CHAPTER 11

Web Based Sentiment Analysis Application with Deep Learning

Yusuf UZUN
Necmettin Erbakan University, Türkiye

Yunus Emre GOKTEPE
Necmettin Erbakan University, Türkiye

Nurgul AYDIN
Necmettin Erbakan University, Türkiye

To Cite This Chapter:
Uzun, Y., Goktepe, Y.E. & Aydin, N. (2023).Web Based Sentiment Analysis Application with Deep Learning. In H. Arikan, Y. Uzun & I. Cetin (Eds.), *Current Studies in Basic Sciences, Engineering and Technology 2023* (pp. 103–109). ISRES Publishing..

Introduction

Sentiment analysis, also known as Opinion Mining, is vital in the field of natural language processing and studies in this field. Emotion analysis, in its most basic definition, is a field of study that deals with certain emotional states and allows these emotional states to be analyzed from texts, photographs, or videos. There are many studies on sentiment analysis in the literature. It is used in areas such as sales forecasts, product ratings, opinion polls, voting result predictions, finance, media, and stockbroking (Tokcaer, 2021). Its increasing importance day by day enables the development of sentiment analysis applications in connection with many fields. In this direction, a web-based sentiment analysis application will be developed.

With its application, it will allow sentiment analysis to be used in areas where it is not effective or has little effect today. For example, with such an application, the emotional state of the students during the lesson can be detected, and analyses such as happy or unhappy can be made. Similarly, employee satisfaction analysis can be performed in the work environment.

Nowadays, sentiment analysis is carried out, especially on social media platforms. For example, product comments, likes, or dislikes on videos, etc. However, it may happen that these emotional expressions made on social media do not reflect reality. This project aims to design a system that everyone can easily access. Nowadays, many homes have internet, and access to the internet has become easier. In this regard, designing a system with a web application is quite simple in terms of user access. However, the main goal is to ensure that sentiment analysis is used not only on social media but also in daily life. With a well-trained system, the accuracy of the analyses can be brought to a successful level. Thanks to this system, for example, it can determine where or in which areas the unhappiness rate is high and enable regulation efforts.

There is a lot of research on sentiment analysis. This research: It was created

by bringing together different techniques such as usage areas, research methods, and methods used.

In the article on face recognition and emotion analysis with deep learning, they implemented deep learning-based face recognition and facial expression recognition. Models developed with VGG-16, AlexNet, and ZFNet architectures were trained, and their success rates were compared. The most successful model was developed with the VGG-16 architecture as a reference, with a success rate of 92.03% (Safalı et al., 2021).

In the study of improving emotion classification performance using deep learning and down-sampling approaches, a data set consisting of 243 thousand user comments belonging to the Turkish e-commerce platform Hepsiburada company was used. In this unbalanced data set, deep learning algorithms were used to improve the classification performance, and an unbalanced data set approach was presented. With the presented approach, the false positive rate was improved from 69% to 90% and the accuracy value was improved from 95.5% to 99% (Santur, 2020).

In comparing different machine learning methods for sentiment analysis in Turkish texts, the classification results of different machine learning algorithms on the TREMO dataset were compared. Sentiment analysis was considered a text classification problem and four approaches were examined: ANN, SVM, RF, and KEYK algorithms. The categories provided by the data set, such as happiness, fear, anger, sadness, disgust, and surprise, were used as the emotion categories examined. According to the experimental study results, the ANN algorithm gave the best results. SVM, RF, and KEYK algorithms showed decreasing performance in this order (Toçoğlu et al., 2019).

In the article on sentiment analysis on Twitter, sentiment analysis was conducted on Twitter data, which is a popular social media site used by many users where users post their status updates in the form of "tweets". To perform this analysis, an intelligent model was created to classify tweets into positive and negative classes using machine learning methods such as Naïve Bayes and Support Vector Machine, and comparative results were given (İlhan et al., 2020).

Sentiment analysis was performed on Twitter datasets, which are widely used in the multi-population-based particle swarm optimization study for sentiment analysis. Emotions in the texts were classified as positive, negative, or ambiguous. Text mining preprocesses were applied to the datasets. In this study, a new optimization-based method was proposed for the classification process, and it was determined through experiments that the classification performance achieved with this method was more successful than the studies in the literature (Aydın et al., 2018).

In the sentiment analysis study using deep learning methods in Turkish texts, offensive language was detected on the OffensEval dataset. LSTM and BiLSTM networks were used as machine learning models in the study. The classification performances of these deep neural networks were evaluated as accuracy, sensitivity, precision, and F-score. The use of a large-sized corpus provided an improvement of over 40% in the F1 score (Yilmaz et al., 2021).

The sentiment analysis study of Twitter messages during the COVID-19 period, it was tried to measure the impact of the measures taken and the applications offered during the Coronavirus period on people using the sentiment analysis method. Important headlines about the coronavirus were published via Twitter. According to the sentiment analysis method, an overview was created by classifying positive and negative comments for the topics collected under 5 headings, and then it was analyzed whether there was a visible change in these topics in the weekly period. For approaches grouped as positive and negative on Twitter data, the logistic regression analysis method, which is accepted in the literature and can provide fast results, was used on approximately 2,000,000 tweets (Sarıman et al, 2020).

The above projects will be considered as examples of web-based sentiment analysis applications and these projects will be used as resources in some stages of the study.

In this study, there will be three different options for sentiment analysis. Analysis

will be made through text, photos, and videos. In the video analysis option, analysis will be made by selecting a video recording. The same method will be valid for emotion analysis with photography. In analysis with text, analysis will be performed according to the entered text. Artificial neural network models will be used for all three cases. Which ANN models will be used will be explained in detail in the later stages of the project.

Python language will be used to train sentiment analysis models. Since the project will be written as web-based, the server of the application will be prepared with Flask. The interface of the project will be implemented with Vue.js.

Web-Based Sentiment Analysis Application

Visual Studio Code platform will be used for web-based sentiment analysis applications. In the project, the training of artificial intelligence models will be carried out with Python. The FER2013 dataset will be used together with the CNN (Convolutional Neural Networks) model, which is one of the ANN models, for video, photo, and emotion analysis. In text sentiment analysis, the LSTM model, which is also one of the ANN models, and the data set called store reviews, consisting of store comments, will be used. Vue.js will be used for the interface and vuetify will be used for interface design. Finally, a server will be created with Flask to display the analyses on the interface using the trained models, and the connection between this server and the interface will be provided by axios.

FER2013 dataset

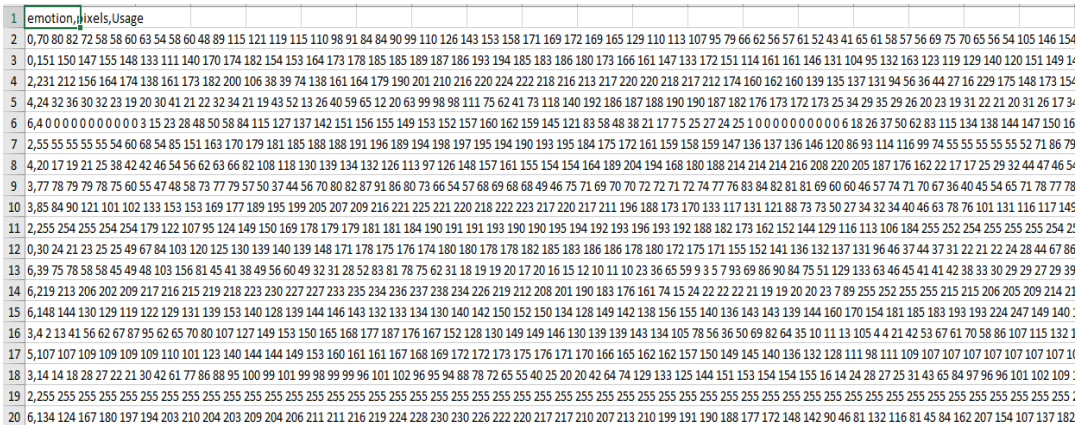
Fer2013 is an open-source dataset created by Pierre-Luc Carrier and Aaron Courville for an ongoing project, later shared for a public Kaggle competition shortly before ICML 2013. This dataset consists of 35,887 grayscale, 48x48 face images. 7 emotional states are labeled. The data set consists of images belonging to the relevant emotional state: 4593 angry (0), 547 hate (1), 5121 fear (2), 8989 happy (3), 6077 sad (4), 4002 confused (5) and 6198 normal (6). It was created (Fer2013, 2023). Sample images of the Fer2013 dataset are shown in Figure 1.

Figure 1.
Fer2013 Sample Images, (Fer2013, 2023)



In the shared data set, each line contains the BGR color codes of the images converted into pixels. The first column of each row indicates the emotion tagged to that image. Figure 2 shows the content of the dataset named Fer2013.

Figure 2.
Fer2013 Content, (Fer2013, 2023)

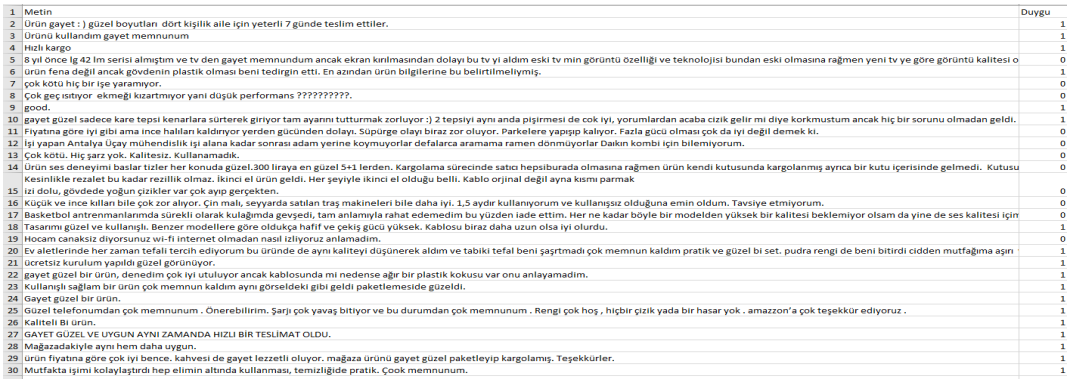


	emotion,pixels,Usage
2	0,70 80 82 72 58 58 60 63 54 58 60 48 89 115 121 119 115 110 98 91 84 84 90 99 110 126 143 153 158 171 169 172 169 165 129 110 113 107 95 79 66 62 56 57 61 52 43 41 65 61 58 57 56 69 75 70 65 56 54 105 146 154
3	0,151 150 147 155 148 133 111 140 170 174 182 154 153 164 173 178 185 185 189 187 186 193 194 185 183 186 180 173 166 161 147 133 172 151 114 161 161 146 131 104 95 132 163 123 119 129 140 120 151 149 14
4	2,231 212 156 164 174 138 161 173 182 200 106 38 39 74 138 161 164 179 190 201 210 216 220 224 222 218 216 213 217 220 220 218 217 212 174 160 162 160 139 135 137 131 94 56 36 44 27 16 229 175 148 173 154
5	4,24 32 36 30 32 23 19 20 30 41 21 22 32 34 21 19 43 52 13 26 40 59 65 12 20 63 99 98 98 111 75 62 41 73 118 140 192 186 187 188 190 190 187 182 176 173 172 173 25 34 29 35 29 26 20 23 19 31 22 21 20 31 26 17 34
6	6,4 0 0 0 0 0 0 0 0 0 0 0 0 3 15 23 28 48 50 58 84 115 127 137 142 151 156 155 149 153 152 157 160 162 159 145 121 83 58 48 38 21 17 7 5 25 27 24 25 1 0 0 0 0 0 0 0 0 0 0 6 18 26 37 50 62 83 115 134 138 144 147 150 16
7	2,55 55 55 55 55 54 60 68 54 85 151 163 170 179 181 185 188 188 191 196 189 194 198 197 195 194 190 193 195 184 175 172 161 159 158 159 147 136 137 136 146 120 86 93 114 116 99 74 55 55 55 55 55 52 71 86 79
8	4,20 17 19 21 25 38 42 42 46 54 56 62 63 66 82 108 118 130 139 134 132 126 113 97 126 148 157 161 155 154 154 164 189 204 194 168 180 188 214 214 214 216 208 220 205 187 176 162 22 17 17 25 29 32 44 47 46 5
9	3,77 78 79 79 78 75 60 55 47 48 58 73 77 79 57 50 37 44 56 70 80 82 87 91 86 80 73 66 54 57 68 69 68 68 49 46 75 71 69 70 70 72 72 71 72 74 77 76 83 84 82 81 81 69 60 60 46 57 74 71 70 67 36 40 45 54 65 71 78 77 78
10	3,85 84 90 121 101 102 133 153 153 169 177 189 195 199 205 207 209 216 221 225 221 220 218 222 223 217 220 217 211 196 188 173 170 133 117 131 121 88 73 73 50 27 34 32 34 40 46 63 78 76 101 131 116 117 145
11	2,255 254 255 254 254 179 122 107 95 124 149 150 169 178 179 179 181 181 184 190 191 191 193 190 190 195 194 192 193 196 193 192 188 182 173 162 152 144 129 116 113 106 184 255 252 254 255 255 255 254 25
12	0,30 24 21 23 25 25 49 67 84 103 120 125 130 139 140 139 148 171 178 175 176 174 180 180 178 178 182 185 183 186 186 178 180 172 175 171 155 152 141 136 132 137 131 96 46 37 44 37 31 22 21 22 24 28 44 67 86
13	6,39 75 78 58 58 45 49 48 103 156 81 45 41 38 49 56 60 49 32 31 28 52 83 81 78 75 62 31 18 19 19 20 17 20 16 15 12 10 11 10 23 36 65 59 9 3 5 7 93 69 86 90 84 75 51 129 133 63 46 45 41 41 42 38 33 30 29 29 27 29 39
14	6,219 213 206 202 209 217 216 215 219 218 223 230 227 227 233 235 234 236 237 238 234 226 219 212 208 201 190 183 176 161 74 15 24 22 22 22 21 19 19 20 20 23 7 89 255 252 255 255 215 215 206 205 209 214 21
15	6,148 144 130 129 119 122 129 131 139 153 140 128 139 144 146 143 132 133 134 130 140 142 150 152 150 134 128 149 142 138 156 155 140 136 143 143 139 144 160 170 154 181 185 183 193 193 224 247 149 140
16	3,4 2 13 41 56 62 67 87 95 62 65 70 80 107 127 149 153 150 165 168 177 187 176 167 152 128 130 149 149 146 130 139 139 143 134 105 78 56 36 50 69 82 64 35 10 11 13 105 4 4 21 42 53 67 61 70 58 86 107 115 132 1
17	5,107 107 109 109 109 109 110 101 123 140 144 144 149 153 160 161 161 167 168 169 172 172 173 175 176 171 170 166 165 162 162 157 150 149 145 140 136 132 128 111 98 111 109 107 107 107 107 107 107 10
18	3,14 14 18 28 27 22 21 30 42 61 77 86 88 95 100 99 101 99 98 99 99 96 101 102 96 95 94 88 78 72 65 55 40 25 20 20 42 64 74 129 133 125 144 151 153 154 154 155 16 14 24 28 27 25 31 43 65 84 97 96 96 101 102 109
19	2,255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255
20	6,134 124 167 180 197 194 203 210 204 203 209 204 206 211 211 216 219 224 228 230 230 226 222 220 217 217 210 207 213 210 199 191 190 188 177 172 148 142 90 46 81 132 116 81 45 84 162 207 154 107 137 182

Store reviews dataset

The data set was created by collecting user opinions from different online stores. Emotional states are divided into two labels: positive and negative. It contains 11622 pieces of data in total. Manual data entry was also done to diversify the data. A total of 27070 pieces of data were provided in the data set. Store review dataset content is shown in Figure 3.

Figure 3.
Store Comments Dataset Content, (Fer2013, 2023)



	Metin	Duygu
2	Ürün gayet :) güzel boyutları dört kişilik aile için yeterli 7 günde teslim ettiler.	1
3	Ürünü kullandım gayet memnunum	1
4	Hızlı kargo	1
5	8 yıl önce lg 42 lm serisi almıştım ve tv den gayet memnundum ancak ekran kırılmasından dolayı bu tv yi aldım eski tv min görüntü özelliği ve teknolojisi bundan eski olmasına rağmen yeni tv ye göre görüntü kalitesi o	0
6	ürün fena değil ancak gövdenin plastik olması beni tedirgin etti. En azından ürün bilgilerine bu belirtilmeliymiş.	1
7	çok kötü hiç bir işe yaramıyor.	0
8	Çok geç ısıtıyor ekmeği kızartmıyor yani düşük performans ??????????.	0
9	good.	1
10	gayet güzel sadece kare tepsi kenarlara sürterek giriyor tam ayarını tutturmak zorluyor :) 2 tepsiyi aynı anda pişirmesi de cok iyi, yorumlardan acaba cizik gelir mi diye korkmustum ancak hiç bir sorunu olmadan geldi.	1
11	Fiyatına göre iyi gibi ama ince halıları kaldırıyor yerden gücünden dolayı. Süpürge olayı biraz zor oluyor. Parkelere yapışıp kalıyor. Fazla gücü olması çok da iyi değil demek ki.	0
12	İşi yapan Antalya Üçay mühendislik işi alana kadar sonrası adam yerine koymuyorlar defalarca aramama ramen dönmüyorlar Daıkın kombi için bilemiyorum.	0
13	Çok kötü. Hiç şarz yok. Kalitesiz. Kullanamadık.	0
14	Ürün ses deneyimi baslar tizler her konuda güzel.300 liraya en güzel 5+1 lerden. Kargolama sürecinde satıcı hepsiburada olmasına rağmen ürün kendi kutusunda kargolanmış ayrıca bir kutu içerisinde gelmedi. Kutusu	0
	Kesinlikle rezalet bu kadar rezillik olmaz. İkinci el ürün geldi. Her şeyiyle ikinci el olduğu belli. Kablo orjinal değil ayna kısmı parmak	
15	izi dolu, gövdede yoğun çizikler var çok ayıp gerçekten.	0
16	Küçük ve ince kılları bile çok zor alıyor. Çin malı, seyyarda satılan traş makineleri bile daha iyi. 1,5 aydır kullanıyorum ve kullanışsız olduğuna emin oldum. Tavsiye etmiyorum.	0
17	Basketbol antrenmanlarımda sürekli olarak kulağımda gevşedi, tam anlamıyla rahat edemedim bu yüzden iade ettim. Her ne kadar böyle bir modelden yüksek bir kalitesi beklemiyor olsam da yine de ses kalitesi için	0
18	Tasarımı güzel ve kullanışlı. Benzer modellere göre oldukça hafif ve çekiş gücü yüksek. Kablosu biraz daha uzun olsa iyi olurdu.	1
19	Hocam canaksiz diyorsunuz wi-fi internet olmadan nasıl izliyoruz anlamadım.	0
20	Ev aletlerinde her zaman tefali tercih ediyorum bu üründe de aynı kaliteyi düşünerek aldım ve tabiki tefal beni şaşırtmadı çok memnun kaldım pratik ve güzel bi set. pudra rengi de beni bitirdi cidden mutfağıma aşırı	1
21	ücretsiz kurulum yapıldı güzel görünüyor.	1
22	gayet güzel bir ürün, denedim çok iyi utuluyor ancak kablosunda mi nedense ağır bir plastik kokusu var onu anlayamadım.	1
23	Kullanışlı sağlam bir ürün çok memnun kaldım aynı görseldeki gibi geldi paketlemeside güzeldi.	1
24	Gayet güzel bir ürün.	1
25	Güzel telefonumdan çok memnunum . Önerebilirim. Şarjı çok yavaş bitiyor ve bu durumdan çok memnunum . Rengi çok hoş , hiçbir çizik yada bir hasar yok . amazzon'a çok teşekkür ediyoruz .	1
26	Kaliteli Bi ürün.	1
27	GAYET GÜZEL VE UYGUN AYNI ZAMANDA HIZLI BİR TESLİMAT OLDU.	1
28	Mağazadakiyle aynı hem daha uygun.	1
29	ürün fiyatına göre çok iyi bence. kahvesi de gayet lezzetli oluyor. mağaza ürünü gayet güzel paketleyip kargolamış. Teşekkürler.	1
30	Mutfakta işimi kolaylaştırdı hep elimin altında kullanması, temizliğide pratik. Çook memnunum.	1

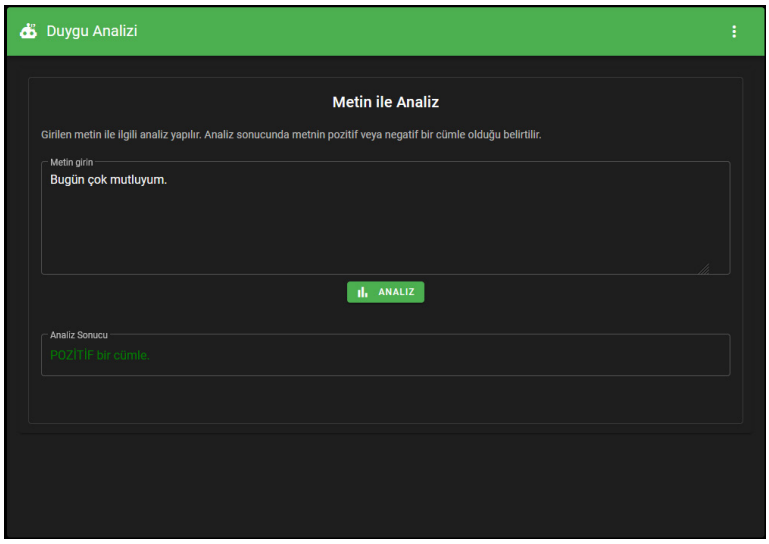
Results and Discussion

In the web-based sentiment analysis project, the data was primarily trained using relevant artificial intelligence models. The store reviews dataset trained with LSTM gave an accuracy rate of 98%, and the FER2013 dataset trained with CNN gave an accuracy rate of 68%. Then, the interface of the project was created using vue and vuetify. Requests were sent from Vue to the Flask API via axios. Flask is written in Python using trained models. When a text analysis is performed in the project interface, it takes up to two seconds for the analysis result to appear on the screen. In photo analysis, this time varies between three and four seconds. However, video analysis takes a little longer, reaching an average of fifteen seconds. My goal with these findings is to provide a basic understanding to evaluate and improve the performance of my sentiment analysis project.

As a result of the findings, some important situations arise. While the data trained with LSTM gives a very successful percentage, it is seen that the data trained with CNN does not give such a successful percentage. In their study titled Comparison of different machine learning methods for sentiment analysis in Turkish texts, it is explained that

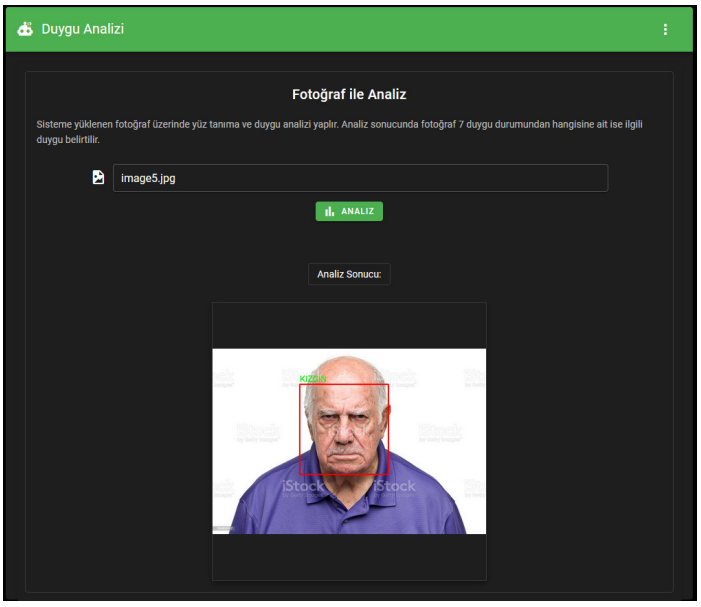
different algorithms can also change the accuracy rates (Toçoğlu et al., 2019). The low success rate at CNN may be due to such a problem. Another situation is that the analysis result is quickly reflected on the interface in text data, but the speed decreases accordingly in photo and video analysis. This may be due to the structure and size of the data. These situations that I obtained because of my findings may be important discussion topics to further improve the success of the project. The screenshot of sentiment analysis with text is shown in Figure 4.

Figure 4.
Sentiment Analysis Example with Text



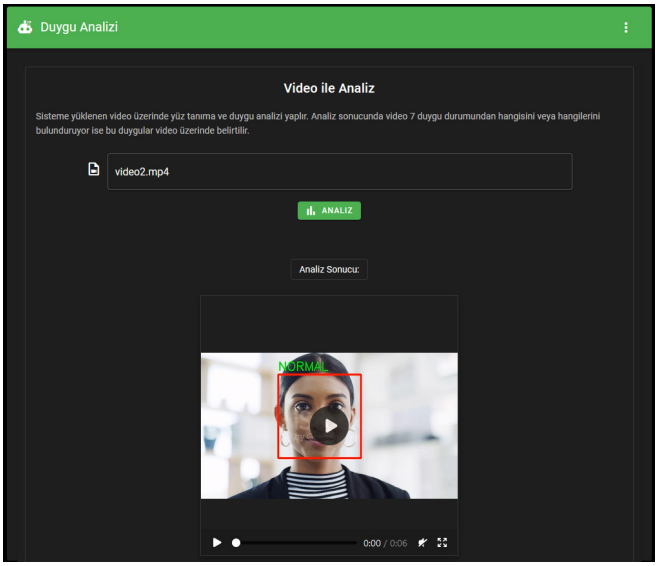
In the text and sentiment analysis section of the project, the text entered is analyzed. As a result of the analysis, if the text is a positive sentence, it is given a positive sentence; if it is a negative sentence, the analysis result is given as a negative sentence. The screenshot of the sentiment analysis with the image process is shown in Figure 5.

Figure 5.
Sentiment Analysis Example with Image



In the emotion analysis with photo section, face recognition, and emotion analysis are performed on the photo selected from the folder. The analyzed photograph is given as the analysis result. It is also saved as a new photo. The screenshot of the video sentiment analysis application is shown in Figure 6.

Figure 6.
Sentiment Analysis Example with Video



In the video emotion analysis section, face recognition and emotion analysis are performed on the video selected from the folder. The analyzed video is given as the analysis result. It is also saved as a new video.

Conclusions

In this study, it was aimed and carried out to perform sentiment analysis with the web interface. The results of the sentiment analysis in the project were largely successful in reflecting reality. In particular, text and photo analysis give better results than video analysis. This may be because the accuracy rate of the model trained with CNN is lower than the accuracy rate of the model trained with LSTM. The analysis result of the text entered from the interface is quickly obtained. The result of analysis with photographs is created in a short time. However, some time must be waited for the result of the video analysis. In emotional analysis with videos and photographs, happy, normal, angry, and sad expressions are easily perceived. In text analysis, an output of around 0.01 is produced for a negative sentence and around 0.991 for a positive sentence. These results reveal the success of the project. However, it shows that it can be taken further.

This study, like every other study, can be developed in line with different thoughts and opinions. However, performance is more important than design and is the main goal. When thought in this direction, the accuracy rate of data trained with CNN can be increased from 68% to around 90%. A different model or a different data set may be used. The data set can be diversified to make each statement easier to perceive. The number of trainings can be increased. Many other solutions are also applicable. For the same purpose, the waiting time for the result of video analysis can be reduced to five seconds. The data received from the user is processed on the flask side and waiting occurs in this part. However, this is due to the size of the data. Data processing time can be reduced with different methods. This study can be an important example of web-based sentiment analysis.

References

Aydin, I., Salur, M., U., Baskaya, F. (2018). Multi-Population Based Particle Swarm Optimization for Sentiment Analysis. *Turkish Informatics Foundation Journal of Computer Science and Engineering,* 11(1), 52-64.

Fer2013. (2023, 11 November), *Fer2013 Dataset Open Source*. https://www.kaggle.com/datasets/msambare/fer2013

Ilhan, N. ve Sagaltici, D. (2020). Sentiment Analysis on Twitter. *HRU Engineering Journal*, 5(2), 146-156.

Safali, Y. and Avaroglu, E. (2021), Facial Recognition and Emotion Analysis with Deep Learning. *European Journal of Science and Technology*, (31), 764-770.

Santur, Y. (2020). Improving Emotion Classification Performance Using Deep Learning and Downsampling Approaches. *Fırat University Journal of Engineering Informatics*, 32(2), 561-570.

Sariman, G. ve Mutaf, E. (2020). Sentiment Analysis of Twitter Messages During Covid-19. *Euroasia Journal of Mathematics, Engineering*, Natural & Medical Sciences International Indexed & Refereed. ISSN: 2667-6702, 7(10), 137-148.

Tokcaer, S. (2021). Sentiment Analysis in Turkish Texts. *Yaşar University E-Journal*, 16 (63), 1514-1534. DOI: 10.19168/jyasar.928843

Tocoglu, M., A., Celikten, A., Aygun, I., Alpkocak, A. (2019). Comparison of Different Machine Learning Methods for Sentiment Analysis in Turkish Texts. *DEUFMD*, 21(63), 719-725.

Yilmaz, S.S., Ozer, I., Gokcen, H. (2021). Sentiment Analysis Using Deep Learning Methods in Turkish Texts. *International Symposium of Scientific Research and Innovative Studies*., 971-982.

Authors

Yusuf UZUN, PhD, is an Assistant Professor of Computer Engineering at Necmettin Erbakan University in Konya, Turkey. He holds a PhD in Mechanical Engineering from Necmettin Erbakan University. His main areas of interest are artificial intelligence, autonomous systems and augmented reality applications. He also works as the Rector's Advisor at Selcuk University.
E-mail: yuzun@erbakan.edu.tr, **ORCID:**0000-0002-7061-8784

Yunus Emre GOKTEPE, PhD, is an Assistant Professor at Necmettin Erbakan University. The author's areas of expertise are bioinformatics, artificial intelligence and machine learning. He is still serving as a faculty member of the Department of Computer Engineering, Seydisehir Ahmet Cengiz Engineering Faculty at Necmettin Erbakan University in Konya, Turkey.
E-mail: ygoktepe@erbakan.edu.tr, **ORCID:** 0000-0002-8252-2616

Nurgul AYDIN, is a Computer Engineer. She has the bachelor's degree in the Department of Computer Engineering, Seydisehir Ahmet Cengiz Engineering Faculty at Necmettin Erbakan University in Konya, Turkey. Her main areas of interest are artificial intelligence, web programming and machine learning.
E-mail: aydinnurgul87@gmail.com,**ORCID:** 0009-0008-7760-6592

Similarity Index
The similarity index obtained from the plagiarism software for this book chapter is 24%.