studies on emotion or sentiment analysis is carried out at three levels. These three types of analysis are document-level, sentence-level and aspect-based. Document-level analysis involves providing a document as input. After a thorough analysis of the entire document, the dominant emotion conveyed in the document is identified. Although multiple entities and emotions may exist within a document, it is ultimately categorised under a single emotional label. This feature makes it a traditional text classification problem. Sentence-level sentiment analysis involves processing a document at a granular level, specifically focusing on individual sentences. In most of the content there is no singular



subject and no singular active state attributed to this subject. Given that analyses at the document and sentence levels tend to remain at a higher level of abstraction, this problem is addressed through feature level analysis [8, 31].

There are also researches on utilizing five distinct categories for the purpose of recognizing emotions in textual data. The categorization of these methods includes keyword, rule, machine learning, deep learning-based, and hybrid approaches [18].

**2.2.1.1. Keyword-Based Approaches** The keyword-based emotion recognition methodology involves the identification of emotional lexicons within a given corpus of text and the subsequent assignment of an emotion label based on the presence of these lexicons. The dominant methodology for this approach involves the use of a keyword identification technique. This technique involves formulating a comprehensive catalogue of emotional lexicons for each emotion label. The current methodology involves an initial pre-processing step followed by the identification of emotion by referring to a predetermined lexicon of emotional terms. Before assigning an emotional classification to each sentence, evaluation is performed to assess the intensity of the emotion and to check for negativity [32].

As an example for this approach, Tao [33] devised a lexicon in question is organized based on the categorization of each word into two categories: a content or an emotion-functional word. The Emotional Functional Words were categorized into three distinct groups based on their linguistic functions: emotion keywords, modifier words, and metaphorical expressions. The study utilized PoS tagging, a semantic tree, and HowNet [34] to establish the correlation between the content word and the emotion-functional words. In order to identify emotions, the initial procedure involved utilizing a POS tagger, verifying the presence of emotion-functional words, assigning a rating to the emotion, allocating weights to each emotion keyword, and subsequently aggregating the amplitudes of the emotion-related keywords throughout all sentences. Nevertheless, a significant number of emotions were still misidentified.

**2.2.1.2. Rule-Based Approaches** Emotion recognition in text through rule-based approaches is dependent on pre-established rules and patterns that are used to characterize emotions present in textual data. Frequently, these are grounded on linguistic characteristics, such as lexical items, grammatical categories, and structural arrangements. An instance of a rule-based methodology could potentially categorize a sentence that incorporates the term "happy" as conveying a favorable emotion. The utilization of rule-based methodologies is characterized by its simplicity and interpretability. However, its effectiveness can be constrained by the extent of rules and the intricacy involved in defining rules for emotions that are more multifaceted.

Lee et al. [35] can serve as an example of this methodology. A model based on rules was proposed by them for the purpose of identifying events that cause emotions in the Chinese language. The researchers created a corpus that was annotated with emotions and their causes. They then computed the distribution of cause event types, position, and keywords. Additionally, they identified seven distinct groups of linguistic cues and developed two sets of linguistic rules. The experimental results indicate that the system exhibited favorable performance with respect to the identification of cause occurrence and cause event recognition.

**2.2.1.3. Machine Learning-Based Approaches** The process of recognizing emotions in the text can be achieved through either machine or classical learning approaches. This entails training models incorporating text and emotion labels, utilizing labeled datasets. Subsequently, these models can be employed to forecast the emotion of an unobserved textual data. Multiple machine learning algorithms, including support vector machines, decision trees, and neural networks, can be employed for this purpose. The choice of algorithm is contingent upon various factors, such as the task's intricacy, the dataset's magnitude, and the intended performance metrics. The flexibility and accuracy of machine learning methods surpass those of rules-based approaches, albeit at the cost of requiring substantial amounts of labeled data and potentially reduced interpretability [36].

The authors Ghazi et al. [37] introduced an innovative hierarchical methodology for the purpose of recognizing emotions. The text input is categorized into two distinct groups: sentences that convey emotions and those that do not, with subsequent analysis of the polarity of the sentences. The ultimate stage entails categorizing emotions that exhibit a negative polarity. Two experiments were conducted to compare the effectiveness of hierarchical and flat classifications. The results of the experiments indicated that the hierarchical approach exhibited superior performance compared to the flat classification approach.

**2.2.1.4. Deep Learning-Based Approaches** Applying deep learning techniques for emotion recognition in textual data involves utilizing multi-layer neural networks to extract features from the text and acquire the ability to predict emotions accurately. As mentioned earlier, the methodologies have gained significant traction in contemporary times owing to their capacity to process vast quantities of data and their cutting-edge efficacy in diverse natural language processing endeavors. Multiple architectures are available for recognizing emotions in text, including convolutional neural networks, recurrent neural networks, and transducers. These methodologies have the potential to yield high precision. However, they necessitate substantial quantities of annotated data and computing resources [18].

Ragheb and colleagues [38] have introduced a deep learning framework for the purpose of identifying emotions in written dialogues. The tripartite dialogue was concatenated and subsequently inputted Bi-LSTM that were trained using average stochastic gradient descent. The initial and final segments of the discourse were subjected to a self-attention mechanism, which was subsequently succeeded by an average pooling operation. The language model was trained using the Wikitext-103 [39] dataset; however, the findings indicated suboptimal performance in the identification of the happy emotion label.

**2.2.1.5. Hybrid Approaches** Hybrid methodologies for detecting emotions in textual data involve the integration of various techniques to enhance the precision and comprehensiveness of emotion identification. An approach that combines rule-based techniques for identifying basic emotions, such as happiness and sadness, with machine

learning or deep learning methods for detecting more intricate emotions, such as fear or anger, is referred to as a hybrid approach. Hybrid methodologies have the potential to capitalize on the advantages of diverse techniques and surpass the performance of singular approaches in isolation [18].

Park et al. [40], introduced and compared two models designed specifically for the task of multi-label emotion identification in tweets. The initial model implemented a linear regression approach that included label distance as a regularization component, whereas the subsequent model employed a logistic regression classifier. The researchers utilized a convolutional neural network to extract features from a distinct Twitter corpus that had been annotated with hashtags in a remote manner. This was done with the aim of acquiring emotive word vectors. In addition, two deep learning models were employed to acquire knowledge about emoji vectors.

### **2.2.2. Data for Emotion Analysis**

In order to conduct an analysis of emotions, it is necessary to have access to a dataset, lexicon, or corpora. The application of emotion labeling can be accomplished through utilization of the provided dataset. Several datasets have been curated for the purpose of conducting an analysis.

The simplicity and intuitiveness of discrete emotion models have led to their widespread acceptance in the field of emotion recognition. Emotions are categorized into discrete classifications. The task of recognizing emotions can be construed as a classification task that involves assigning one or more emotion labels to a given expression. As of present, a unanimous agreement has yet to be reached regarding the delineation of fundamental emotional categories. Paul Ekman's research on six fundamental emotions [28], namely anger, disgust, fear, happiness, sadness, and surprise are widely recognized as the universal emotions. Subsequently, Plutchik introduced a wheel of emotions [41] to delineate the interconnections among fundamental emotions, taking into account the impact of emotional

correlation. The wheel categorizes eight fundamental emotions, namely joy, sadness, fear, anger, trust, disgust, surprise and anticipation.

The dimensional emotion model quantifies emotional states using numerical dimensions, whereby each emotional state is depicted as a vector with multiple dimensions. The PAD model is a frequently employed dimensional model that delineates emotions into three distinct dimensions, namely pleasure, arousal, and dominance. According to Russell's circumflex model, the two dimensions of Valence and Arousal have been proposed as the most distinct emotions. The task of constructing a dimensional emotional corpus of high quality is a challenging one, and the current selection of publicly available corpora is restricted in scope [42–44].

Data for emotion analysis can be categorized into two distinct classes based on the length of the text: short text and long text. The majority of early studies on emotion analysis and classification employ brief textual content, such as news headlines and micro-blogs. Several studies have incorporated lengthy textual materials, including news, blog posts, children's stories, and amalgamations of datasets. The previously mentioned datasets are characterized by ease of manipulation, yet may pose challenges in accurately discerning affective states [45–48].

### **3. RELATED WORK**

This thesis focused on data, a crucial component of emotion analysis. Under this section, the study examines data generation methods from various languages that illustrate the previously mentioned reasons. Despite the abundance of resources available for the English language, there is a dearth of resources for the Turkish language. The majority of the sources observed were generated through the process of translation from the English language [49].

The study of emotions in writing has been greatly facilitated by the vast amount of data available. Individuals have increasingly used news websites, social media platforms, blogs, and online forums as a means of expressing their emotions and opinions. Extensive research has been conducted on polarity analysis utilizing this type of data. Within the context of polarity analysis, information is categorized as either positive, negative, or neutral. In contrast, emotion analysis entails a fragmented polarity analysis structure. For instance, “I was scared”, “This job really pissed me off” sentences are assigned a negative label. The initial statement conveys a sense of fear, while the subsequent statement connotes anger. This is the main difference between sentiment and emotion analysis [50, 51].

English is widely recognized as a global language and boasts a plethora of data that can serve as a foundation for analysis. Consequently, it is the language that offers the most abundant resources for emotion analysis. The ISEAR [52] dataset is widely recognized and frequently employed in numerous studies. The dataset comprises 7,666 sentences and encompasses seven distinct emotion labels. The labels, as mentioned above, encompass a range of emotions: anger, joy, disgust, fear, shame, surprise, and sadness.

An additional dataset that is available for analysis is the AFINN [53] lexicon. It encompasses a vocabulary of over two thousand entries, which have been categorized as either positive or negative in nature. Furthermore, several other datasets exist, such as Alm [54], Aman [55], SemEval-2007 [46], SemEval-2018 [56], SemEval-2019 [57], and Neviarouskaya [58, 59]. The datasets encompass a range of emotions, such as anger, sadness, joy, surprise, shame, guilt, anticipation, optimism, pessimism, confidence, and love. The ISEAR dataset is the sole

example of a balanced dataset, while the Alm and SemEval-2007 datasets exhibit the lowest number of samples. The SemEval-2019 dataset exclusively comprises textual dialogues.

Chaffer and Inkpen's study [60] involved the extraction of six distinct emotions from a heterogeneous dataset comprising various sources, including news headlines, fairy tales, and blogs. The study conducted by Kouloumpis, Wilson, and Moore [61] aimed to assess the impact of incorporating features in supervised learning methodologies for the purpose of Twitter sentiment analysis. The attainment of this objective was facilitated through the implementation of three distinct structures for Twitter messages. The training datasets utilized comprised two different components, namely labeled words and expressions. In addition, the models were tested using a dataset that had been annotated.

Corporas have been developed specifically for these fields of study. Various are available, but the most widely recognized and utilized is the WordNet [62]. It comprises sets of synonyms, also known as synsets. A synset consists of a collection of synonyms with distinct semantic significance. The development of WordNet-Affect [63] involved the utilization of numerous key terms, amounting to several hundred, which were subsequently allocated to distinct categories of emotion. In the subsequent phase entails scrutinizing the WordNet lexicon for words that bear resemblance to the fundamental terms and allotting each of those words the identical affective classification as the corresponding root term. To facilitate sentiment analysis, the SentiWordNet [63] tool incorporated positive, negative, and neutral polarity valence scores into the pre-existing WordNet lexicon. The SenticNet [64] framework employs a concept-level approach to conduct emotion analysis, which endeavors to encapsulate the intricate nature of emotional expression. The dictionary comprises distinct sections, each of which is assigned a polarity score that spans from -1 to 1. The lexicographical resource is equipped with labels that delineate the affective, dispositional, and attitudinal dimensions exhibited by each lexical item. The MPQA [65] lexicon is an additional resource for polarity classification. The lexicon, as mentioned earlier, comprises a total of 8222 entries and has been categorized as either positive, negative, or neutral. Furthermore, it is worth noting that several lexicons are at one's disposal, including but not limited to the Bing Liu Dictionary [66], the NRC Word-Emotion Association

Dictionary [67], the NRC Impact Intensity Dictionary [68], the NRC Value, Arousal and Dominance (VAD) Dictionary [69], the NRC Hashtag Emotion Dictionary [70], and the Sentiment140 Dictionary [71].

Besides English, emotion lexicons can be found in a few other languages. These include Spanish [72–76], German [77–80], Polish [81–83], Chinese [84, 85], Italian [86], Portuguese [87], Dutch [88], Indonesian [89], Greek [90] and Croatian [91].

The number of studies focused on data collection for emotion analysis in Turkish is notably lower in comparison to those conducted in English. In the majority of studies, data was generated through the process of translation from numerous English resources into Turkish rather than originating *de novo*.

Vural et al. [92] performed sentiment analysis using the SentiStrength infrastructure made for English and translated into many languages. Meral and Diri [93] manually tagged 8,500 tweets and performed sentiment analysis with Random Forest, Support Vector Machine, and Naive Bayes classifiers from word capture perspective and machine learning methods. Mayda and Aytakin [94] developed an opinion-mining model for the comparison task for competition analysis in social media.

Akba et al. [95] performed emotion classification on the movie dataset with the help of feature selection algorithms and support vector machine with completely unsupervised machine learning techniques and achieved approximately 84% success. Çetin and Amasyalı [96] conducted many experiments on Turkish Twitter data and compared the results of the experiments. Özsert and Özgür [97] developed a semi-automatic method to determine word polarities, and Türkmenoğlu and Tantuğ [98] focused on deciding which of the dictionary-based and machine learning-based approaches performs better in Turkish.

Uçan [99] translated the SentiWordNet sentiment dictionary into Turkish using three different translation engines.

The study conducted by Atlı et al. [100] introduced the NAYALex lexicon, which utilizes the NRC Emotion Lexicon to identify a wide range of emotions (such as hope, anxiety, love,



pessimism, optimism, anger, fear, and sadness) from textual data. The authors suggest that this approach has the potential to identify additional emotions from texts that are shared. The NRC Emotion Lexicon is a compilation of 6,469 English words, each of which is linked to at least one of 38 distinct emotions. This lexicon was developed based on Plutchik's Theory of Emotions [41].

In addition to the translation datasets, Açııcı [101] requested that 500 university students recount a personal memory pertaining to seven distinct emotional categories. The dataset comprises a collection of around 3,200 documents, as provided by her. In contrast, Demirci [102] collected a corpus of six thousand tweets containing Ekman's research on six basic emotions [28] on the social media platform Twitter. The individual conducted emotion classification by utilizing machine learning algorithms on the tweets that were acquired. The individual achieved optimal outcomes while analyzing the data using Naive Bayes, Support Vector Machine, and k-Nearest Neighbors techniques. Specifically, the Support Vector Machine approach yielded the highest success rate of 69.2%. The sole existing dataset for conducting emotion analysis in Turkish language is TREMO [49, 103]. This dataset represents the initial Turkish corpus designed exclusively for the purpose of implementing emotion analysis methodologies. The study is predicated on a survey that involved 5,000 participants who were requested to compose narratives concerning their most salient memories of the six emotions posited by Ekman's research [28].

Furthermore to these studies, lexicon studies have also been carried out using news in Turkish as data. Sağlam et al. [104]'s study introduces two original contributions: the development of a polarity-scored Turkish sentiment lexicon through the utilization of online news media, and the evaluation of existing lexicons on texts that are not specific to any one subject. The polarity lexicon of SWNetTR, consisting of 27,000 words, was enhanced to include an additional 10,000 words, resulting in a total of 37,000 words. The findings indicated that SWNetTR, which consisted of English translations, attained a polarity classification accuracy of 60.6%. In contrast, SWNetTR-PLUS, which was created throughout the course of this research, acquired a higher accuracy of 72.2%. Another study from Sağlam et al. [8] extended the previously mentioned SWNetTR-PLUS and achieve a successful analysis

rate comparable to translation-based approaches. The lexicon is extended to SWNetTR++, with a word count of 49K, achieving higher success rates compared to SWNetTR and SWNetTR-PLUS. The approach uses synonym-antonym datasets and a tone propagation algorithm to compute tone scores.

Additional study of Mertoğlu et al. [105] is about detecting fake news in Turkish using lexicons. It gathers a large corpus of labeled Turkish news texts and generates four distinct lexicons. The models are based on the agglutinative structure of the Turkish language and generate root words, suffix forms, and raw words with POS tags. With a recall ratio of 0.929, the raw form model was the most effective at detecting false news, according to the results. The models that combine root words and raw words with part-of-speech identifiers also perform well. The suffix model, which lacks content awareness, obtains a high recall of 0.814, indicating that suffixes carry significant information in agglutinative languages and should be considered in studies such as emotion or sentiment analysis.

## 4. PROPOSED METHOD

Since available resources were observed to be insufficient to conduct emotion analysis in Turkish, this research was initiated to create a comprehensive data set. The scarcity of such sources necessitated this study, which tries to fill this gap in the literature by creating a comprehensive data set that can be used to conduct sentiment analysis in Turkish. During the data generation process, a series of experiments were conducted on various components. In this section, a comprehensive and detailed description of all the work done will be provided.

### 4.1. Creating a Dataset

Given the research objective at hand, it has been determined that the data produced in prior studies was not suitable for the purpose of conducting emotional analysis. Hence, it was deemed more beneficial to generate the necessary data from scratch. The initial phase of this research involved the identification of the data source. Given the multifaceted nature of emotions depicted in literary works, it was deemed suitable to use books as a way to perform emotion analysis.

The research used a corpus consisting of 100 unique literary pieces. Appendix-1 offers a collection of the books that have been discussed. During the process of picking novels, careful consideration was made to ensure that the chosen books adequately represent the diverse spectrum of emotions experienced in Turkish.

The ePub format was used for the extraction of sentences from books. The extraction of individual sentences was simplified through the use of the Python ebooklib<sup>11</sup> library [106]. The books underwent various processing techniques, resulting in the extraction of a total of 660,000 sentences. The sentences extracted from the books are shown in Table 4.1.

---

<sup>11</sup><https://pypi.org/project/EbookLib/>

Table 4.1 Example sentences from data.

Sentences
Alsın benim kitapları okusun işte! dedi baba.
Altı ay önce boşandık dedim.
Ama size söyledim: Onun intikamını almanızı istiyorum.
Anne ne oldu burada? diye sordu.
...
Anne ne olursun söyleme, böyle konuşma dedim ve ağladım.
Anne, beni niye suçluyorsun, ne yaptım ki ben? dedim.
Alın sinir ağrı nedir? diye sordu Amparo.
Ellerini oynattı ve parmaklarıyla uzandı.
Sürekli kıkırdayıp gülümsüyordu.
...

Subsequent to the extracting of said sentences, a process was initiated with the aim of identifying the emotions expressed in the Turkish language. In this study, a compilation of emotions in the Turkish language was generated through the utilization of Ekman's [28] and Plutchik's researches on [41] emotional categories, which were further refined through the inclusion of Turkish synonyms and antonyms associated with emotions. This emotion list contained 213 different emotions. The list mentioned above is observable in the Table 4.2.

After the extraction process of sentences which can be seen in Figure 4.1 and the identification of sentiments in Turkish were completed, the first task was to search for these emotions in the sentences. However, the process of parsing sentences from ePub format resulted in errors related to the reading procedures of Turkish characters. As an example, the character denoting "ı" has undergone a modification to represent the letter "i". Furthermore, techniques for cleaning Twitter data [107] were employed during the process of word cleaning. Throughout the investigation, a comprehensive record of words linked to each emotion was concurrently maintained. The initial findings of this inquiry revealed the co-occurrence of distinct emotions and the frequency.

Table 4.2 Emotion list used in this study.

kafa karışıklığı	hayal kırıklığı	can sıkıntısı	kafayı takma	vicdan azabı	memnuniyetsiz	motivasyonlu	alçakgönüllü	samimiyetsiz
anlaşılmama	ciddiyetsiz	aşağılamama	tereddütsüz	böbürlenme	düşüncesiz	hassasiyet	hayırsever	memnuniyet
motivasyon	sahiplenme	soğukkanlı	teslimiyet	çekincesiz	nezaketsiz	vicdansız	adaletsiz	anlaşılmama
aşağılama	başarısız	beğenmeme	depresyon	dikkatsiz	düşünceli	ferahlama	hoşnutsuz	küçümseme
melankoli	saldırgan	tedbirsiz	anksiyete	sıkıntılı	endişesiz	vicdanlı	temkinli	bağımsız
beklenti	ciddiyet	dengesiz	dikkatli	duyarsız	güvensiz	huzursuz	isteksiz	karamsar
kararsız	kaygısız	kayıtsız	keyifsiz	kırılgan	korkusuz	kötümser	merhamet	paranoya
rahatsız	sabırsız	saygısız	sevgisiz	tedbirli	tedirgin	tereddüt	tutarsız	utanmama
vazgeçme	yüceltme	alınmama	endişeli	gurursuz	pervasız	suçluluk	şüphesiz	beğenme
agresif	aidiyet	bağımlı	çekince	çaresiz	cesaret	fanatik	gerilim	gücenme
heyecan	histeri	iğrenme	perişan	imrenme	intikam	iyimser	ızdırap	kararlı
kınamak	kıskanç	nezaket	rekabet	sadakat	sempati	sinirli	sükunet	tutarlı
ümitsiz	umutsuz	utangaç	uyumsuz	yenilgi	onursuz	acımama	azimsiz	dostluk
durgun	tiksin	adalet	alınma	beğeni	bencil	bıkkın	bitkin	çılgın
cinnet	cömert	dalgın	dehşet	dürüst	empati	gergin	hasret	hayran
hayret	hiddet	hoşnut	hüsran	huysuz	inatçı	kızgın	kurnaz	mahcup
mahzun	minnet	mutsuz	nefret	onurlu	pişman	samimi	şaşkın	şefkat
şehvet	sevinç	şükran	takdir	tevazu	üzüntü	hırslı	aşksız	kinsiz
utanma	acıma	cesur	çoşku	garez	gıpta	gurur	güven	heves
hüzün	huzur	ilham	itaat	kabul	kaygı	keder	keyif	kibir
korku	merak	mutlu	özgür	özlem	panik	rahat	sabır	sakin
saygı	sevgi	stres	şüphe	tutku	utanç	üzgün	zulüm	aciz
aksi	arzu	azap	azim	neşe	öfke	şevk	ümit	umut
uyum	zevk	aşk	kin	şok	yas	haz		

During the execution of these procedures, the TF-IDF scores of the terms that co-occur with emotions were also computed. The TF-IDF metric is a mathematical formula that entails the product of two distinct variables. The objective of this computation is to estimate the probability of a term's appearance while considering the normalization of its frequency by the overall frequency of the document or set of documents [108]. The rationale behind this metric is that a term's significance in a document is directly proportional to its frequency of occurrence, thereby warranting a higher weight. In the event that a term is present in a significant number of other documents, it is likely not a distinctive identifier. Consequently, it is advisable to allocate a reduced weight to said term. The math formula for this metric is in Equation 1. The variable "t" is used to represent the terms, "d" is used to represent

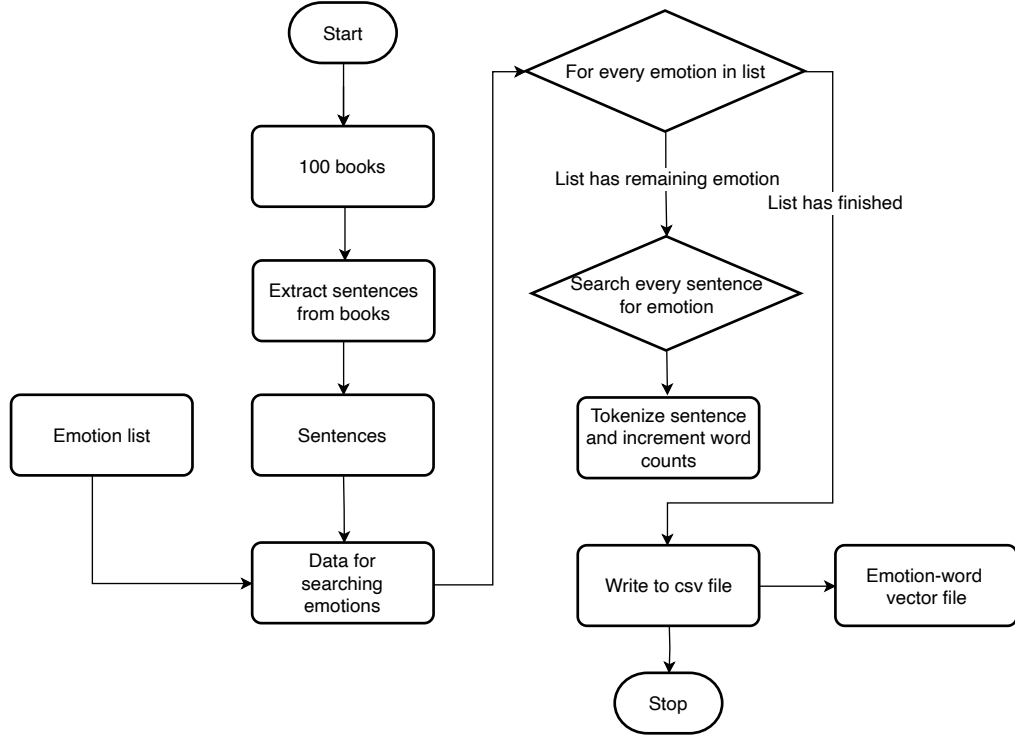


Figure 4.1 Flowchart of data preparation.

each individual document, and “D” represents the entire collection of documents. The  $tf$  formula includes the frequency of words with emotion. Every single one of the documents comes from the emotion list. Therefore, each emotion is a document. The formula  $N_t$  is the number of documents that includes the term  $t$ . When calculating the  $idf$  formula, it is found by dividing the number of all documents by the number of documents containing that word.  $tf\_idf$  is obtained by multiplying the results of  $tf$  and  $idf$  with each other.

$$tf = |t|_d$$

$$N_t = \sum_{d=1}^D \mathbb{I}_{t \in d} \tag{1}$$

$$idf = \ln\left(\frac{|D|}{N_t}\right)$$

$$tf\_idf = tf \times idf$$

For example, if the  $tf$  value is 5 if an X word appears 5 times in conjunction with the Y emotion. If the word X is on a list of 10 distinct emotions and we are dealing with a total of about 213 emotions, the  $idf$  value is calculated with Equation 2. The  $tf\_idf$  is determined by multiplying 3 by 5 and arriving at an approximation of 15 as can be seen at Equation 3.

$$N_t = 10, D = 213, tf = 5$$

$$idf = \ln\left(\frac{213}{10}\right) \simeq 3.06 \quad (2)$$

$$tf\_idf = 5 * 3.06 \simeq 15.3 \quad (3)$$

The initial study's findings indicate that emotions have the potential to be categorized into distinct groups. Initially, it was recognized that the dimensions of the word vectors associated with each emotion ought to be uniform. Consequently, 660,000 sentences were analyzed to extract words that corresponded to 213 distinct emotions. To accomplish this task, a vector was generated by extracting the initial five letters of each term [49, 109].

For example as a summary of the methods used so far, in the sentence “Aşk daha üstündür, diye cevap verdi genç Balıkçı ve denizin dibine daldı; Amy da ağlaya ağlaya bataklıktan uzaklaştı.”, all punctuation marks are first removed. As a result, the sentence becomes “Aşk daha üstündür diye cevap verdi genç Balıkçı denizin dibine daldı Amy da ağlaya ağlaya bataklıkta uzaklaştı”. Then all letters are converted to lower case. After the sentence becomes “aşk daha üstündür diye cevap verdi genç balıkçı denizin dibine daldı ruh da ağlaya ağlaya bataklıktan uzaklaştı”, the whitespaces are removed. As a result, the sentence “aşk daha üstündür diye cevap verdi genç balıkçı denizin dibine daldı amy da ağlaya ağlaya bataklıkta uzaklaştı” is obtained. After this sentence is obtained, it is checked whether there are names in the stopwords or nltk names dataset. Since the stopwords “ve”, “daha” ve “diye” and the noun “amy” are found, these elements are also removed from the sentence. As the last sentence to be searched for emotion, “aşk üstündür cevap verdi genç balıkçı denizin dibine daldı da ağlaya ağlaya bataklıktan uzaklaştı”. In this sentence, the presence of the emotion

“aşk” in the emotion list is seen. Since the emotion “aşk” is found in this sentence, the word counting step is continued. Since the word “aşk” in the sentence should not be in the emotion vector “aşk”, this sentence is translated as “what üstündür cevap verdi genç balıkçı denizin dibine daldı da ağlaya ağlaya bataklıktan uzaklaştı”. Any word other than Turkish can be used for this change. Because after the words are separated, it is also checked whether the word is Turkish or not. In the emotion vector “aşk”, the number of the other words in the sentence “üstün, cevap, verdi, genç, balık, deniz, dibin, daldı, ağlay, ağlay, bataklık, uzakla” are increased one by one. This sentence is analysed for all other emotions in the list. For this, the sentence is analysed again by putting the emotion word again. The reason for this is to see the relationship between one emotion to another and not to spend the sentence only on one emotion. However, since only the emotion “aşk” is mentioned in the sentence, the research process for this sentence is completed here.

After processing all the sentences the magnitude of the vector was determined to be 213x28,268. A data was generated through the computation of the term frequency-inverse document frequency values for individual words. The words in the vector are arranged in descending order according to their idf values. A part of the data created in the first step can be seen in the Table 4.3.

Table 4.3 A snippet form the data from the first try.

<b>Emotion</b>	<b>Related Words</b>
kafa karışıklığı	‘dolgu’: [3,6.037], ‘türle’: [2, 4.72], ‘kazar’: [1, 2.47], ‘venüs’: [2, 1.98], ‘yönle’: [2, 1.92], ...
...	...
nefret	‘ediyo’: [358, 57.55], ‘ederd’: [70, 55.81], ‘ettiğ’: [186, 35.39], ‘ondan’: [93, 32.77], ‘bende’: [82, 30.65], ...
...	...
aşk	‘tanrı’: [361, 105.04], ‘bir’: [2844, 56.88], ‘kadın’: [366, 54.61], ‘roman’: [54, 46.11], ‘hayat’: [183, 33.72], ...





dissociated from the node and establish a novel cluster in distinct scenarios. This is achieved by traversing to the leaf word and examining its tone status with an upper node, provided that both exhibit positive value. A total of 900 groups were eliminated through manual selection, and those groups that did not occur in at least 100 sentences were also excluded. The outcome of this procedure yielded a grand total of 253 clusters that were utilized. Upon analysis of these groups, it was observed that certain words exhibited a common root sense, albeit with varying affixes, across multiple groups. The process of segregating the groups involved the consideration of values that held substantial importance. The values selected for this purpose were -0.1 and 0.1, which were deemed to be of notable significance. Given these specified values, in cases where the tone of a word fell within the range of said values, the group was permitted to proceed without undergoing division. Thus, our objective was to avoid excessive fragmentation. After this work, you can see a part of the final data in the Table 4.4

Table 4.4 A snippet form the data after graph clustering.

<b>Emotion</b>	<b>Related Words</b>
(alçakgönüllülük, alçakgönüllü)	‘kendi’: [85, 42.35], ‘tanrı’: [39, 40.10], ‘göste’: [46, 37.57], ‘karşı’: [49, 30.39], ...
(hassasiyeti, hassasiyetimiz, hassasiyetimizle, hassasiyet)	‘göste’: [12, 9.80], ‘hisse’: [8, 7.02], ‘elino’: [4, 6.22], ‘büyük’: [8, 6.04], ...
(saldırganlık)	‘öfkeden’: [4, 4.74], ‘öfkeyle’: [4, 4.74], ‘öfkelenir’: [4, 4.74], ‘öfke’: [4, 4.74], ...
...	...
(tedirginliğimiz)	‘geçti’: [1, 1.13], ‘zaman’: [1, 0.60], ‘acele’: [0, 0], ‘edeli’: [0, 0],...
(düşüncesiz, düşüncesizce)	‘davra’: [11, 10.01-], ‘kendi’: [20, 9.96], ‘patav’: [2, 8.90], ‘olduğ’: [16, 8.55],...
...	...

### 4.3. Clustering Emotion Groups

Following the collection of data, it was proceeded to categorize these distinct emotional groups into clusters. Initially, the data was prepared for the algorithms. The initial step involved identifying solely the emotional state and the term frequency values associated with said emotion within the rows of the dataset, as well as the corresponding words present in the columns. Consequently, a matrix consisting of words associated with emotions was established. The K-Means [111] algorithm, widely recognized as a prominent clustering algorithm, was employed to derive a result from the input data. The analysis revealed that a considerable number of autonomous clusters were combined into a single cluster. Consequently, the initial set of extracted words underwent a preprocessing step wherein any stopwords that were not identified by the NLTK [112] library were eliminated. The process of identifying stopwords involved an examination of the idf values associated with each word. Following the removal of stopwords, the data was reintroduced into the clustering algorithm, however, a satisfactory outcome was not achieved. Subsequently, it was acknowledged that the K-Means algorithm's efficacy is limited when confronted with a high number of dimensions. It was determined that the utilization of the PCA [113] algorithm would be a novel approach for this purpose.

Prior to being fed into any algorithm, the data underwent a scaling process. The data was subjected to scaling using the MinMaxScaler [114] function from the `sklearn.preprocessing`<sup>12</sup> library in Python, prior to being inputted into the algorithms.

To ascertain the number of output features yielded by the PCA algorithm, it is necessary to predefine the number of components as a percentage value. The objective of this study was to investigate the number of features selected by the PCA algorithm for each component value within the range of 0.70 to 0.99, prior to finalizing a singular decision. Figure 4.3 displays a graphical representation of this phenomenon. The horizontal axis represents the value of the component, while the vertical axis indicates the number of features that will be produced.

---

<sup>12</sup><https://scikit-learn.org/stable/modules/preprocessing.html>

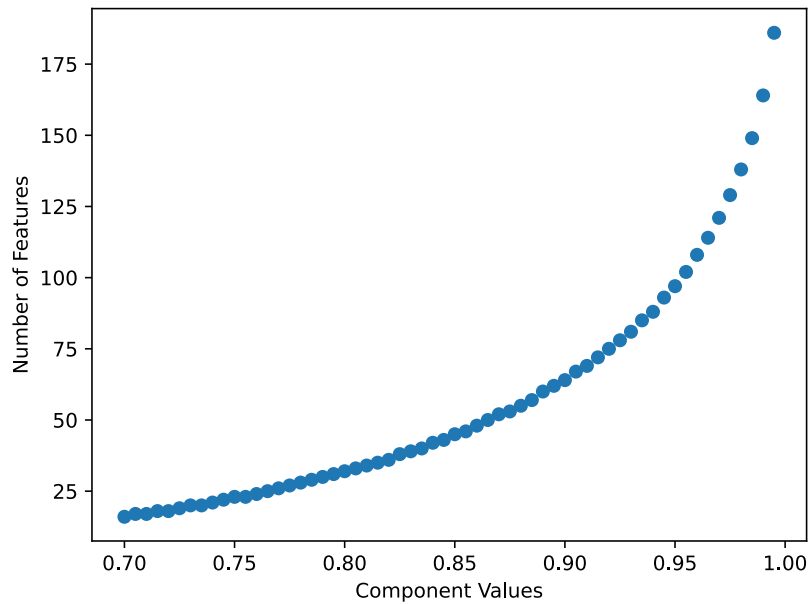


Figure 4.3 Graph for PCA output.

The rationale for utilizing the silhouette score pertains to the ambiguity surrounding the identity of the features derived from PCA, and the uncertainty regarding which features would yield optimal outcomes for K-Means. The silhouette score is a metric that quantifies the degree of cohesion and separation of data points within a cluster. In other words, the higher the silhouette score, the more successful clustering was performed. Apart from the silhouette score, the Calinski-Harabasz and Davies-Bouldin scores are also significant metrics. The present investigation incorporates the silhouette score as a metric [115–117].

To operationalize the silhouette score, a framework was devised to incorporate each component value generated by PCA into the K-Means algorithm, from which the corresponding silhouette scores were extracted. The prior configuration generates a graphical representation for every constituent value, thereby enabling the observation of the silhouette score for each individual cluster. The provided Figure 4.4 displays a graph representing the component value of 0.94.

Subsequently, the obtained values underwent individual testing. Divergent clustering outcomes were derived based on the evaluation metrics of silhouette scores and PCA component values. Nevertheless, due to the attempted values, the emotions aggregated

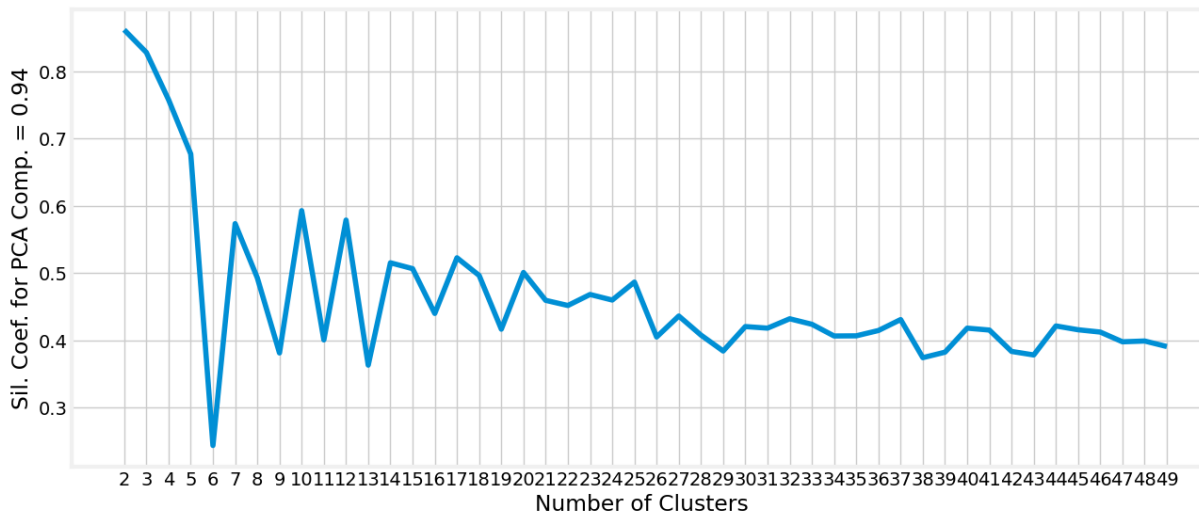


Figure 4.4 Silhouette score graph for PCA 0.94.

together, leading to an unsuccessful outcome. Various clustering algorithms have been experimented with, in addition to K-Means. The algorithms includes hierarchical, mini-batch, spectral clustering and NLTK’s GAAC clustering.

The methodology of hierarchical clustering is distinct from that of partition-based clustering in that it constructs a binary merge tree, commencing from the leaves and culminating at the root. In order to execute a hierarchical clustering algorithm, it is necessary to select a linkage function that establishes the distance between any two subsets and utilize the fundamental distance between individual elements. The term “dendrogram” is commonly used to refer to this particular graphical representation [118].

Typically, spectral clustering entails extracting the leading eigenvectors of a matrix that is constructed based on the inter-point distance or other relevant attributes. These eigenvectors are subsequently utilized to partition the data points into distinct clusters [119].

Mini batch K-Means algorithms use small random batches of examples of a fixed size to store in memory. Each iteration, a new random sample is obtained and used to update the cluster. The learning rate decreases with the number of iterations, and the effect of new examples is reduced as the number of iterations increases. Convergence can be detected when no changes in the clusters occur in several consecutive iterations [120].

The GAAC clustering algorithm initiates by considering each of the N vectors as individual clusters. The process involves the iterative merging of pairs of clusters that exhibit the closest centroids [121].

#### 4.4. Clustering with Cosine Distance

Subsequently, in light of the algorithms' insufficient efficacy, the subsequent course of action was to employ clustering utilizing direct cosine distance computation. The cosine similarity metric is widely employed in the domain of information retrieval and other associated scholarly fields. The metric in question employs a vector of terms from a given text to construct a model, and the similarity between two texts is established by calculating the cosine distance - as shown in Equation 4 - of their corresponding term vectors [122]. Vectors have been generated from the words present in it, along with emotional expressions, for this particular objective. Binary cosine distance values were computed between the preceding vectors. A combination of groups that are in close proximity to one another is performed, followed by the consideration of the merged groups as a single entity in the subsequent step, wherein computations are executed with the remaining clusters. The aforementioned procedures were iterated until the point where a one cluster remained. Prior to performing the cosine distance computation, a process of data scaling was implemented. In this context, scaling operations are attempted both column and row wise.

$$\cos(x, y) = 1 - \frac{x \cdot y}{\|x\| * \|y\|} \quad (4)$$

During the initial four iterations, the clustering process demonstrated a high degree of success in merging the clusters. However, in subsequent iterations, due to insufficient data and a paucity of information, the clustering process resulted in the merging of clusters at a single point, leading to a snowball effect. In Table 4.5 you can see the first 4 iterations of this experiment.

Table 4.5 First four iterations of cosine distance clustering.

Clusters	Cosine Distance
(‘umursadığımı’, ‘umursadığım’) — (‘vazgeçmesine’, ‘vazgeçmesini’, ‘vazgeçmeyeceğinden’, ‘vazgeçmeyeceklerini’)	0.021
(‘tutarsızlığı’, ‘tutarsızlıkların’, ‘tutarsızlıkları’, ‘tutarsızlığın’, ‘tutarsızlık’, ‘tutarsızlıklarına’, ‘tutarsızlıklar’, ‘tutarsızlıktır’, ‘tutarsız’) — (‘huzursuzuz’, ‘huzursuzum’, ‘huzursuz’)	0.051
(‘hüzünlüydü’, ‘hüzünlüyüm’, ‘hüzünde’, ‘hüzünleneceğiz’, ‘hüzünle’, ‘hüzünlüyüz’, ‘hüzün’, ‘hüzünlü’, ‘hüzünden’, ‘hüzünlüdür’, ‘hüzünlerimizi’, ‘hüzünlerin’) — (‘kaygılarla’, ‘kaygımızı’, ‘kaygıdır’, ‘kaygılarımız’, ‘kaygımız’, ‘kaygısı’, ‘kaygılıyım’, ‘kaygısıyla’, ‘kaygılandırılmaktadır’, ‘kaygılarımızı’, ‘kaygıların’, ‘kaygılıdırlar’, ‘kaygıdan’, ‘kaygılı’, ‘kaygılara’, ‘kaygısında’, ‘kaygınız’, ‘kaygılardan’, ‘kaygılar’, ‘kaygımızın’, ‘kaygısının’, ‘kaygıyla’, ‘kaygılanıyorlar’, ‘kaygısıdır’, ‘kaygılıyız’, ‘kaygısal’, ‘kaygının’, ‘kaygı’, ‘kaygıyı’, ‘kaygısına’, ‘kaygısını’, ‘kaygıya’, ‘kaygınızı’, ‘kaygılarıyla’, ‘kaygılarından’, ‘kaygıları’, ‘kaygılıları’)	0.0511
(‘bunaldık’, ‘bunaldıkların’, ‘bunalınca’, ‘bunalarak’, ‘bunalıp’, ‘bunaldı’, ‘bunalmıştır’, ‘bunaldıkları’, ‘bunalmış’, ‘bunaldım’, ‘bunalanların’, ‘bunalanlar’, ‘bunalan’, ‘bunaldığı’, ‘bunaldığını’, ‘bunaldığımızda’, ‘bunaldığımız’) — (‘tutarsızlığı’, ‘tutarsızlıkların’, ‘tutarsızlıkları’, ‘tutarsızlığın’, ‘tutarsızlık’, ‘tutarsızlıklarına’, ‘tutarsızlıklar’, ‘tutarsızlıktır’, ‘tutarsız’) (‘huzursuzuz’, ‘huzursuzum’, ‘huzursuz’)	0.052

At this point, one of our hypotheses regarding the poor performance of clustering algorithms is the use of the first five letters of words. One of the rationales behind this decision stems from the analysis of frequently used terms within the configuration that forms the common groupings. Subsequently, the decision was made to forgo the process of reversing the data generation procedures and instead opt to take the first five letters of words instead. It is also noteworthy that emotional expressions retain their original semantic connotations, while the

dimensionality of the word vectors increases to 175,814. The previously mentioned search and *tf\_idf* calculation procedures were repeated using the newly obtained data.

Given the lack of significant deviation from the previous outcome, a novel quadrilateral testing framework was devised as a final endeavor to effectuate clustering via this methodology. In the present experimental framework, the merged clusters were disentangled and multiple associated values were observed for each fundamental component. As a result, approximately 2,000 rows were generated. Upon imposing the prerequisite that said emotions possess no less than 200 distinct dimensions, the resultant dataset comprises 205 unique emotions and a total of 175,814 columns namely D3. Furthermore, upon revisiting the original dataset and imposing the criterion of having a minimum of 213 emotions conveyed through at least 50 sentences, a dataset comprising 175,814 entries was generated from the initial 167 namely D1. Two additional datasets were created by eliminating columns with *tf\_idf* values lower than 5.3 from the original data. These two new datasets,  $205 \times 20010$  and  $167 \times 20010$ , each had 20,010 columns after the columns with *tf\_idf* values less than 5.3 were eliminated from the additional 175,814 of these two data namely D2 and D4. The study was resumed, encompassing a collective of four distinct datasets. The study was resumed, encompassing a collective of four distinct datasets. The datasets features can be seen in Table 4.6.

Table 4.6 The names and features of datasets.

	<b>Using all the words</b>	<b>Using words after dropping</b>
<b>Emotion list without suffix</b>	D1	D2
<b>Emotion list with suffix</b>	D3	D4

The identical cosine distance operations were employed on all four datasets. A Java-based interface has been developed to enhance the visibility of said operations. The present analysis consolidates sentences that incorporate both the selected emotions and common words, thereby synthesizing the binary phrases that generate shared vocabulary and the words that express the target affective states. The cosine distance values pertaining to the two chosen emotions can also be observed. The interface can be seen at Figure 4.5.



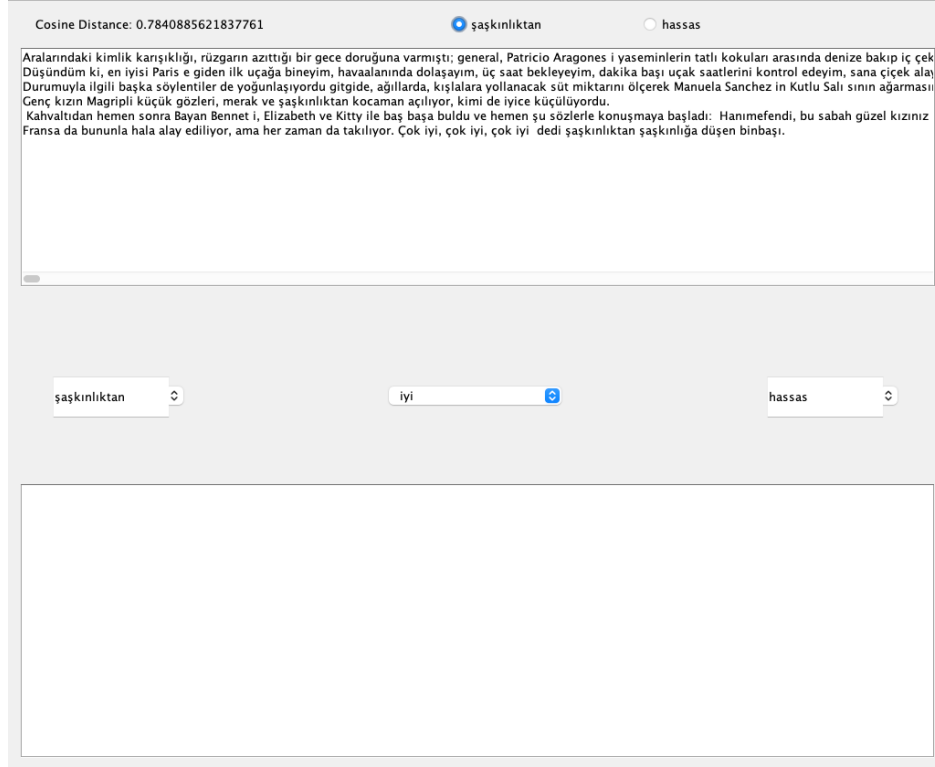


Figure 4.5 Interface for finding common words.

Upon scrutiny of the word review interface, it was observed that the absence of clustering was solely attributed to the excessive dimensionality of the data. The observation was made that emotions tend to cluster together based on shared vocabulary. However, despite the presence of common words, their weight is insufficient and the data's high dimensionality prevents effective clustering. At this juncture, it was determined that the extracted data was unsuitable for the creation of a lexicon or an emotion clustering.

#### 4.5. Embedding with Generated Data

Upon conducting a thorough analysis on alternative methods of data evaluation, it was discovered that the data in question could be effectively utilized for the purpose of word embedding. Word embedding refers to a numerical representation of a word that is utilized to extract various relationships associated with it [123].

Further, a decision was made to develop an interface based on the established cosine distance calculation algorithm while keeping the data constant. Through this interface, the inputted textual data was processed to extract the most proximate emotional states.

The purpose of the user interface, which can be seen in Figure 4.6, is so that tests can be performed in real time and experiments can be conducted using a variety of text data. The desired text data for emotion extraction is entered into the text area section. When the “Submit” button is pressed, it will generate it automatically on the back-end of the interface. The text data are segmented using words in accordance with the structure of the emotion-word vector. In a similar manner, these words have their own distinct term frequencies. This structure that was created as a result is scaled to its own. To determine the cosine distances between the scaled data and the emotion-word vectors, calculations are performed. Under the submit button is a list containing, according to the approximated distance values, the emotions that are closest to those with the closest distance. The scale is used to determine the biggest cosine distance value, which is then used to determine the threshold value. When the “Generate Chart” option is used with emotions that exceed the previously established threshold value, a pie chart of the ratios is displayed. The process can be summarized as Figure 4.7.

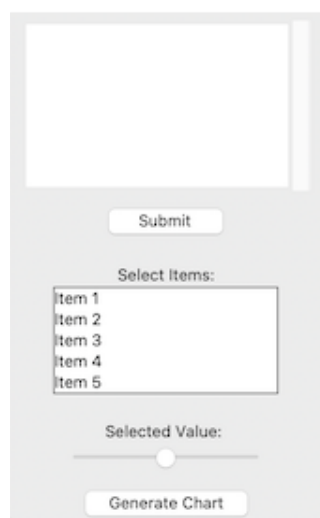


Figure 4.6 GUI process.

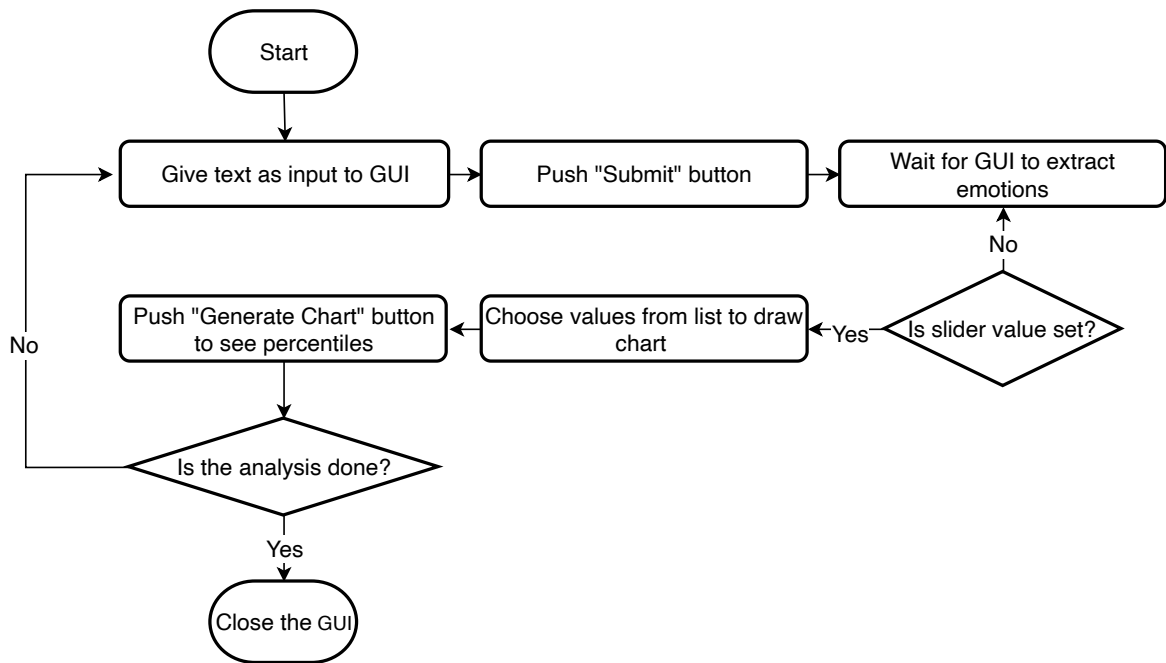


Figure 4.7 GUI process.

## 5. EXPERIMENTAL RESULTS

Following the development of the interface, an experiment was conducted to evaluate the effectiveness of the interface in paragraphs extracted from 41 different literary works. The paragraphs can be seen in Appendix-2. The experiment involved the application of the previously obtained D1, D2, D3, D4. As a result of the analysis of the test results, the ratios presented in the table below were obtained. Table 5.1 shows that the D2 dataset gives the most favourable results.

Table 5.1 The result of emotion list checking.

<b>Emotion List</b>	<b>Ratio</b>
List of root emotions	27/41
List of emotions with suffixes	8/41
Equal for both	6/41

After deciding on the list of emotions and the data to be used, which is D2, the scaling method was used again. Here, it was aimed to make a clear decision about row or column scale. For this purpose, using the same paragraphs, the scale methods to be created by taking column, row or maximum value from the data were tested. As a result, taking the maximum value gave the most successful result. The results are shown in Table 5.2.

Table 5.2 The result of scale methods.

<b>Scale Method</b>	<b>Ratio</b>
Column wise	9/41
Row wise	5/41
Maximum value	27/41

### 5.1. Performance Evaluation

The evaluation of the data and the model was carried out using an interface. Specifically, the interface was used to evaluate short passages taken from a total of 50 different novels.

During this experimental research, the interface was used as a tool for conducting a review with the ultimate goal of identifying the five closest emotions. Following the text analysis, a group of four people participated in a voting process to determine whether the emotions identified in the texts were adequately addressed. This voting process was facilitated by the use of a table format containing the texts and the emotions considered. These texts were also entered into ChatGPT, which was asked to provide 5 different emotions. The results of these experiments are presented in Table 5.3. The texts can be seen in the Appendix-3. The graphical representation of an entered paragraph is also observable in Figure 5.1.

Table 5.3: Results from the model and ChatGPT.

<b>Text</b>	<b>Emotions from Model</b>	<b>Emotions from ChatGPT</b>
T-1	sıkıntılı, panik, kin, vazgeçme, öfke	korku, şaşkınlık, endişe, panik, umutsuzluk
T-2	hüzün, merhamet, keder, durgun, küçümseme	hüzün, iç çekişme, kaygı, nostalji, umutsuzluk
T-3	histeri, aşk, arzu, hayran, beğenme	aşk, bağlılık, mutluluk, şükran, heyecan
T-4	şehvet, sadakat, alçakgönüllü, arzu, sevgi	aşk, hassasiyet, hüzün, özlem, içsel güzellik
T-5	keyifsiz, cesur, aciz, hüzün, minnet	kaybolmuşluk, özlem, aşkın öğrettikleri, hayatın kısa süresi, farklılıkların güzelliği
T-6	dikkatli, tereddüt, tedbir, kararsız, panik	korku, rahatlama, şüphe, içgüdü, farkındalık
T-7	zulüm, hüzün, sükunet, melankoli, kararsız	korku, endişe, güçsüzlük, umutsuzluk, iç savaş
T-8	hasret, sıkıntılı, aciz, samimi, pişman	nostalji, ilgi, hayal kırıklığı, kendini yitirme hissi, keder

Continued on next page

Table 5.3: Results from the model and ChatGPT. (Continued)

T-9	temkinli, vicdan azabı, aksi, şok, sabırsız	şefkat, merak, üzüntü, şok, endişe
T-10	hayret, şaşkın, duyarsız, dehşet, endişeli	utanç, hüznün, korku, endişe, kaygı
T-11	ümitsiz, gerilim, keder, cesur , huzursuz	çaresizlik, umutsuzluk, karmaşa, içsel çaba, belirsizlik
T-12	kırılgan, dengesiz, melankoli, hüznün, kararsız	duygusal soğukluk, içine işleme, geleceğin karanlık yüzü, tanıklık etme, duygu karmaşası
T-13	beklenti, iyimser, umutsuz, sabır , hayal kırıklığı	umutsuzluğun dönüşümü, umut, sevinç, fırsatları yakalama, gelecek umudu
T-14	güvensiz, dehşet, hoşnutsuz, tiksinti, umutsuz	sessizlik, korku, endişe, sıkıntı, tehdit
T-15	iyimser, şüphesiz, beğenme, özgür , güven	cesaret, kararlılık, macera, keşif, içsel dönüşüm
T-16	melankoli, zulüm, çaresiz, korkusuz, karamsar	umutsuzluk, kayıp, boşluk, içsel yolculuk, yön bulma, rehberlik
T-17	şehvet, neşe, mutlu, tutku, sevgi	sevgi, mutluluk, tutku, hüznün, anlam
T-18	uyumsuz, utangaç, bencil, şüphesiz, histeri	yalnızlık, boşluk, yabancılaşma, adımsama, anlam arayışı
T-19	sıkıntılı, empati, samimi, mutsuz, hayal kırıklığı	yıkıcılık, sarmalama, yükselme, keşfetme, yardımcı olma
T-20	ızdırap, empati, korku, melankoli, zulüm	karanlık, dehşet, ürperti, gerilim, sürükleme

Continued on next page

Table 5.3: Results from the model and ChatGPT. (Continued)

T-21	tereddüt, kabul, itaat, merak, hayret	merak, ilham, duygusallık, şaşkınlık, huzur
T-22	iyimser, keyif, mutlu, coşku, neşe	heyecan, hızlanan kalp ritmi, sıcaklık, tutku, bakışlar
T-23	beklenti, iyimser, umutsuz, hayal kırıklığı, sabır	umutsuzluk, umut ışığı, direnç, inanç, umut
T-24	melankoli, kaygısız, şehvet, aşk, tutku	cinsellik, tutku, yasak arzular, sarılma, karanlığın cazibesi
T-25	şehvet, arzu, keder, pişmanlık, kıskançlık	aşk, kıskançlık, kontrol, sınır, trajik
T-26	bıkkın, hasret, karamsar, korkusuz, dengesiz	melankoli, sessizlik, ezgi, hayatın anlamını sorgulama, kırılabilirlik
T-27	tutarsız, kararsız, azim, rekabet, dengesiz	aşk, gurur, çelişki, çözüm bulma, iç çatışma
T-28	beklenti, keder, umut, çaresiz, öfke	korku, isyan, yasak aşk, özgürlük, çelişki
T-29	dehşet, yenilgi, çılgın, öfke, huzursuzluk	koku, korku, savaş, adrenalin, ölüm
T-30	durgun, ümitsiz, dikkatli, uyumsuz, acıma	boşluk, yalnızlık, sorgulama, anlamsızlık, seslerin etkisi
T-31	tiksinti, çekince, durgun, şüphesiz, tutarsız	iğrençlik, mide bulantısı, zehirlenme, çekicilik, merak
T-32	beklenti, iyimser, umutsuz, hayal kırıklığı, sabır	umutsuzluk, çaresizlik, ağırlık, enerji, umut

Continued on next page

Table 5.3: Results from the model and ChatGPT. (Continued)

T-33	gerilim, tiksinti, utanç, aciz, sabır	adalet, ırkçılık, içsel çatışma, tanıklık etmek, mücadele
T-34	uyumsuz, keder, arzu, gerilim, huzursuz	melankoli, özlem, yalnızlık, akış, sarma hissi:
T-35	tiksinti, hoşnutsuz, güvensiz, melankoli, dehşet	umutsuzluk, korku, endişe, yıkım, gözlem
T-36	duyarsız, düşüncesiz, özlem, bencil, küçümseme	boşluk, uzaklaşma, anlam sorgulama, isyan, duygusuzluk
T-37	şok, gerilim, tiksinti, hayret, gergin	şaşkınlık, dehşet, yabancılaşma, iç çatışma, dönüşüm
T-38	kırılgan, suçluluk, çaresiz, gerilim, şüphe	ağırlık hissi, mücadele, kurtarma arzusu, karanlık, özgürlük
T-39	samimi, dostluk, sukünet, heves, sadakat	sevgi, mutluluk, destek, bağlılık, dayanışma
T-40	keder, zulüm, tedbirli, huzursuz, umut	acı, kalp kırıklığı, kararsızlık, umu, karmaşa
T-41	zulüm, hüznün, sukünet, melankoli, kararsız	karanlık, hayal kırıklığı, solgunluk, boğuculuk, umutsuzluk
T-42	şehvet, tutku, sadakat, arzu, anlaşılma	aşk, heyecan, sevdâ, anlam, yücelme
T-43	keder, özlem, beklenti, hayal kırıklığı, kırılgan	boşluk, sıkıntı, kayıp, huzursuzluk, içsel yolculuk
T-44	şükran, adaletsiz, ızdırap, aksi, pişman	pişmanlık, yük, izler, cesaret, yeniden başlama

Continued on next page



Table 5.3: Results from the model and ChatGPT. (Continued)

T-45	hasret, hayal kırıklığı, başarısız, özlem, hüznün	ihamet, yara, hüznün, inançsızlık, mücadele
T-46	ilham, takdir, hayırsever, iyimser, şükran	lham, kanatlar, büyü, hayal dünyası, yaratıcılık
T-47	ızdırap, karamsar, hoşnutsuz, histeri, şükran	acı, gözyaşları, keder, huzur, çaba
T-48	üzüntü, keder, tedirgin, pişman, heyecan	nostalji, kayıp, hatıralar, yüzleşme, inşa
T-49	tiksinti, bağımsız, hoşnutsuz, dehşet, öfke	korku, dehşet, hayatta kalma mücadelesi, gerilim, tehlike
T-50	aşağılama, aciz, düşüncesiz, duyarsız, korkusuz	hüznün, umutsuzluk, kabul, güçlü durma, sosyal bağlılık

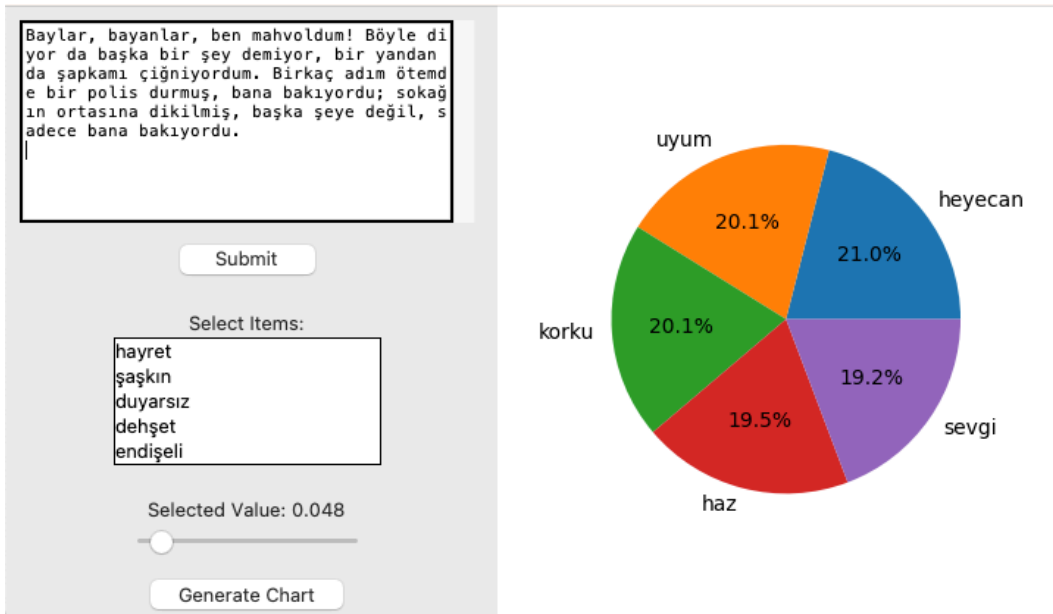


Figure 5.1 GUI output for the paragraph.

As can be seen in Table 5.3, some results from ChatGPT express behaviour rather than emotion. The shortness of the texts and the fact that almost none of the texts contain exactly

5 dominant emotions caused ChatGPT not to extract emotions completely. After a certain point, it extracted the words in the text as emotions.

Table 5.4 shows the votes given by the participants for the first 5 emotions for 50 different texts used in the test. It is seen that the totals of the votes, where the totals can be maximum 20, are minimum 9 and maximum 19 in these results. When this situation is expressed as a ratio, it is seen that a minimum of 45% and a maximum of 95% success is obtained from the model for 50 texts. The shortness of the texts and the high success rate in removing the emotion words in the text revealed the success of this model.

Table 5.4 The results from participants' votes for every text.

<b>Text No</b>	<b>1<sup>st</sup> emotion</b>	<b>2<sup>nd</sup> emotion</b>	<b>3<sup>rd</sup> emotion</b>	<b>4<sup>th</sup> emotion</b>	<b>5<sup>th</sup> emotion</b>
T-1	4	3	3	0	3
T-2	4	2	4	2	0
T-3	2	4	2	4	3
T-4	1	3	0	3	4
T-5	3	2	3	4	4
T-6	2	4	2	3	4
T-7	4	3	0	4	1
T-8	3	2	3	4	2
T-9	3	2	2	4	0
T-10	3	3	1	4	3
T-11	4	4	4	0	4
T-12	1	2	4	4	0
T-13	4	4	0	3	1
T-14	3	3	4	0	4
T-15	4	1	1	4	4
T-16	4	1	3	0	2
T-17	2	4	4	4	4
T-18	4	2	3	0	3

T-19	2	1	4	1	3
T-20	4	1	4	0	3
T-21	1	1	0	4	4
T-22	3	4	4	4	4
T-23	4	4	1	0	3
T-24	2	3	4	4	4
T-25	1	4	1	4	4
T-26	2	1	4	2	3
T-27	3	4	1	0	4
T-28	4	1	4	4	2
T-29	4	1	3	0	4
T-30	3	4	1	1	3
T-31	4	1	1	0	4
T-32	4	4	1	0	4
T-33	4	3	2	2	3
T-34	1	3	4	0	2
T-35	2	4	4	0	4
T-36	4	4	1	0	1
T-37	1	1	4	0	3
T-38	4	4	2	3	3
T-39	3	1	1	4	1
T-40	4	3	1	0	4
T-41	4	1	4	1	3
T-42	2	4	1	3	1
T-43	4	4	1	4	1
T-44	2	4	3	1	4
T-45	1	4	2	1	4
T-46	4	1	1	3	1

T-47	4	4	3	0	1
T-48	4	2	4	0	2
T-49	4	0	4	4	1
T-50	3	4	1	3	0

Table 5.5 presents the summary of participant number and count of agreed emotions for a total of 50 texts. In the first cell, two participants verified the existence of five distinct emotions within a set of nine texts. Two to three participants concurred that the text contains at least two emotions. The consensus among all participants regarding at least one emotional reaction to the test outcomes is evident in Table 5.3 and Figure 5.2. All emotions articulated in the text were collectively accepted by a minimum of two individuals. In this study, two participants approved two feelings within a total of six texts. Additionally, three emotions were accepted across 23 texts, four emotions within 12 texts, and five emotions within a total of nine texts.. Based on the analysis of emotions supported by three participants, it was found that, on average, 16 out of the 50 texts were accurately categorised with two emotions, 25 texts were categorised with three emotions, eight texts were categorised with four emotions, and one text was categorised with five emotions accurately. Given the shortness of these texts, it was observed that a remarkably favourable outcome was achieved, as the emergence of 4-5 emotions was not anticipated. If we assume that unanimous approval from all four individuals is necessary, it is also evident that these four individuals shared at least one common emotion, and there was no instance where they unanimously disapproved of any of 5 emotions. In essence, it can be observed that the provided text contains a minimum of one emotional expression based on the generated data. When this scenario is viewed in terms of the consensus reached by two individuals, it also yields a significantly higher success rate. In instances where two individuals express their approval, it is observed that a collective decision is reached based on a minimum of two emotions.

Table 5.5 Emotion to number of participants agreement.

	<b>2 participants</b>	<b>3 participants</b>	<b>4 participants</b>
<b>5 emotions</b>	9	1	0
<b>4 emotions</b>	12	8	3
<b>3 emotions</b>	23	25	8
<b>2 emotions</b>	6	16	22
<b>1 emotions</b>	0	0	17
<b>0 emotions</b>	0	0	0

### Correlation Between the Emotions and Participants' Votes

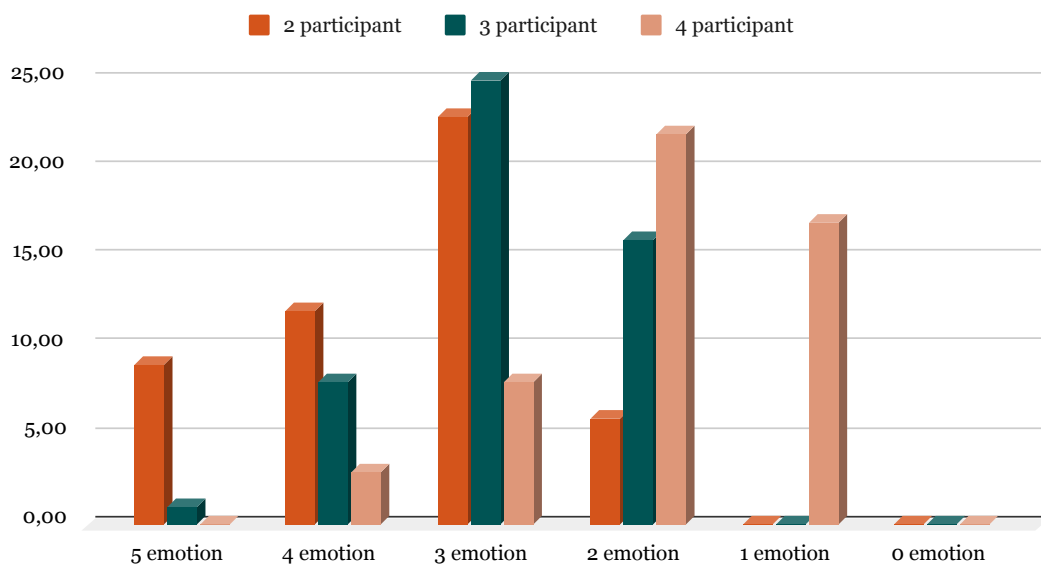


Figure 5.2 Graph showing the number of emotions agreed with the number of participants.

## 6. CONCLUSION

This thesis presents a comprehensive emotion analysis in Turkish through a series of research studies and experiments. The data collected from these endeavors were utilized to inform the development of emotion analysis techniques specific to the Turkish language. In the course of these experiments, it was deemed imperative to generate novel data that had not been previously employed. In order to achieve the intended objective, a total of 213 distinct Turkish emotions were taken into account and further steps were initiated. To generate this dataset, a sample of 100 distinct literary compositions was chosen, from which a cumulative sum of 660,000 sentences was extracted. A total of 175,814 distinct terms were identified from the sentences. A vector structure comprising emotion words was constructed using the extracted words and corresponding emotions. Based on the employed data structure, sentences were identified to represent various emotions.

Before starting data analysis, the decision was made to exclude any emotional content from a subset of sentences consisting of fewer than 50 instances. Consequently, a cumulative count of 167 emotions was gathered. Furthermore, it was determined that reducing the number of columns was necessary due to the potential issue of having 175,814 columns, which would lead to a significant amount of sparsity in the data. The TF-IDF method was employed for the purpose of conducting this reduction. After the completion of the operations, columns with a TF-IDF value lower than 5.3 are eliminated. Consequently, a matrix comprising 20,010 columns and 167 rows was constructed.

The present study has examined the possibility of clustering emotions based on the data generated. Multiple clustering methods have been employed for this purpose. Several algorithms, including K-Means, GAAC, and spectral, have been tested. The PCA algorithm was employed with caution during the process of selecting features. The silhouette score metric was employed to assess the cohesion of the clusters obtained through the clustering process. Nevertheless, the attempts made did not yield successful clustering.

Subsequently, the objective was to perform clustering on this dataset utilizing cosine distance by modifying both the data and the emotion. The objective is to merge emotions that are most similar to each other by examining the cosine distances between the two vectors representing the respective emotion-words. While initial iterations showed successful mergers, subsequent iterations revealed that the clusters coalesced into a singular cluster and exhibited a snowball effect.

After these experiments, a test interface for realtime testing was developed. The interface performs the following tasks: parsing the input test data, constructing and normalising a vocabulary dictionary, and computing the cosine distance between pre-existing emotion word vectors. Furthermore, the emotions that exhibited the highest proximity were determined based on the cosine distance values obtained from the interface. Afterwards, a visual representation in the form of a pie chart was generated to depict the distribution of these emotions.

In order to conduct cross validation for the study, the emotions extracted from 50 distinct texts were put to evaluation by a group of four individual participants. During the course of the study, it was noted that a collective of four participants arrived at a mutual agreement regarding the occurrence of at least one emotion from a set of five emotions that were derived from the texts sampled. This discovery indicates that the present emotion expressed in the text was accurately identified. Furthermore, the mentioned texts were inputted into ChatGPT with the objective of discerning the five prevailing emotions exhibited within the text. Both the model and ChatGPT have successfully detected prevalent emotions. Nevertheless, owing to the brevity of the texts, ChatGPT exhibited limitations in accurately discerning emotions. Consequently, it erroneously attributed the words within the sentence as indicative of an emotion, despite only detecting a few emotions.

In summary, the objective of this thesis was to construct a comprehensive lexicon encompassing a total of 167 distinct emotional states. This study represents the initial study in Turkish to encompass a diverse range of emotions. This study made significant advancements in emotion analysis for the Turkish language by creating a highly

comprehensive dataset that surpasses the existing TREMO dataset, which contains only six emotions.

There exists the potential to augment the quantity of books employed in subsequent research endeavours, as well as to incorporate a greater degree of emotional content within the gathered data. Furthermore, the utilisation of machine learning and deep learning algorithms can be employed for the analysis of this data. Moreover, the incorporation of bigram structure can be implemented during the process of data generation. Additionally, the attainment of a higher degree of specificity in the outcomes can be accomplished through the enlistment of supplementary participants for the experimental study.



## REFERENCES

- [1] Alex Thomas. *Natural Language Processing with Spark NLP: Learning to Understand Text at Scale*. O'Reilly Media, **2020**.
- [2] Dođan KÜÇÜK and Nursal ARICI. Dođal dil işlemede derin öğrenme uygulamaları üzerine bir literatür çalışması. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 2(2):76–86, **2018**.
- [3] Andrea Danyluk and Scott Buck. Artificial intelligence competencies for data science undergraduate curricula. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 9746–9747. **2019**.
- [4] Masato Hagiwara. *Real-world natural language processing: practical applications with deep learning*. Simon and Schuster, **2021**.
- [5] James Allen. *Natural language understanding*. Benjamin-Cummings Publishing Co., Inc., **1995**.
- [6] Noam Chomsky. Chomsky, n. 1959. a review of bf skinner's verbal behavior. *language*, 35 (1), 26–58., **1959**.
- [7] Ekaterina Kochmar. *Getting started with Natural Language Processing*. Manning, **2022**.
- [8] Fatih Sağlam. *Otomatik Duygu Sözlüğü Geliştirilmesi ve Haberlerin Duygu Analizi*. Ph.D. thesis, Fen Bilimleri Enstitüsü, **2019**.
- [9] Sadi Evren SEKER. Dođal dil işleme (natural language processing). *YBS Ansiklopedi*, 2(4):14–31, **2015**.
- [10] Cumali Türkmenođlu. *Türkçe metinlerde duygu analizi*. Ph.D. thesis, Fen Bilimleri Enstitüsü, **2016**.

- [11] Alebachew Chiche and Betselot Yitagesu. Part of speech tagging: a systematic review of deep learning and machine learning approaches. *Journal of Big Data*, 9(1):1–25, **2022**.
- [12] Dastan Hussen Maulud, Subhi RM Zeebaree, Karwan Jacksi, Mohammed A Mohammed Sadeeq, and Karzan Hussein Sharif. State of art for semantic analysis of natural language processing. *Qubahan Academic Journal*, 1(2):21–28, **2021**.
- [13] Diksha Khurana, Aditya Koli, Kiran Khatter, and Sukhdev Singh. Natural language processing: State of the art, current trends and challenges. *Multimedia tools and applications*, 82(3):3713–3744, **2023**.
- [14] Adrian Wong, Joseph M Plasek, Steven P Montecalvo, and Li Zhou. Natural language processing and its implications for the future of medication safety: a narrative review of recent advances and challenges. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 38(8):822–841, **2018**.
- [15] Sima Sharifirad and Stan Matwin. When a tweet is actually sexist. a more comprehensive classification of different online harassment categories and the challenges in nlp. *arXiv preprint arXiv:1902.10584*, **2019**.
- [16] Oliver Baclic, Matthew Tunis, Kelsey Young, Coraline Doan, Howard Swerdfeger, and Justin Schonfeld. Artificial intelligence in public health: Challenges and opportunities for public health made possible by advances in natural language processing. *Canada Communicable Disease Report*, 46(6):161, **2020**.
- [17] Anima Pradhan, Manas Ranjan Senapati, and Pradip Kumar Sahu. Comparative analysis of lexicon-based emotion recognition of text. In *Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021*, pages 671–677. Springer, **2023**.

- [18] Nourah Alswaidan and Mohamed El Bachir Menai. A survey of state-of-the-art approaches for emotion recognition in text. *Knowledge and Information Systems*, 62:2937–2987, **2020**.
- [19] Bing Liu. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167, **2012**.
- [20] Marouane Birjali, Mohammed Kasri, and Abderrahim Beni-Hssane. A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226:107134, **2021**.
- [21] Shiliang Sun, Chen Luo, and Junyu Chen. A review of natural language processing techniques for opinion mining systems. *Information fusion*, 36:10–25, **2017**.
- [22] Iti Chaturvedi, Erik Cambria, Roy E Welsch, and Francisco Herrera. Distinguishing between facts and opinions for sentiment analysis: Survey and challenges. *Information Fusion*, 44:65–77, **2018**.
- [23] Doaa Mohey El-Din Mohamed Hussein. A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences*, 30(4):330–338, **2018**.
- [24] Gizem Gezici and Berrin Yanıkoğlu. Sentiment analysis in turkish. *Turkish natural language processing*, pages 255–271, **2018**.
- [25] Nida Manzoor Hakak, Mohsin Mohd, Mahira Kirmani, and Mudasir Mohd. Emotion analysis: A survey. In *2017 international conference on computer, communications and electronics (COMPTELIX)*, pages 397–402. IEEE, **2017**.
- [26] L Jean Harrison-Walker. The effect of consumer emotions on outcome behaviors following service failure. *Journal of Services Marketing*, **2019**.
- [27] Saif M Mohammad et al. Tracking sentiment in mail: How genders differ on emotional axes. *arXiv preprint arXiv:1309.6347*, **2013**.

- [28] Paul Ekman, Tim Dalgleish, and Mick Power. Handbook of cognition and emotion. *Chichester, UK: Wiley*, **1999**.
- [29] Phoebe C Ellsworth. Appraisal theory: Old and new questions. *Emotion Review*, 5(2):125–131, **2013**.
- [30] Hatice Gunes and Maja Pantic. Automatic, dimensional and continuous emotion recognition. *International Journal of Synthetic Emotions (IJSE)*, 1(1):68–99, **2010**.
- [31] Bing Liu. Sentiment analysis: mining sentiments, opinions, and emotions. *Cambridge, UK: Cambridge University*, **2015**.
- [32] Samar Al-Saqqa, Heba Abdel-Nabi, and Arafat Awajan. A survey of textual emotion detection. In *2018 8th International Conference on Computer Science and Information Technology (CSIT)*, pages 136–142. IEEE, **2018**.
- [33] Jianhua Tao. Context based emotion detection from text input. In *Eighth International Conference on Spoken Language Processing*. **2004**.
- [34] Zhendong Dong and Qiang Dong. Hownet-a hybrid language and knowledge resource. In *International conference on natural language processing and knowledge engineering, 2003. Proceedings. 2003*, pages 820–824. IEEE, **2003**.
- [35] Sophia Yat Mei Lee, Ying Chen, and Chu-Ren Huang. A text-driven rule-based system for emotion cause detection. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, pages 45–53. **2010**.
- [36] Jianhua Zhang, Zhong Yin, Peng Chen, and Stefano Nichele. Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion*, 59:103–126, **2020**.
- [37] Diman Ghazi, Diana Inkpen, and Stan Szpakowicz. Hierarchical versus flat classification of emotions in text. In *Proceedings of the NAACL HLT 2010*

*workshop on computational approaches to analysis and generation of emotion in text*, pages 140–146. **2010**.

- [38] Waleed Ragheb, Jérôme Azé, Sandra Bringay, and Maximilien Servajean. Lirmm-advance at semeval-2019 task 3: attentive conversation modeling for emotion detection and classification. In *SemEval: Semantic Evaluation in NAACL-HLT*, pages 251–255. **2019**.
- [39] Merity Stephen, Xiong Caiming, Bradbury James, and Richard Socher. Pointer sentinel mixture models. *Proceedings of ICLR*, **2017**.
- [40] Ji Ho Park, Peng Xu, and Pascale Fung. Plusemo2vec at semeval-2018 task 1: Exploiting emotion knowledge from emoji and# hashtags. *arXiv preprint arXiv:1804.08280*, **2018**.
- [41] Robert Plutchik. *The psychology and biology of emotion*. HarperCollins College Publishers, **1994**.
- [42] Albert Mehrabian. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, 14:261–292, **1996**.
- [43] JA Russell. A circumplex model of affect journal of personality and social psychology 39. *161-178*, **1980**.
- [44] Sutjipto Arifin and Peter YK Cheung. Affective level video segmentation by utilizing the pleasure-arousal-dominance information. *IEEE Transactions on Multimedia*, 10(7):1325–1341, **2008**.
- [45] Hugo Liu, Henry Lieberman, and Ted Selker. A model of textual affect sensing using real-world knowledge. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 125–132. **2003**.

- [46] Carlo Strapparava and Rada Mihalcea. Semeval-2007 task 14: Affective text. In *Proceedings of the fourth international workshop on semantic evaluations (SemEval-2007)*, pages 70–74. **2007**.
- [47] Alastair J Gill, Robert M French, Darren Gergle, and Jon Oberlander. The language of emotion in short blog texts. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 299–302. **2008**.
- [48] Xin Kang and Fuji Ren. Understanding blog author’s emotions with hierarchical bayesian models. In *2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC)*, pages 1–6. IEEE, **2016**.
- [49] Mansur Alp Tocoglu and Adil Alpkocak. Tremo: A dataset for emotion analysis in turkish. *Journal of Information Science*, 44(6):848–860, **2018**.
- [50] Rosa Meo and Emilio Sulis. Processing affect in social media: A comparison of methods to distinguish emotions in tweets. *ACM Transactions on Internet Technology (TOIT)*, 17(1):1–25, **2017**.
- [51] Muhammad Abdul-Mageed and Lyle Ungar. Emonet: Fine-grained emotion detection with gated recurrent neural networks. In *Proceedings of the 55th annual meeting of the association for computational linguistics (volume 1: Long papers)*, pages 718–728. **2017**.
- [52] Klaus R Scherer and Harald G Wallbott. ” *Evidence for universality and cultural variation of differential emotion response patterning*”: Correction. American Psychological Association, **1994**.
- [53] Finn Årup Nielsen. afinn project. *DTU Compute Technical University of Denmark*, **2017**.
- [54] Cecilia Ovesdotter Alm, Dan Roth, and Richard Sproat. Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing*, pages 579–586. **2005**.

- [55] Saima Aman and Stan Szpakowicz. Identifying expressions of emotion in text. In *Text, Speech and Dialogue: 10th International Conference, TSD 2007, Pilsen, Czech Republic, September 3-7, 2007. Proceedings 10*, pages 196–205. Springer, **2007**.
- [56] Saif Mohammad, Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. Semeval-2018 task 1: Affect in tweets. In *Proceedings of the 12th international workshop on semantic evaluation*, pages 1–17. **2018**.
- [57] Ankush Chatterjee, Kedhar Nath Narahari, Meghana Joshi, and Puneet Agrawal. Semeval-2019 task 3: Emocontext contextual emotion detection in text. In *Proceedings of the 13th international workshop on semantic evaluation*, pages 39–48. **2019**.
- [58] Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. Compositionality principle in recognition of fine-grained emotions from text. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 3, pages 278–281. **2009**.
- [59] Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. @ am: Textual attitude analysis model. In *Proceedings of the naacl hlt 2010 workshop on computational approaches to analysis and generation of emotion in text*, pages 80–88. **2010**.
- [60] Soumaya Chaffar and Diana Inkpen. Using a heterogeneous dataset for emotion analysis in text. In *Advances in Artificial Intelligence: 24th Canadian Conference on Artificial Intelligence, Canadian AI 2011, St. John's, Canada, May 25-27, 2011. Proceedings 24*, pages 62–67. Springer, **2011**.
- [61] Efthymios Kouloumpis, Theresa Wilson, and Johanna Moore. Twitter sentiment analysis: The good the bad and the omg! In *Proceedings of the international AAAI conference on web and social media*, volume 5, pages 538–541. **2011**.
- [62] George A Miller. *WordNet: An electronic lexical database*. MIT press, **1998**.

- [63] Fabrizio Sebastiani and Andrea Esuli. Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th international conference on language resources and evaluation*, pages 417–422. European Language Resources Association (ELRA) Genoa, Italy, **2006**.
- [64] Raoul Biagioni. *The SenticNet sentiment lexicon: Exploring semantic richness in multi-word concepts*, volume 4. Springer, **2016**.
- [65] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of human language technology conference and conference on empirical methods in natural language processing*, pages 347–354. **2005**.
- [66] Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. **2004**.
- [67] Saif Mohammad and Peter Turney. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, pages 26–34. **2010**.
- [68] Saif Mohammad and Svetlana Kiritchenko. Wikiart emotions: An annotated dataset of emotions evoked by art. In *Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018)*. **2018**.
- [69] Saif Mohammad. Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 english words. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: Long papers)*, pages 174–184. **2018**.
- [70] Saif M Mohammad and Svetlana Kiritchenko. Using hashtags to capture fine emotion categories from tweets. *Computational Intelligence*, 31(2):301–326, **2015**.



- [71] Svetlana Kiritchenko, Xiaodan Zhu, and Saif M Mohammad. Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research*, 50:723–762, **2014**.
- [72] Jaime Redondo, Isabel Fraga, Isabel Padrón, and Montserrat Comesaña. The spanish adaptation of anew (affective norms for english words). *Behavior research methods*, 39(3):600–605, **2007**.
- [73] Hans Stadthagen-González, Pilar Ferré, Miguel A Pérez-Sánchez, Constance Imbault, and José Antonio Hinojosa. Norms for 10,491 spanish words for five discrete emotions: Happiness, disgust, anger, fear, and sadness. *Behavior research methods*, 50:1943–1952, **2018**.
- [74] José A Hinojosa, Natalia Martínez-García, Cristina Villalba-García, Uxía Fernández-Folgueiras, Alberto Sánchez-Carmona, Miguel A Pozo, and Pedro R Montoro. Affective norms of 875 spanish words for five discrete emotional categories and two emotional dimensions. *Behavior research methods*, 48:272–284, **2016**.
- [75] Hans Stadthagen-Gonzalez, Constance Imbault, Miguel A Pérez Sánchez, and Marc Brysbaert. Norms of valence and arousal for 14,031 spanish words. *Behavior research methods*, 49:111–123, **2017**.
- [76] Pilar Ferré, Marc Guasch, Natalia Martínez-García, Isabel Fraga, and José Antonio Hinojosa. Moved by words: Affective ratings for a set of 2,266 spanish words in five discrete emotion categories. *Behavior research methods*, 49:1082–1094, **2017**.
- [77] Melissa LH Võ, Markus Conrad, Lars Kuchinke, Karolina Urton, Markus J Hofmann, and Arthur M Jacobs. The berlin affective word list reloaded (bawl-r). *Behavior research methods*, 41(2):534–538, **2009**.
- [78] Philipp Kanske and Sonja A Kotz. Leipzig affective norms for german: A reliability study. *Behavior research methods*, 42:987–991, **2010**.

- [79] Benny B Briesemeister, Lars Kuchinke, and Arthur M Jacobs. Discrete emotion norms for nouns: Berlin affective word list (denn–bawl). *Behavior research methods*, 43:441–448, **2011**.
- [80] David S Schmidtke, Tobias Schröder, Arthur M Jacobs, and Markus Conrad. Angst: Affective norms for german sentiment terms, derived from the affective norms for english words. *Behavior research methods*, 46:1108–1118, **2014**.
- [81] Monika Riegel, Małgorzata Wierzba, Marek Wypych, Łukasz Żurawski, Katarzyna Jednoróg, Anna Grabowska, and Artur Marchewka. Nencki affective word list (nawl): the cultural adaptation of the berlin affective word list–reloaded (bawl-r) for polish. *Behavior Research Methods*, 47:1222–1236, **2015**.
- [82] Małgorzata Wierzba, Monika Riegel, Marek Wypych, Katarzyna Jednoróg, Paweł Turnau, Anna Grabowska, and Artur Marchewka. Basic emotions in the nencki affective word list (nawl be): New method of classifying emotional stimuli. *PLoS One*, 10(7):e0132305, **2015**.
- [83] Kamil K Imbir. Corrigendum: Affective norms for 4900 polish words reload (anpw\_r): Assessments for valence, arousal, dominance, origin, significance, concreteness, imageability and, age of acquisition. *Frontiers in Psychology*, 12:707540, **2021**.
- [84] Liang-Chih Yu, Lung-Hao Lee, Shuai Hao, Jin Wang, Yunchao He, Jun Hu, K Robert Lai, and Xuejie Zhang. Building chinese affective resources in valence-arousal dimensions. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 540–545. **2016**.
- [85] Zhao Yao, Jia Wu, Yanyan Zhang, and Zhenhong Wang. Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 chinese words. *Behavior research methods*, 49:1374–1385, **2017**.

- [86] Maria Montefinese, Ettore Ambrosini, Beth Fairfield, and Nicola Mammarella. The adaptation of the affective norms for english words (anew) for italian. *Behavior research methods*, 46:887–903, **2014**.
- [87] Ana Paula Soares, Montserrat Comesaña, Ana P Pinheiro, Alberto Simões, and Carla Sofia Frade. The adaptation of the affective norms for english words (anew) for european portuguese. *Behavior research methods*, 44:256–269, **2012**.
- [88] Agnes Moors, Jan De Houwer, Dirk Hermans, Sabine Wanmaker, Kevin Van Schie, Anne-Laura Van Harmelen, Maarten De Schryver, Jeffrey De Winne, and Marc Brysbaert. Norms of valence, arousal, dominance, and age of acquisition for 4,300 dutch words. *Behavior research methods*, 45:169–177, **2013**.
- [89] Agnes Sianipar, Pieter Van Groenestijn, and Ton Dijkstra. Affective meaning, concreteness, and subjective frequency norms for indonesian words. *Frontiers in psychology*, 7:1907, **2016**.
- [90] Elisavet Palogiannidi, Polychronis Koutsakis, Elias Iosif, and Alexandros Potamianos. Affective lexicon creation for the greek language. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2867–2872. **2016**.
- [91] Bojana Ćoso, Marc Guasch, Pilar Ferré, and José Antonio Hinojosa. Affective and concreteness norms for 3,022 croatian words. *Quarterly Journal of Experimental Psychology*, 72(9):2302–2312, **2019**.
- [92] A Gural Vural, B Barla Cambazoglu, Pinar Senkul, and Z Ozge Tokgoz. A framework for sentiment analysis in turkish: Application to polarity detection of movie reviews in turkish. In *Computer and information sciences III: 27th international symposium on computer and information sciences*, pages 437–445. Springer, **2013**.

- [93] Meriç Meral and Banu Diri. Twitter üzerinde duygu analizi. *IEEE*, 22:23–25, **2014**.
- [94] İslam Mayda and Çiğdem Aytekin. Sosyal medyada rekabet analizi için karşılaştırma görevine yönelik fikir madenciliği modeli. *JAMMO ISSN 2146-3328*, 7(27):414–425, **2013**.
- [95] Fırat Akba, Alaettin Uçan, Ebru Akcapinar Sezer, and Hayri Sever. Assessment of feature selection metrics for sentiment analyses: Turkish movie reviews. In *8th European Conference on Data Mining*, volume 191, pages 180–184. **2014**.
- [96] Mahmut Çetin and M Fatih Amasyalı. Supervised and traditional term weighting methods for sentiment analysis. In *2013 21st Signal Processing and Communications Applications Conference (SIU)*, pages 1–4. IEEE, **2013**.
- [97] Cüneyd Murad Özsert and Arzucan Özgür. Word polarity detection using a multilingual approach. In *Computational Linguistics and Intelligent Text Processing: 14th International Conference, CICLing 2013, Samos, Greece, March 24-30, 2013, Proceedings, Part II 14*, pages 75–82. Springer, **2013**.
- [98] Cumali Türkmenoglu and Ahmet Cüneyd Tantug. Sentiment analysis in turkish media. In *International Conference on Machine Learning (ICML)*. **2014**.
- [99] Alaattin Ucan. Automatic sentiment dictionary translation and using in sentiment analysis. *MSc, Hacettepe University, Ankara, Turkey*, **2014**.
- [100] ATLI Yakup and Nagehan İlhan. Duygu analizi için yeni bir sözlük; nayalex duygu sözlüğü. *Avrupa Bilim ve Teknoloji Dergisi*, (27):1050–1060, **2021**.
- [101] Mansur Alp Tocoglu and Adil Alpkocak. Emotion extraction from turkish text. In *2014 European Network Intelligence Conference*, pages 130–133. IEEE, **2014**.
- [102] Sinem Demirci. *Emotion analysis on Turkish tweets*. Master’s thesis, Middle East Technical University, **2015**.

- [103] Mansur Alp Toçoğlu and Adil Alpkocak. Lexicon-based emotion analysis in turkish. *Turkish Journal of Electrical Engineering and Computer Sciences*, 27(2):1213–1227, **2019**.
- [104] Fatih Sağlam, Hayri Sever, and Burkay Genç. Developing turkish sentiment lexicon for sentiment analysis using online news media. In *2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA)*, pages 1–5. IEEE, **2016**.
- [105] Uğur Mertoğlu and Burkay Genç. Lexicon generation for detecting fake news. *arXiv preprint arXiv:2010.11089*, **2020**.
- [106] Algirdas Laukaitis, Darius Plikynas, and Egidijus Ostasius. Sentence level alignment of digitized books parallel corpora. *Informatica*, 29(4):693–710, **2018**.
- [107] Belal Abdullah Hezam Murshed, Suresha Mallappa, Osamah AM Ghaleb, and Hasib Daowd Esmail Al-ariki. Efficient twitter data cleansing model for data analysis of the pandemic tweets. *Emerging Technologies During the Era of COVID-19 Pandemic*, pages 93–114, **2021**.
- [108] Akiko Aizawa. An information-theoretic perspective of tf–idf measures. *Information Processing & Management*, 39(1):45–65, **2003**. ISSN 0306-4573. doi:[https://doi.org/10.1016/S0306-4573\(02\)00021-3](https://doi.org/10.1016/S0306-4573(02)00021-3).
- [109] Fazli Can, Seyit Kocberber, Erman Balcik, Cihan Kaynak, H Cagdas Ocalan, and Onur M Vursavas. Information retrieval on turkish texts. *Journal of the American Society for Information Science and Technology*, 59(3):407–421, **2008**.
- [110] Muazzez Şule Karaşlar, Fatih Sağlam, and Burkay Genç. Tone density based sentiment lexicon for turkish. In *The International Conference on Artificial Intelligence and Applied Mathematics in Engineering*, pages 500–514. Springer, **2022**.

- [111] Lopamudra Dey, Sanjay Chakraborty, Anuraag Biswas, Beepa Bose, and Sweta Tiwari. Sentiment analysis of review datasets using naive bayes and k-nn classifier. *arXiv preprint arXiv:1610.09982*, **2016**.
- [112] Steven Bird. Nltk: the natural language toolkit. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, pages 69–72. **2006**.
- [113] Takio Kurita. Principal component analysis (pca). *Computer Vision: A Reference Guide*, pages 1–4, **2019**.
- [114] Hera Shaheen, Shikha Agarwal, and Prabhat Ranjan. Minmaxscaler binary pso for feature selection. In *First International Conference on Sustainable Technologies for Computational Intelligence: Proceedings of ICTSCI 2019*, pages 705–716. Springer, **2020**.
- [115] Murat Aslanyürek and Altan Mesut. Kümeleme performansını ölçmek için yeni bir yöntem ve metin kümeleme için değerlendirmesi. *Avrupa Bilim ve Teknoloji Dergisi*, (27):53–65, **2021**.
- [116] Emre Tolga Ayan, Rabia Arslan, Muhammed Said Zengin, Hacı Ali Duru, Sedat Salman, and Batuhan Bardak. Turkish keyphrase extraction from web pages with bert. In *2021 29th Signal Processing and Communications Applications Conference (SIU)*, pages 1–4. IEEE, **2021**.
- [117] Ketan Rajshekhkar Shahapure and Charles Nicholas. Cluster quality analysis using silhouette score. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*, pages 747–748. IEEE, **2020**.
- [118] Frank Nielsen. *Hierarchical Clustering*, pages 195–211. Springer International Publishing, Cham, **2016**. ISBN 978-3-319-21903-5. doi:10.1007/978-3-319-21903-5\_8.
- [119] Deepak Verma and Marina Meila. A comparison of spectral clustering algorithms. *University of Washington Tech Rep UWCSE030501*, 1:1–18, **2003**.

- [120] Javier Béjar Alonso. K-means vs mini batch k-means: a comparison. **2013**.
- [121] Vidya Setlur and Jock D. Mackinlay. Automatic generation of semantic icon encodings for visualizations. page 541–550, **2014**. doi:10.1145/2556288.2557408.
- [122] Faisal Rahutomo, Teruaki Kitasuka, and Masayoshi Aritsugi. Semantic cosine similarity. In *The 7th international student conference on advanced science and technology ICAST*, volume 4, page 1. **2012**.
- [123] Hande Aka Uymaz and Senem Kumova Metin. Emotion-enriched word embeddings for turkish. *Expert Systems with Applications*, 225:120011, **2023**.

## APPENDIXES

### Appendix-1: Literary Works Used in This Study

- Japon Masalları - Anonim
- Mutlu Ölüm - Albert Camus
- Sürgün ve Krallık - Albert Camus
- Veba - Albert Camus
- Binbir Hayalet - Alexandre Dumas
- Yabancı - Albert Camus
- Kamelyalı Kadın - Alexandre Dumas
- Monte Kristo Kontu - Alexandre Dumas
- Küçük Prens - Antoine de Saint-Exupery
- Empati - Adam Fawer
- Üç Silahşör - Alexandre Dumas
- Saatleri Ayarlama Enstitüsü - Ahmet Hamdi Tanpınar
- Olasılıksız - Adam Fawer
- Beatrice'den Sonra Birinci Yüzyıl - Amin Maalouf
- Afrikalı Leo - Amin Maalouf
- Doğunun Limanları - Amin Maalouf
- Ölü Ruhlar - Abbas Maroufi
- Işık Bahçeleri - Amin Maalouf
- Semerkant - Amin Maalouf
- Hayatın Kaynağı - Ayn Rand
- Tanios Kayası - Amin Maalouf
- Bizim Büyük Çaresizliğimiz - Barış Bıçakçı
- Günlerin Köpüğü - Boris Vian
- Peter Schlemihl'in Garip Hikayesi - Adelbert von Chamisso
- Amerikan Sapığı - Bret Easton Ellis



- Kaptan Yemeğe Çıktı ve Tayfalar Gemiyi Ele Geçirdi - Charles Bukowski
- Kasabanın En Güzel Kızı - Charles Bukowski
- Postane - Charles Bukowski
- Büyük Umutlar - Charles Dickens
- Sıradan Delilik Öyküleri - Charles Bukowski
- Müşterek Dostumuz - Charles Dickens
- İki Şehrin Hikâyesi - Charles Dickens
- Oliver Twist - Charles Dickens
- Dövüş Kulübü - Chuck Palahniuk
- Hayalet Yazılar - David Mitchell
- Körleşme - Elias Canetti
- Açlık Sanatçısı - Franz Kafka
- Dönüşüm - Franz Kafka
- Piyanist - Elfriede Jelinek
- Taşrada Düğün Hazırlıkları - Franz Kafka
- Mavi Oktav Defteri - Franz Kafka
- Huzursuzluğun Kitabı - Fernando Pessoa
- Sevgili Milena - Franz Kafka
- Benjamin Button'ın Tuhaf Hikayesi - F. Scott Fitzgerald
- Muhteşem Gatsby - F. Scott Fitzgerald
- Vadideki Zambak - Honore de Balzac
- Albaya Mektup Yazan Kimse Yok - Gabriel Garcia Marquez
- Aşk ve Öbür Cinler - Gabriel Garcia Marquez
- Başkan Babamızın Sonbaharı - Gabriel Garcia Marquez
- Evde Kalmış Kız - Honoré de Balzac
- Genç Werther'in Acıları - Goethe
- Ateşten Gömlek - Halide Edib Adıvar
- Sinekli Bakkal - Halide Edib Adıvar
- 1Q84 - Haruki Murakami
- Yolpalas Cinayeti - Halide Edib Adıvar

- Blbl ldrmek - Harper Lee
- İmknsızın Őarkısı - Haruki Murakami
- HaŐlanmış Harikalar Diyarı ve Dnyanın Sonu - Haruki Murakami
- Sınırın Gneyinde GneŐin Batısında - Haruki Murakami
- Yaban Koyununun İzinde - Haruki Murakami
- Sahilde Kafka - Haruki Murakami
- Gurur ve nyargı - Jane Austen
- Emma - Jane Austen
- Akıl ve Tutku - Jane Austen
- AŐk ve Gurur - Jane Austen
- Los Angeles Yolu - John Fante
- avdar Tarlasında ocuklar - Jerome David Salinger
- Byk Alık - John Fante
- Beyaz DiŐ - Jack London
- Sakın Yatađın Altına Bakma - Millas
- Őeker Portakalı - Jose Mauro de Vasconcelos
- Beni Asla Bırakma - Kazuo Ishiguro
- Krkl Vens - Leopold von Sacher-Masoch
- Uzak Tepeler - Kazuo Ishiguro
- Ademden nceki YaŐam - Margaret Atwood
- Damızlık Kızın yks - Margaret Atwood
-  rnek yk ve Bir nsz - Miguel De Unamuno
- GılgamıŐ Destanı - Muzaffer Ramazanođlu
- Sis - Miguel De Unamuno
- Tehlikeli oyunlar - Ođuz Atay
- Karısını Őapka Sanan Adam - Oliver Sacks
- Dorian Gray'in Portresi Oscar Wilde
- Btn Masallar, Btn ykler - Oscar Wilde
- Lord Arthur Savile'in Suu Oscar - Wilde
- Aldatmak - Paulo Coelho

- On Bir Dakika - Paulo Coelho
- Kazanan Yalnızdır - Paulo Coelho
- Simyacı - Paulo Coelho
- Zahir - Paulo Coelho
- Hayalet Şehir - Patrick McGrath
- Babilî düşünmek - Richard Brautigan
- Koku - Patrick Süskind
- Gülün Adı - Umberto Eco
- Kan Davası - Reşat Nuri Güntekin
- Alacakaranlık - Sâdık Hidâyet
- Foucault Sarkacı - Umberto Eco
- Ağaçkakan - Tom Robbins
- Önceki Günün Adası - Umberto Eco
- Otomatik Portakal - Anthony Burgess
- Prag Mezarlığı - Umberto Eco
- Doktor Hastalandı - Anthony Burgess
- Marifetler - Ursula K. Le Guin
- Sineklerin Tanrısı - William Golding
- Aylak Adam - Yusuf Atılgan
- Yaz Ortasında Ölüm - Yukio Mişima
- Fi - Akilah Azra Kohen

## Appendix-2: Paragraphs used for testing

- Odası kan kırmızısı boyalıydı. Kapı kilitliydi, penceresi kapalıydı ve içeri hiçbir ışık sızılmıyordu. Odanın ortasında yüksek bir yatak vardı, üzerinde siyah bir örtü vardı ve odanın tam ortasındaki küçük bir masa üzerinde mum ışığı yanıyordu. Bu odada öldüğüne inanılan kadın, diye düşündü John Mortimer. Birçok insan burada can vermiş olabilir. Neden bu odada ölmüş olmaları gerektiğini bilmiyorum ama kesin olan bir şey var, burada kötü bir şeyler olmuş. Birden, kapının kilidi kırıldı gibi bir ses duydu ve sonra bir adamın fısıltısı duyuldu. Seni burada tutacağım, dedi adam. Burada öleceksin. John kalbi hızla çarpmaya başladı ve odanın sırtına yaslandı. Kapı korkunç bir gıcırtyla açıldı ve içeri bir adam girdi. Adamın elinde bir bıçak vardı ve yüzünde karanlık bir ifade vardı. John geriye doğru çekildi ama duvarda bir şeyler hissetti. Daha fazla geriye gidemeyeceğini anladı. Adam yaklaştı ve bıçağı kaldırdı. Yardım! diye bağırdı John. Yardım edin! Adam bıçağı indirdi ama vücuduna değmedi. John, bıçağın bir santimetre kadar ötesinde, duvarda bir kolu olduğunu fark etti. O anda, kolu itti ve bir kapı açıldı. John dışarı çıktı ve nefes nefese kaldı. Kendini güvende hissetmek için koşarak uzaklaştı. Ancak, o kırmızı odadan kurtulduğunu düşünürken bile, kalbinin hızlı atışı durmamıştı.
- Yağmurun sesi camlara vuruyordu ve sokaklarda hiç kimse yoktu. Midori, kahvesini yudumlarırken dışarıdaki yağmura bakıyordu. Onun gibi pek çok insan da yağmuru seviyordu ama bugün, yağmurda güzel bir şey bulamıyordu. Tam tersine, yağmurun hüzünlü bir havası vardı. Midori, içindeki hüznle başa çıkmaya çalışıyordu. O, hayatında birçok kayıp yaşamıştı ve bu kayıpların izleri hala kalbini acıtıyordu. Yağmurlu günlerde ise kayıplarını daha fazla hissediyordu. Midori, kahvesini bitirdi ve pencereyi açtı. Soğuk hava içeriye dolarken, Midori kendini camın önüne doğru eğdi. Sokakta tek bir insan bile yoktu ve bu sessizlik onu daha da üzdü. Birkaç dakika sonra Midori pencereyi kapattı ve yavaşça yatağına doğru yürüdü. Yatağın içine girer girmez, hüznün içinde boğuldu. Gözlerinden yaşlar süzülüyordu ve hayatın ona ne kadar acımasız olduğunu düşünüyordu. Midori, hayatının hüzünlü olduğunu kabul etti.

Belki de, hayatta herkesin biraz hüzünlü olduğunu düşündü. Ama hüzünlü olmanın da bir anlamı vardı. Belki de, hüzünlü olmak insanları daha da güçlü kılıyordu. Midori, hayatı boyunca yaşadığı kayıpların izlerini hala taşısa da, bu hüznü onun hayatını kontrol etmesine izin vermeyecekti. Yağmurun hüzünlü havasına rağmen, Midori kararlıydı. Hayatını en iyi şekilde yaşamaya kararlıydı.

- Seni sevdiğimi söylemek istiyorum, ama kelimeler yetmiyor. Sana olan hislerimi ifade edebilecek kelime yok gibi geliyor. Kalbimdeki bu derin sevgiyi nasıl anlatabilirim ki sana? Adam, genç kadının ellerini tuttu ve gözlerindeki aşk dolu bakışlarını ona yansıttı. Kadın, bu anı ömrü boyunca unutamayacağını biliyordu. Adamın bakışları, içindeki sevgiye dair her şeyi söylüyordu. Seni sevmek, hayatımın anlamı oldu. Seni düşünmeden bir gün bile geçiremiyorum. Seninle her şey daha anlamlı, daha güzel oluyor. Seni sevmek, hayatın bütün güzelliklerini keşfetmek gibi bir şey. Kadın, adamın sözlerine gözlerindeki yaşlarla karşılık verdi. Onun da içindeki sevgi, aşk dolu hislerle doluydu. İkisi de birbirini sonsuza kadar seveceklerini biliyordu. Aşkın gücü, onları birbirine bağlıyordu. Dünyanın her köşesinde, her dilde anlatılan bu duygu, her zaman aynıydı. İkisi de bu güçlü duyguya teslim olmuştu ve hayatlarını birleştirmeye karar vermişlerdi.
- Mutluluğu aramak, hayatın en büyük amacıdır. Hayatta ne kadar başarılı olursanız olun, ne kadar mal varlığınız olursa olsun, eğer mutlu değilseniz, hayatınızın bir anlamı yoktur. Dalai Lama, öğrencilerine hayatın gerçek anlamının mutluluk olduğunu öğretiyordu. Ona göre, mutluluk sadece maddi zenginlikle veya başarıyla elde edilemezdi. Gerçek mutluluk, içimizdeki huzuru ve sevgiyi bulmakla mümkündü. Mutlu olmak için önce kendimizi tanımamız gerekir. Kendimizi tanıdığımızda, neye ihtiyacımız olduğunu daha iyi anlayabiliriz. Ve bu da bizi mutlu eder. Dalai Lama, mutluluğun herkesin hakkı olduğunu düşünüyordu. Ona göre, insanların mutluluğu bulabilmesi için öncelikle sevgi, saygı ve hoşgörü gibi değerleri benimsemeleri gerekiyordu. Mutlu olmak için başkalarına yardım etmek gerekir. Başkalarının acılarını hafifletmek, onlara destek olmak, onların mutluluğu için çalışmak, bizi de mutlu eder.

- Yalnızlık, içinde kaybolduđun bir denizdir. Kimileri bođulur bu denizde, kimileri ise yüzmeyi öğrenir. Ama hiçbirimiz tam olarak kurtulamayız bu denizden. Yalnızlıkla yaşamak zorunda kalırız ve bu bizim hayatımızın bir parçasıdır. Ancak yalnızlık, bizi kendimizle yüzleştirdiđi için bize çok şey öğretir. Kendimizi tanımamızı, iç dünyamızı keşfetmemizi sağlar. Belki de yalnızlık, hayatta en değerli öğretmenlerimizden biridir. Onunla barışık olmayı öğrenirsek, hayatın zorluklarına daha güçlü bir şekilde karşı koyabiliriz. Çünkü yalnızlık, bize iç gücümüzü keşfetme fırsatı verir.
- Aşk, insanın kalbinde doğan bir çiçektir ve özenle sulanması gereken bir bitkidir. Ama aşkın çiçeđi, bazen ayrılıkla solabilir. İnsanın sevdiđi kişiye duyduđu özlem, kalbindeki aşkın en büyük kanıtıdır. Bir insanın hayatındaki kayıplar, onun karakterini belirler. Hayatta karşılaşılan her zorluk, insanın güçlenmesine ve olgunlaşmasına yardımcı olur. Ama unutulmamalıdır ki, her kayıp bir fırsattır ve hayatın değerini anlamak için bir hatırlatmadır.
- Hayat, insanın içindeki tutkuları takip ettiđi sürece anlamlıdır. Bazen yaşamın zorlukları, tutkularımızı engelleyebilir ama yine de pes etmemeliyiz. Yaşamda, her şey mümkündür ve her şeyin bir çözümü vardır. Aşk, hayatın anlamını belirleyen en güçlü duygudur. Aşk, insanı mutlu eder ama aynı zamanda acı çekmesine de neden olabilir. Ama yine de, aşkın varlıđı hayata anlam katar. Aşk ne kadar zor olsa da, onun güzelliđini de inkar edemeyiz. İnsanların kalplerindeki en büyük güçlerden biri olan aşk, hayatımızda unutulmaz anılar bırakabilir. O anıların içinde en değerlileri, sevdiklerimizle paylaştığımız anılardır. Birini sevmek, onunla bir ömür geçirmek istemek, herkesin arayışında olduđu bir şeydir.
- Hayat, bir satranç tahtası gibidir. Her adımın sonucu, sonraki adımlarınızı belirler. Hayatta başarılı olmak için, öncelikle bir plana sahip olmalısınız. Ama planınızı deđiştirmek zorunda kalırsanız, esnek olmalısınız. Satranç, hayatın bir metaforudur. Satranç oynarken, rakibinizi takip eder ve onun hamlelerine karşılık verirsiniz. Hayatta da, başarılı olmak için başkalarını anlamak ve onların hareketlerine karşılık vermek önemlidir.

- Bir insanın yalnızlık hissi, ruhunun derinliklerinde var olan bir ıđlık gibidir. Ama yalnızlık, aynı zamanda bir insanın i dnyasını keřfetmesini ve kendisiyle barıřmasını sađlar. İnsanın iindeki farklı duygular, yalnız kaldıđında kendini gsterir.
- Bir insanın gemiři, onun geleceđini belirler. Ama gemiřte yapılan hataların telafisi iin insanın kendini affetmesi gerekir. Gemiřteki hatalar, insanın bugnn etkileyebilir ama geleceđini řekillendirmesine izin vermemeli. Sevgi, insanların birbirine verdiđi en deđerli hediyelerden biridir. Sevgi, bir insanın bařka bir insana hissettirdiđi en gzel duygudur. Ama sevginin gc, insanı bazen deli gibi hareket etmeye sevk edebilir.
- G, insanları deđiřtirir. Bir insanın gce sahip olması, onun karakterini ortaya ıkarır. Bazı insanlar gce sahip olduklarında, ktye kullanırlar. Ama bazı insanlar, glerini kullanarak dnya iin daha iyi bir yer yaratırlar. Adalet, insanların eřit olduđu bir dnya iin mcadele etmek demektir. Adalet, insanların haklarını savunmak, onları korumak ve onlara eřit fırsatlar sunmak demektir. Adalet, insanların birbirlerine karřı sorumluluklarını yerine getirmeleri demektir.
- Karanlık, insanın korkularını ve endiřelerini besler. Ama karanlık aynı zamanda insanın i dnyasına yolculuk yapmasını sađlar. Korkularımızla yzleřmek, kendimizi keřfetmek ve i dnyamızı keřfetmek iin karanlıkla yalnız kalmamız gerekir. Umut, insanın en deđerli varlıklarından biridir. Umut, insanın zorluklarla bařa ıkmasını, hayallerinin peřinden gitmesini ve hayatta kalmak iin mcadele etmesini sađlar. Umut, insanların iindeki umutsuzluđu yok eder ve onlara yeni bir bařlangı sađlar.
- Hayat, insanlar iin bir yolculuktur. Her yolculuk, insanın kendisini keřfetmesini ve hayatındaki amacını bulmasını sađlar. Hayatta ne kadar ilerlersek ilerleyelim, önemli olan yolda karřılařtıđımız engellerle nasıl bařa ıktıđımızdır. Kendine inanmak, insanın kendini keřfetmesi ve hayatındaki amacını bulması iin önemlidir. İnsanların hayatta bařarılı olması iin önce kendilerine inanmaları gerekir. Kendine inanmak, insanın zorluklarla bařa ıkmasını sađlar ve hayallerine ulařmasına yardımcı olur.

- Hayat, insanlar için bir maceradır. Macera, insanların hayatlarında karşılaştıkları zorluklarla başa çıkmaları, yeni şeyler keşfetmeleri ve kendilerini geliştirmeleri için gereklidir. Macera, insanların hayatlarına heyecan katar ve onları ileriye doğru itmeye yardımcı olur. Bir insanın hayatındaki en önemli şey, kendini keşfetmesidir. Kendini keşfetmek, insanın iç dünyasına yolculuk yapması, zayıflıklarını kabul etmesi ve kendisini geliştirmesi için gereklidir. Kendini keşfetmek, insanın hayatta başarılı olması ve mutlu olması için önemlidir.
- Hayatın anlamı, insanların hayatlarına kattığı anlamla belirlenir. Hayat, insanların kendilerini keşfetmeleri, hayallerinin peşinden gitmeleri ve başkalarına yardım etmeleri için bir fırsattır. Hayatın anlamı, insanların hayatta bir iz bırakması için mücadele etmeleri ve hayatlarına anlam katmaları için gereklidir. Hayatta ne kadar başarılı olduğumuz, insanların hayatlarına kattığımız değerle belirlenir.
- Sokaklar sessizdi. Yalnızca ayaklarının çıkardığı tıkırtıları duyuyordu. Etrafta kimse yoktu. Bir an önce evine gitmek istedi ama adımları yavaşladı. Kalbi hızla çarpıyordu. Ardında bir şeyin izini sürdüğü hissine kapıldı. Korku içinde geriye dönüp baktı ama hiçbir şey göremedi. Belki de sadece kafasını oynattığından böyle hissetmişti. Tekrar yürümeye başladı ama kalbi hala hızlı atıyordu. Eve ulaşır kapıyı açtığı anda rahat bir nefes aldı. Kapıyı kapattı ve bir an için gözlerini kapattı. Neden korktuğunu bilemiyordu ama içinde bir şeylerin farklı olduğunu hissetmişti.
- Güneş gözlerimi kamaştırıyordu ve uyandım. Başımın hafifçe döndüğünü hissettim, yatağın kenarındaki candan dışarıyı izledim. Sokakları ve evleri görebiliyordum ama insan yoktu, her yer sessiz ve boştu. Sonra kulağımda bir çıtırtı duydum, bir şeyler hareket ediyordu ama ne olduğunu tam olarak anlayamadım. Daha sonra fark ettim ki bu çıtırtıların kaynağı sadece benim nefes almamdan kaynaklanıyordu. İşte o an, kendimden korkmaya başladım.
- Gece havasının serinliği yanaklarımda hissediyordum. Yıldızların ışıkları, geceye gizemli bir hava katıyordu. Ayağa kalkıp, denizin kıyısındaki kayaların üzerine çıktım ve saatlerce suyun üzerine bakarak düşüncelere daldım. Ne kadar zaman geçtiğini



bilmiyorum ama beni geri getiren şey, ufukta beliren bir siluet oldu. Yaklaştıkça, o siluetin bir insan olduğunu fark ettim. Çok geçmeden, benimle konuşmak için yanıma geldi. Fakat bu kişi, sıradan bir insan değildi. Derisi siyahtı ve gözleri gökyüzü mavisiydi. Onunla konuşmaya başladıkça, içimdeki korku yerini merak bıraktı. Kimdi bu ve neden buradaydı?

- Birçok insan, hayatın anlamını arar ama aslında hayatın anlamı aranacak bir şey değildir, yaşanacak bir şeydir. Her insanın hayatı kendine özgüdür ve sadece kendi yolunu çizebilir. Bu yüzden, bir başkasının hayatını yaşamak yerine, kendimizin hayatını yaşamalıyız.
- Bazen insanlar, hayatlarının kontrolünü kaybettiklerini hissederler. Hayat, onları sürüklemeye başlar ve her şey o kadar hızlı gelişir ki, durdurmak veya geriye dönmek imkansız gibi görünür. Ancak, hayatın kontrolünü tekrar ele geçirmek mümkündür. Yapmanız gereken tek şey, durmak ve derin bir nefes alarak kendinizi yeniden odaklamak. Sonra, bir adım atın ve yeniden harekete geçin. Hayatın kontrolü size geri dönecektir.
- Bir insanın sevgisi, başka bir insanın hayatını değiştirebilir. Sevgi, güçlü bir silahtır ve onu kullanan kişi, büyük şeyler başarabilir. Aynı şekilde, sevgi eksikliği de hayatı zehir edebilir. Etrafımızdaki insanları sevmeli ve onlara değer vermeliyiz. Böylece, hayatımızda gerçek anlamda mutlu ve tatmin olmuş olabiliriz.
- Dünya, keşfedilmeyi bekleyen muhteşem bir yerdir. Doğanın güzelliği, insanların yarattığı sanat ve kültür, her yerde kendini gösterir. Ancak, dünyada yaşayan insanlar birbirlerini yok ediyorlar. Savaşlar, çatışmalar, nefret ve ayrımcılık, dünyayı zehirliyor. Buna karşın, insanlık olarak yapabileceğimiz en iyi şey, birbirimize sevgi ve saygı göstermek ve dünyayı daha iyi bir yer haline getirmektir.
- Gözlerimi kapattım ve o eski günlere gittim. O zamanlar, tüm dünya benim önümde açık gibi görünüyordu. Her şey mümkündü ve hiçbir şey imkansız değildi. Ama şimdi, yıllar geçtikçe, o umut dolu genç adam kaybolmuş gibi hissediyorum. Hayat beni

yıprattı ve umudumu kaybettirdi. Ama şimdi, bu eski anılar geri gelince, tekrar umutla doluyorum ve hayatın bana sunabileceklerini yeniden keşfetmeye hazır hissediyorum.

- Yaşamımın en zor zamanlarından birindeydim. Dünya beni terk etmişti ve yalnız kalmıştım. Ama o an, içimdeki gücü keşfettim ve kendi kendime şöyle dedim: Hayatta kalmak için mücadele etmek zorundasın. Her gün, bir sonraki adımı atmak için kendimi zorladım. Ve sonunda, zor zamanlar geride kaldı ve hayat benim için daha iyi hale geldi.
- Ve sonra, havanın keskin soğuşunda, Jane bir köpek yavrusu gördü. Yavru sırt üstü yatıyor, bacakları havada çırpınıyor ve inlemeler çıkarıyordu. Jane kedi köpek fark etmeksizin hayvanları sevmezdi, ama bu yavru görmezden gelemeyeceği kadar sevimliydi. Yanına yaklaşarak ona dokundu ve bir çığlıkla geri çekildi. Yavru her yerinden kanlıydı. Korkuyla onu yakalayıp arabaya getirdi. Yolda yavrunun ne kadar zayıf olduğunu fark etti ve hemen veterinerine götürdü. Veteriner, yavrunun aşırı açlık ve yaralar nedeniyle ölmek üzere olduğunu söyledi. Jane, yavruyu iyileştirmek için elinden geleni yapmaya karar verdi.
- Dünya, onları yok eden bir virüs tarafından ele geçirilmişti. İnsanlar, yalnızca kendilerine zarar veren bu hastalıktan korunmak için, sığınaklara kapatılmışlardı. Ancak burada, düzenli olarak uygulanan aşılama sayesinde insanlar yaşamaya devam edebiliyordu. Ama bu sığınaklar da güvenli değildi. Bir gün, sığınakların birinde aşılama bir hata yapıldı ve virüs sızdı. Kısa sürede herkes enfekte oldu ve birbirlerini öldürmeye başladılar. Sığınak, bir cehenneme dönmüştü. Ölü bedenler, kanlı duvarlar ve acı çeken insanlar her yerdedi. İnsanlığın sonu yakındı.
- Adeline, Kuzey Afrika'da bir çöl kasabasında kendi hayatını kurmaya çalışırken, İkinci Dünya Savaşı sırasında Fransızların işgali altındaki Tunus'ta yaşayan diğer kadınlarla da bağlantı kurar. Adeline, işgal altındaki Tunus'ta Fransız askerlerinin gözünden kaçmaya çalışırken, aralarında bir müttefik askeri olan Amerikalı pilot Kyle'ın da bulunduğu bir grup insanla tanışır. Kyle, Adeline'e aşık olur ve ona evlenme teklif eder, ancak Adeline kendini ona tam olarak açamaz. Adeline, içindeki duyguları

keşfetmek ve sevgisi hakkında daha fazla düşünmek için zaman isteyerek teklifi reddeder. Ancak Adeline, Kyle'ın teklifini reddettikten sonra bile, onunla olan bağı daha da güçlenir ve birlikte hayatta kalmak için mücadele etmeye devam ederler.

- Kafası karışık. Kendisine ne olup bittiğini anlayamadı. Aniden ortadan kaybolan, yerine bir yığın duman bırakan gizemli bir varlık mıydı yoksa sadece hayalinde mi gördü? Bu soruların cevaplarını aramaya çalıştı, ama daha fazla kafa karışıklığına sebep oldu. O gece hiçbir şekilde uyuyamadı ve yatağında kıvranıp durdu. Düşünceleri, hayalleri ve gerçekleri arasında kayboldu.
- Bir kadın kendini, hayatta tek yapabileceği şey olan, birini sevmeye adanmıştı. Ama o adam, tüm dünyanın bildiği bir insan, yalnızca bir hayranı olarak kalacaktı. O kadın, bir yandan aşkını yaşarken diğer yandan kendi kendine, onun için hazırlanmış bir senaryoda rol alıyor gibiydi. Kaderinin, aşkının tek sahibi olduğunu düşündüğünde, hayatının yanıldığını göstermek için onu sarsacak bir sınav hazırlar. Ve o gün, kendi kendine söylediği yalanların gerçeği örttüğü, kalbinin acısının bütün bedenini sardığı bir gün olur.
- Kapı kilitlendi, hapisane çıkmazı düşüncesi ona çaresizlik hissi veriyordu. İnanılmaz bir ağırlık yüklendi göğsüne, sanki nefes almaktan aciz kalmıştı. Sırtındaki omuzlarında dev bir yük taşıyormuş gibi hissetti. Kendisini aşağı çeken bir şey vardı, çaresizlik duygusu ağırlaştıkça ağırlaşıyordu. Dünyanın her yerinde zulüm, adaletsizlik ve şiddet vardı. Özgürlüğü elinden alınmış, bağımlılık yaratmış bir hapisane hayatı. Bütün bu düşünceler, karanlık ve karamsar bir ruh haliyle karışık bir şekilde içinde dolanıyordu.
- Köpekler aşkı dünyada temsil ediyor. Seni ölesiye sevdiklerini hissettiriyorlar, her zaman yanında olacaklarını söylüyorlar. Çoğu zaman köpekler insanlardan daha fazla aşkı hak ediyorlar. Onlar sana hiçbir beklentiyle yaklaşmazlar, sadece sevgi verirler. Sana ne kadar sevildiğini hissettirirler ve karşılığında sadece biraz yemek ve sevgi isterler.

- İki saat boyunca karşımda oturan adamın yüzünde belli belirsiz bir gülümseme belirdi ve sadece bir anlığına -ama bir anlamda kesinleşerek- bakışlarımız birbirine değdi. Sonra ağır ağır başını salladı. Ben de başımı salladım; fakat neye salladığımı bilmiyordum. Kendimizi bilinmeyen bir yolda, belirsiz bir geleceğe sürükleyen bu görünmez gücün ellerindeyiz, diye düşündüm ve biraz daha soğuk ter döktüm.
- Kütüphane neredeyse boştu, sadece garip birkaç kişi mevcuttu. Ama o garip kişilerin arasında, birden, belirgin bir şekilde farklı birisi olduğunu hissetti. Bu kişi, karanlık bir köşede, bir kitabı elinde tutarak sessizce okuyordu. Nora, bu kişinin kendisine bir şey söylemeden önce onun yanına gitmek istedi. Ancak yaklaştıkça, bu kişinin yüzünün, yüzü olmayan bir şeyin yerine geçtiğini fark etti.
- Kafası karışmıştı. Önce içinde yaşadığı dünyayı kaybetmişti, sonra da kendisini yeniden yaratmak için kendisine bir dünya inşa etmeye çalışıyordu. İnsanların, her şeyin ne kadar acımasız olduğunu anlamaları için bir sebepleri olmalıydı. Bu dünyada, herkes her şeyi yapabiliyordu, ama her şeyin bir bedeli vardı. İnsanlar, bu bedeli ödemekten korktukları sürece, dünyanın doğru işlenmesi mümkün değildi.
- Kendilerine yapılanları unutmak için birbirlerine sarıldılar. Tüm şehir, aynı korkuyu yaşıyordu. Ancak bu korku, yalnızca bir hastalık ya da salgın değildi. Bu, tüm insanların dünyasını sarsan bir durumdu. Hastalığın, insanlar üzerindeki etkileri, zamanla daha da kötüleşti. Birçok insan, hastalığın etkisiyle kör oldu. Tüm şehir, birbirine karışmış, kör insanlarla doluydu. Herkes, körleşmenin neden olduğu korkuyu yaşıyordu.
- Dorian, resmin birçok zaman kendisinden daha fazla yaşadığını düşündü; çünkü onun üzerindeki yaşam izlerini taşıyordu. Kendi gençliğinin ve güzelliğinin ölümsüzlüğünü elde edebileceğini düşündü. Her zaman genç kalacağım, ve resmim değişecek diye düşündü. Böylece yaşlanmak, hastalanmak ve ölmek onun için anlamsız hale geldi. Ama sonunda, yanlış yolda ilerlediğinin farkına varacaktı.
- Ölmekte olan adamın yüzünde şefkatli bir ifade vardı. Ona bu ifadeyi kazandıran, yatağında yatarken onu bekleyen ölümlüyle yüzleştiği için değildi. Asıl sebep, ölüme

rağmen kalanları düşündüğü içindi. Bu adamın sadece birkaç günü kalmıştı ve farklı insanlarla ilgili düşünceleri gidip geliyordu. Eski arkadaşlarından bazılarını düşündü. Bazılarını hatırlayamıyordu, diğerleri ise artık hayatta değildi. Aklına kızı da geldi. Kendisine çok benzeyen bir kızdı ve onun için endişeleniyordu. İki arasında hep bir uzaklık vardı ve bu uzaklığı kapatamamışlardı. Ama şimdi ölüm yaklaşınca, ona yakınlaşmak istiyordu. Belki de birbirlerini anlamaya başlayacaklardı.

- Jean Valjean, yalnız kalmak istiyordu. Kendisini açığa vurmamak istemedi. Fakat bu başarısız oldu. Sıkıntısı öylesine büyüktü ki, yanbaşındaki insanlar bile ondan etkilendi. Biraz daha zaman geçti ve nihayet bir gün, ağır ağır konuşmaya başladı. Konuşmaya başladıktan sonra birşeyler anlatmak istediği belli oluyordu. İnsanlar onu dinlediler, konuşmasının devamında gözyaşlarına boğuldu. Bu yüzden söylediklerini tam olarak anlayamadılar ama kendisini anlattığı anlaşılıyordu. Şöyle dedi: 'Sana çok şey borçluyum. Bu benim suçum, ama sen bana iyi davrandın. Beni kovmadın, bana ekmek verdin. Ben suçluyum, beni hapiste tutabilirsin. Ama hayatımı kurtardın. O yüzden ne istersen yaparım. Benden ne istersen iste, her şeyimi sana veririm. Sana teşekkür etmek istiyorum. Ne istersen yaparım, bana sadece hayatımı geri ver.'

### Appendix-3: Paragraphs used for testing with participants

- **T-1:** Neden bu odada ölmüş olmaları gerektiğini bilmiyorum ama kesin olan bir şey var, burada kötü bir şeyler olmuş. Birden, kapının kilidi kırıldı gibi bir ses duydu ve sonra bir adamın fısıltısı duyuldu. Seni burada tutacağım, dedi adam. Burada öleceksin.
- **T-2:** Onun gibi pek çok insan da yağmuru seviyordu ama bugün, yağmurda güzel bir şey bulamıyordu. Aksine, yağmurun hüzünlü havası vardı. İçindeki hüznle başa çıkmaya çalışıyordu. O, hayatında birçok kayıp yaşamıştı ve bu kayıpların izleri kalbini acıtıyordu.
- **T-3:** Seni sevmek, hayatımın anlamı oldu. Seni düşünmeden bir gün bile geçiremiyorum. Seninle her şey daha anlamlı, daha güzel oluyor. Seni sevmek, hayatın bütün güzelliklerini keşfetmek gibi bir şey.
- **T-4:** Aşk, insanın kalbinde doğan bir çiçektir ve özenle sulanması gereken bir bitkidir. Ama aşkın çiçeği, bazen ayrılıkla solabilir. İnsanın sevdiği kişiye duyduğu özlem, kalbindeki aşkın en büyük kanıtıdır.
- **T-5:** Kendimi yitik hissettim, kaybolmuş bir geminin kaptanı gibi. Ama biliyorum ki, aşkın bana öğrettiği her şeyi hatırlayacağım. Onun varlığı bana birçok şey öğretti. Dünyanın güzelliklerini, insanların farklılıklarını, hayatın kısa süresini.
- **T-6:** Tekrar yürümeye başladı ama kalbi hala hızlı atıyordu. Eve ulaşıp kapıyı açtığı anda rahat bir nefes aldı. Kapıyı kapattı ve bir an için gözlerini kapattı. Neden korktuğunu bilemiyordu ama içinde bir şeylerin farklı olduğunu hissetmişti.
- **T-7:** Her insanın hayatında bir karanlık taraf vardır. O taraf, kendini her fırsatta hatırlatır ve sizi yutmak istediğini söyler. Ne kadar direnirseniz direnin, o karanlık taraf sizi asla bırakmaz.

- **T-8:** Gözlerimi kapattım ve o eski günlere gittim. O zamanlar, tüm dünya benim önümde açık gibi görünüyordu. Ama şimdi, yıllar geçtikçe, o umut dolu genç adam kaybolmuş gibi hissediyorum. Hayat beni yıprattı ve umudumu kaybettirdi.
- **T-9:** Jane kedi köpek fark etmeksizin hayvanları sevmezdi, ama bu yavru görmezden gelemeyeceği kadar sevimliydi. Yanına yaklaşarak ona dokundu ve bir çığlıkla geri çekildi. Yavru her yerinden kanlıydı.
- **T-10:** Baylar, bayanlar, ben mahvoldum! Böyle diyor da başka bir şey demiyor, bir yandan da şapkamı çığniyordum. Birkaç adım ötemde bir polis durmuş, bana bakıyordu; sokağın ortasına dikilmiş, başka şeye değil, sadece bana bakıyordu.
- **T-11:** Karanlığın derinliklerine daldığımda, ruhumda bir çarpınış hissediyorum. Zihnimin karmaşası, içimdeki umutsuzluğun girdabına sürüklüyor beni.
- **T-12:** Duygusal soğukluk, her sayfada bizi sararak içimize işledi. Geleceğin karanlık yüzüne tanıklık ederken duygu karmaşasıyla boğuştuk.
- **T-13:** Kırık kanatlarımı umut tohumlarıyla onarıyorum. İçimde yeşeren umut, kalbimdeki çiçeklerin açmasına neden oluyor. Hayatın sunduğu fırsatları kucaklayarak, geleceğe dair umutla ilerliyorum.
- **T-14:** Gece yarısı sessizliği içinde, korkunun izleri sırtımda dolaşüyor. Korkuyla dolu bir düşünce, ruhumda endişenin karanlık gölgesini yaratıyor. Ayak izlerimi takip eden bir tehdit, içimdeki sıkıntıyı arttırıyor.
- **T-15:** İçimdeki cesaret yolculuğumda beni ileriye doğru iter. Zorlukların üstesinden gelme isteği, maceraya atılmamı sağlar. Kendimi keşfetme yolunda, içsel dönüşümün şafağına doğru ilerlerim.
- **T-16:** Denizin dalgalarıyla yüz yüze gelirken, umutsuzluğun adasında kaybolurum. Hayatta kaybettiklerim, içimde bir boşluk bırakır. İçsel yolculuğum, umutsuzluğun karanlık sularında bir pusula gibi yol gösterir.

- **T-17:** Birlikte geçirdiğimiz o anlar, yaşadığımız her duyguyla doluydu. Aşkımızın yoğunluğu, kalbimizi kavururken gözyaşlarımızla birleşti ve bizim için anlamını hiç yitirmeyen bir anıya dönüştü.
- **T-18:** Yalnızlık ve anlam arayışı, içimizi saran bir boşluğun hissiyle bizi ele geçirdi. Yabancılaşma, her satırda bir adım daha atmamıza sebep oldu.
- **T-19:** Aşkın yıkıcı etkileri, her bir cümlede bizi sarıp sarmaladı. İçimizde yükselen duygular, sevginin zorlu yollarını keşfetmemizi sağladı.
- **T-20:** Karanlık ve dehşet içimize işledi. Her cümlede tüylerimizi ürperten bir gerilim hissiyle bizi sürükledi.
- **T-21:** Bilgelik ve yaşam deneyimleri, her satırda bizi içine çekerek düşünmeye sevk etti. Hayatın anlamını sorgularken duygu dolu bir yolculuğa çıktık.
- **T-22:** İlk buluşmada yaşanan heyecanla birlikte, kalbin ritmi hızlanır. Gözlerdeki bakışlar, ellerin temasında hissedilen sıcaklık tutkunun habercisidir.
- **T-23:** Umutsuzluğun ortasında bile bir umut ışığı parlar. Zorlu zamanlarda insanın içindeki direnç, inanç ve umut, hayata tutunma gücünü verir. Geleceğe dair umutla, her yeni güne adım atarız.
- **T-24:** Cinsellik, tutku ve yasak arzularla dolu bir dünyanın kapılarını açar. Görülmemiş duygusal sınırları zorlarken, bedenler birbirine sıkı sıkıya sarılır. Aşk ve tutku arasındaki ince çizgide dolaşırken, karanlığın cazibesine kapılırız.
- **T-25:** Aşkın gölgesinde, kıskançlık alevlenir ve insanoğlunu kontrol etmek için her türlü sınırı zorlar. Bir anda içimize işleyen bu duygu, trajik olaylara sebep olabilir.
- **T-26:** Melankoli, sessizliğin derinliklerinde yankılanan bir ezgidir. İç dünyamızda yankılanan bu ezgi, hayatın anlamını sorgularken duygusal bir yoğunlukla bizi kuşatır. Bir piyano notası gibi, sessizliğin içinde kırılğan bir şekilde çalar.
- **T-27:** Fakat bu aşk, onun gururunu da zedeliyordu. Bu iki duygu arasında gidip gelirken bir yandan da çözüm bulmaya çalışıyordu.



- **T-28:** Hem korku hem de isyan dolu bir kalple, yasak aşkı ve özgürlük arzusunu içinde taşıyordu. İçindeki bu çelişkiler, onu gerçeği aramaya ve karanlığın içinde umut ışığı yakmaya yönlendiriyordu.
- **T-29:** Kan ve toprağın kokusu burun deliklerime doldu. İçimdeki korku, savaşın acı gerçekleriyle birleşti. İleriye doğru adım atarken kalbim hızla çarpmaya başladı, sanki ölümlle dans ediyormuşum gibi.
- **T-30:** Sahile doğru yürürken içimde bir boşluk hissi vardı. Dalgaların kıyıya vuran sesleri, yalnızlığımı daha da derinleştiriyordu. varoluşumu sorguluyor, anlamsızlığın içinde kayboluyordum.
- **T-31:** Burnuma dolan iğrenç kokular, midemi bulandırıyor. İnsanların üzerindeki çürümüşlük, tüm duyularımı zehirliyordu. Bu koku, beni bir şekilde çekiyordu ve meraka kapılıyordum.
- **T-32:** Umutsuzluk ve çaresizlik, kalbimde ağır bir yük gibi hissettirdi. Ancak, kahramanların inançlarının ışığında doğan umut damlacıkları, içime yeni bir enerji veriyordu. Umut nehrinin kıyısında ayakta durmaya çalışıyordum.
- **T-33:** Adaletin çığlıkları kulaklarımda yankılanıyordu. Irkçılığın kök saldıği bir toplumda, masumiyetin çiğnendiğine tanık olmak, içimde farklı duygular yaratıyordu. Adaletin zaferi için savaşmaktan başka çarem yoktu.
- **T-34:** Zamanın akışı, her şeyi sarıp sarmalayan bir melankoli yaratıyordu. Nesiller boyu süren yalnızlık, içimde bir özlem hissi uyandırıyor.
- **T-35:** İnsanlığın yıkımını gözlemek, içimde bir umutsuzluk fırtınası yaratıyordu. Gözlerimle gördüğüm manzaralar, geleceğe dair korku ve endişelerimi artırıyor.
- **T-36:** Duygusuz bir toplumun içinde, boşluk hissiyle boğuşuyordum. Mutluluk haplarıyla donatılmış bir dünyada, gerçek hislerimden uzaklaşıyor, yaşamın anlamını sorguluyordum. İnsanlık adına bir isyan ateşi içimde yanıp duruyordu.

- **T-37:** Gregor Samsa'nın bedeninde meydana gelen dönüşüm, içinde bir şaşkınlık ve dehşet yaratıyordu. Kendini yabancılaşmış bir dünyada bulan Gregor, iç çatışmalar yaşıyordu.
- **T-38:** Frodo Baggins'in içinde bir ağırlık hissediyordu. Taşdığı yüzük, dünyayı karanlığın pençesinden kurtarmak için verdiği mücadeleyi yansıtıyordu.
- **T-39:** Meg, Jo, Beth ve Amy'nin içinde kardeşlik bağına dair sevgi ve mutluluk vardı . Yaşadıkları zorluklara rağmen birbirlerine destek oluyorlardı.
- **T-40:** Gözyaşlarım yanağımı ıslatırken, içimdeki acıyı dindiremiyordum. Kalbim kırılmıştı, hayatımın en zor kararını vermek zorundaydım. İçimdeki karmaşayı düşünerek, geleceğim için bir umut ışığı aradım
- **T-41:** İçindeki karanlık bulutlar dağılmıyordu. Uzun bir süredir bir hayal bile kuramamıştı. Yüreğindeki umut ışığı giderek solmaya başlamış, onu boğan karanlık bir sis haline gelmişti.
- **T-42:** Aşkın büyüsunü, her hücresinde hissediyordu. Gözlerindeki parıltı, sevdasının izlerini taşıyordu. Bu güçlü duygu, hayatına anlam katıyordu, onu yüceltiyordu.
- **T-43:** Hep bir şeylerin eksik olduğunu hissediyordu. İçindeki boşluk, yüreğini sıkıyordu. Bir parçanın kayıp olduğunu biliyordu ve bu eksiklik, huzursuz bir ruh haline dönüşüyordu. Kayıp olanı aramak, içsel bir yolculuğa çıkmak zorundaydı.
- **T-44:** Yıllar geçtikçe, geçmişteki hataların yükü artıyordu. Pişmanlık duygusu, kalbinde derin izler bırakmıştı. Geçmişteki yanlış kararları düzeltmek yeni bir başlangıç yapma cesaretini buldu.
- **T-45:** Bir zamanlar güvendiği insanın ihanetine uğramıştı. Kalbindeki yara, derin bir hüzne dönüşmüştü. İhanetin acısı, onu bir kez daha inanmaktan alıkoyuyor, kırık bir güveni tamir etmek için bir mücadele veriyordu

- **T-46:** Kelimelerin ahenk içinde olması ona ilham veriyordu. Yaratıcılığının kanatlarına binerek düşlere yolculuk yapıyordu. Sözcüklerin büyüüne kapılıp, yeni dünyalar yaratıyordu.
- **T-47:** Gözyaşları, yüreğindeki acının bir ifadesiydi. Kedere gömülen kalbi, huzuru yeniden bulmak için çabalıyordu.
- **T-48:** Belleği, kaybolan anılarla doluydu. İçindeki geçmişin izleri, kaybolmaya yüz tutan hatıraları canlandırıyor. Kayıp anılarla yüzleşmek, kendi kimliğini yeniden inşa etmek için bir adımdı
- **T-49:** Gece yaratıklarının, kan emici varlıkların ve korkunç olayların mekanıdır. Kasaba halkı, yaşadıkları dehşetin farkına varırken, hayatta kalma mücadelesine girer.
- **T-50:** Yoksul, zavallı basit insanlardı. Şimdiye kadar bir çok derdi, kederi birlikte yaşamışlardı. O yüzden artık yakınmamayı da öğrenmişlerdi. Orada öylece durup, başlarını eğmiş, acıya katlanıyorlardı.