

# Deep-Cov19-Hate: A Textual-Based Novel Approach for Automatic Detection of Hate Speech in Online Social Networks throughout COVID-19 with Shallow and Deep Learning Models

Cem BAYDOGAN, Bilal ALATAS\*

**Abstract:** The use of various online social media platforms rising day by day caused an increase in the correct or incorrect information shared by users, especially during COVID-19. The introduction of COVID-19 on the world agenda gave rise to an overall bad reaction against East Asia (esp. China) in online social media platforms. The social media users who spread degrading, racist, disrespectful, abusive, discriminatory, critical, abuse, harsh, offensive, etc. posts accused the Asian people of being responsible for the outbreak of COVID-19. For this reason, the development of the Hate Speech Detection (HSD) system was necessary in order to prevent the spread of these posts about COVID-19. In this article, a textual-based study on COVID-19-related hate speech (HS) sharing in online social networks was carried out with Shallow Learning (SL) and Deep Learning (DL) methods. In the first step of this study, typical Natural Language Processing (NLP) pipeline was applied for gathered two different datasets. This NLP pipeline was performed using bag of words, term frequency, document matrix, etc. techniques for features extraction representing datasets. Then, ten different SL and DL models were fine-tuned for HS datasets related to COVID-19. Accuracy, precision, sensitivity, and F-score performance measurement criteria were calculated to compare the performance of the SL and DL algorithms for the problem of HSD. The RNN, one of the models proposed for the first and second dataset in HSD, prevailed with the highest accuracy values of 78.7% and 90.3%, respectively. Due to the promising results of all approaches operated in the HSD, they are forecasted to be chosen in the solution of many other social media and network problems related to COVID-19.

**Keywords:** COVID-19; hate speech detection; shallow and deep learning models; social media analysis; social network problems

## 1 INTRODUCTION

The COVID-19 outbreak that emerged in the city of Wuhan (in China) has had a great anxiety impact on social media platforms all over the world. In almost all social media outlets, Asian people have been blamed for the COVID-19 outbreak that occurred in the early December 2019. Against the people of this region accusing, discriminatory, incriminating, racist, insulting, humiliating, and other HS posts have started to spread very rapidly from social networks. As a result of the COVID-19 outbreak, China and other Asian countries all over the world have become notorious on all online social media platforms (trend topic on Twitter). Therefore, the proposed HSD system has become necessary to automatically detect HS contents and prevent the spread of these shares. Especially in just a few days, more than millions of posts on Twitter are another reason why the automatic HSD system needs to be developed because it is not possible to detect and prevent this incredible number of HS contents under human supervision [1].

The increase in the use of social media outlets has led to exploitation in HS organizations that advocate or criticize hate and offensive materials. Every other day, social media giants continue their efforts to detect and prevent the spread of offensive, anarchical, or propagation HS materials without slowing down [2]. In the face of crises such as COVID-19, which unexpectedly broke out on the world agenda, technology giants had to keep their existing technological structures up-to-date to prevent the spread of sharing HS materials. Due to the rapid upsurge in the number of HS materials and the impossibility to detect manually, the development of the HSD system has become mandatory. It is seen that it is specially designed to detect the HS related to COVID-19 content circulating in social media with artificial intelligence (AI) supported content-based analysis technologies. In this study, NLP, SL, and DL semantic content-based analysis techniques were

applied to develop an automated HSD system to deal with this crisis. Moreover, these semantic content-based analysis methods, consisting of the-state-of-the-art approaches, yielded the study to achieve promising results. It can be easily concluded that this study, which creates a creative synthesis for the approaches used and future studies, is excellent for the automatic COVID-19-related HSD system.

HS has been one of the biggest problems with social networks even before COVID-19 [3]. HS covers all content such as images or texts that are derogatory, critical, sometimes even blasphemous, and insulting, targeting a person or a group. HS continues to be one of the most important problems in social networks from past to present. In fact, content shared on social networks can be seen as freedom of expression. However, there are very important boundaries between freedom of expression and HS [4, 5]. Without considering this situation, many social media users may intentionally or unintentionally spread or support HS content. The best example of this was seen in the shared HSs that social media users for COVID-19 pointed to East Asian societies targeted on social networks as responsible for this outbreak. HS user network interaction created from social media shares during the COVID-19 outbreak is demonstrated in Fig. 1 [6].

The speed and number of content shared on social media platforms entail the development of the HSD system. Although promising news reported on social networks is our biggest source of motivation in defeating the COVID-19 outbreak, the abuse of social media has brought many social network problems such as HS. In addition, it cannot be denied that HS content shared during the COVID-19 process has caused a crisis on all social media platforms. The proposed HSD system to deal with this crisis was well-timed and its development was inevitable. In the literature, state-of-the-art SL and DL algorithms in the content-based modified systems are applied for automatic HSD. SL and DL algorithms are

well-suited for the automatic HSD problem. Detailed information on these algorithms is also highlighted in section 3. In summary, we have implemented a synthesis of social network analysis techniques suitable for HSD system analysis in the COVID-19 outbreak.

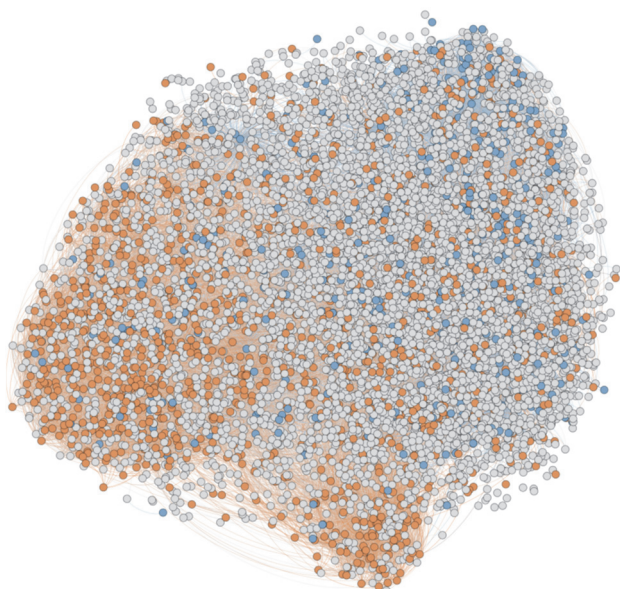


Figure 1 HS distributions on social media during COVID-19 [6] (HS-orange, counter HS-blue, and neutral-grey)

**The contribution of this study:** As a result of our research, it was determined that DL models were very little applied among the proposed methods to solve the HSD problem associated with COVID-19. As our most important goal, more effective solutions were sought to prevent the spread of HS. In accordance with this purpose, we could summarize the contribution of the applied models and our article to the literature and science to satisfy this existing need as follows:

- (1) In order to increase the solution performance in the HSD problem, by applying DL methods in addition to SL models, higher performance results were obtained for HSD related to COVID-19.
- (2) The problem of HSD was modelled as a classification task. Selected feature extraction methods, SL algorithms, and DL models for HSD, which is currently one of the most recent social network problems related to COVID-19, were applied all together for the first time.
- (3) This article is very well-timed to prevent HS spread in social networks while the COVID-19 outbreak continues.
- (4) The use of DL models, which can be easily adapted to solve many other social network problems related to COVID-19, provided more accurate, effective, and reliable results.
- (5) Instead of a single method, ten different SL and DL state-of-the-art methods were operated to find solutions for the HSD.
- (6) It is shown that by most of the essential measurement metrics, DL networks characteristically outperformed SL approaches.

After this part of the article, it is organized as follows: In the second section of the study, studies on the spread and detection of HS during the COVID-19 outbreak are

analyzed. Since the researched topic is very new and fresh, the number of studies is very small. Moreover, the number of methods employed is limited in the literature. The SL and DL models chosen in the third section of the study are listed. These state-of-the-art models picked are applied for HSD about COVID-19. In the fourth section of the study, information about the focused datasets is presented. Then, first applied NLP steps are elucidated. In the next operation, for HSD about COVID-19, the applied feature extraction methods and the approaches chosen for classification are explained in detail. The numerical values of the test results are motivated in section five. The comparison of the results attained with the SL and DL models is highlighted with tables and graphs. In the last section of the paper, attention was drawn to the implications of the importance and necessity of the article. Future works are also underlined in the recommendation subtitle.

## 2 BACKGROUND OF HSD THROUGHOUT COVID-19

Many different studies were carried out in online social media platforms, the use of which increases before and during the COVID-19 outbreak [7-9]. One of these study topics was the HSD system. Considering the importance of HSD, both for societies and world balance, studies on this topic were rare and untouched. In other words, the topic of HSD related to COVID-19 was sufficiently clarified. From the next paragraph, all aspects of HSD research related to COVID-19 were highlighted.

During the COVID-19 outbreak, many different fields of study emerged in many various topics from health sciences to computer sciences, from economy to politics [10-12]. While researches applied in many different fields were especially focused on health, HS materials related to COVID-19 continued to be shared on social media. At the same time, public datasets related to COVID-19 continued to be shared in order to receive the support of all researchers around the world. The study summarizing these public datasets was located in [5].

Ziems et al. gathered about 31M tweets on Twitter during the COVID-19 crisis [6]. They also presented the dataset they prepared by filtering the HS tags, especially targeting the borders of Asia and China, to the use of researchers [13]. They accomplished their work using the BERT and GloVe word-embedding models. They tested the performance of working on their proposed dataset using only one classifier (Logistic Regression). The biggest drawback of this study was that they did not try any other classifier SL or DL methods. Our study seconded this study in this aspect. In fact, this is exactly what the authors stated in their article.

Vidgen et al. conducted a study to analyze the social media reaction against East-Asian people in the COVID-19 outbreak [14]. Researchers created an open-access dataset for other researchers to implement different classifiers or analysis techniques. This open-access dataset was included in [15]. Researchers using 7 different models yielded 88% precision. SL techniques were never used in the study. In addition, only LSTM, one of the DL approaches, was applied. In this respect, our study completed, supported, and satisfied this study.

In another completed study, Awal et al. collected about 40M anti-social abusive tweets and contributed to the literature. Using lexicon-based approaches and the Perspective-API, they achieved the automatic labelling of these data [16]. Hardage et al. practiced a toxic and HSD study on COVID-19 related tweets using GloVe and CNN methods [17]. Specifically, they aimed to increase the false positive evaluation criterion.

In the background of HSD research, there were few reviews and quantitative studies that emphasized the importance of detecting, monitoring, and controlling HS spread on social media during COVID-19 [18-20]. Other studies clearly explained why it is necessary to develop HSD systems, in particular [21, 22]. The power of social media in our age was emphasized. Requirements were listed for the necessity of increasing social media studies. In another study, 5G conspiracy theory was put forward in social media in the COVID-19 bio-crisis [23]. Another study emphasized the importance of coping with cyber racism in the COVID-19 process [24]. In other review articles, in another unexpected crisis, it was clearly observed that the attitudes and behaviours of people on social media had a global impact [25, 26].

In another study, Cotik et al. carried out for HSD in the COVID-19 outbreak [27]. In the study conducted using tweets in Spanish, the F-score yielded 75% of the evaluation value. The use of a single approach indicated that this study was the first version. Nemes et al. fulfilled basic-NLP steps in the sentiment analysis study related to COVID-19. They also strengthened their work using RNN architecture. The numerical values of all experiments obtained in the study were highlighted in [28].

When the background of the HSD system during COVID-19 was scanned, it was understood in the literature that these HSD systems still needed to be developed. It is estimated that the proposed state-of-the-art approaches will fill this gap in the literature.

### 3 METHODOLOGY

In this section, the steps followed for the HSD system related to COVID-19 were listed. The first step of the applied methodology consisted of the characteristics and numerical information of the COVID-19 data used in the study. Then, the applied NLP and text pre-processing process was explained. Finally, the performance of state-of-the-art SL and DL approaches selected in the HSD study were motivated.

#### 3.1 Characteristic of the COVID-19 Data

This HSD study focused on 2 different datasets collected during the COVID-19 period. These datasets used consisted of pre-tagged tweets. These datasets collected throughout COVID-19 revealed that there is a major real-world social problem that needs to be solved on social media.

The first COVID-19 dataset-1 consisted of 2319 tweets and four different classes [13]. The other COVID-19 dataset-2 consisted of 20 thousand (20K) tweets and five different classes [15]. The tweets in these COVID-19 datasets were created from English tweets. Label names

and numerical information on the COVID-19 datasets are illustrated in Fig. 2.

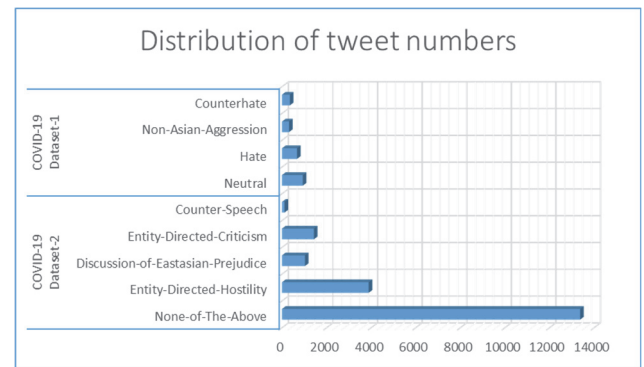


Figure 2 Distribution of used COVID-19 datasets according to labels

#### 3.2 Pre-processing and NLP Steps of HSD System

In the second step of the applied methodology, pre-processing and NLP pipelines were employed for 2 different COVID-19 datasets referenced. First of all, the data from the source was converted to the appropriate formats (.csv and .xlsx) to be able to apply the NLP pipelines. When the original data were looked into, there were many details in the concentrated datasets. The columns suitable for our study were filtered.

Then, the following pre-processing steps were performed for COVID-19 tweets on the filtered columns.

- **Punctuation erasure** (the punctuation in the COVID-19 corpus was deleted),
- **Number(s) filter** (numeric values in the COVID-19 text were filtered out),
- **n-char(s) filter** (by choosing  $n = 3$ , expressions with  $n\text{-char} < 3$  were separated from the corpus),
- **Stop-word-filter** (custom words should be selected for this filter),
- **Case-converter** (lowercase conversion applied to all COVID-19 corpus).

After the pre-processing stages were completed, stemmer, bag of words, TF-IDF, and document matrix were fulfilled for feature extraction. The TF (term-frequency) value operated in the study was calculated according to Eq. (1).

$$TF(t, corpus) = \frac{(\text{count of 't' in corpus})}{(\text{number of bag in corpus})} \quad (1)$$

$t$  and  $corpus$  represented each term and the entire COVID-19 dataset, respectively.

#### 3.3 Approaches for COVID-19

For the HSD system developed for COVID-19, classifier selection was operated in the third step of the methodology. In this selection, SL and DL methods, which are among state-of-the-art approaches used in the literature, were preferred. These classifiers were selected from state-of-the-art algorithms, which are applied in solving many real-world problems and most recommended for artificial intelligence supported systems. These approaches were listed under sub-headings.

### 3.3.1 SL-based Approaches

Shallow learning (SL), also known as Machine Learning, was a frequently used classifier in the literature. In this study, state-of-the-art SL approaches were chosen to ensure the reliability of the HSD system. In this study, the most reliable SL approaches were applied for HSD, which was one of the problems caused by the unexpected COVID-19 crisis in social media.

These preferred approaches were 5 different state-of-the-art SL approaches. These were listed as follows: Support vector machines (SVM), k-nearest neighbour (KNN) algorithm, naive Bayes (NB), random forest (RF), and logistic regression (LR). These methods were abundantly motivated in the literature. In the study, it was not preferred to motivate these approaches in detail in order

not to bore the reader. However, the comparative use of these selected approaches in the HSD system related to COVID-19 was highly appreciated.

### 3.3.2 DL-based Approaches

The other approach applied in the HSD system for COVID-19 was DL-based approaches. DL-based approaches were reliable methods used in solving many engineering problems. In this study, 5 different state-of-the-art DL-based algorithms were preferred to solve the HSD problem. The promising results of these approaches in other studies were very important in terms of increasing the reliability of our study.

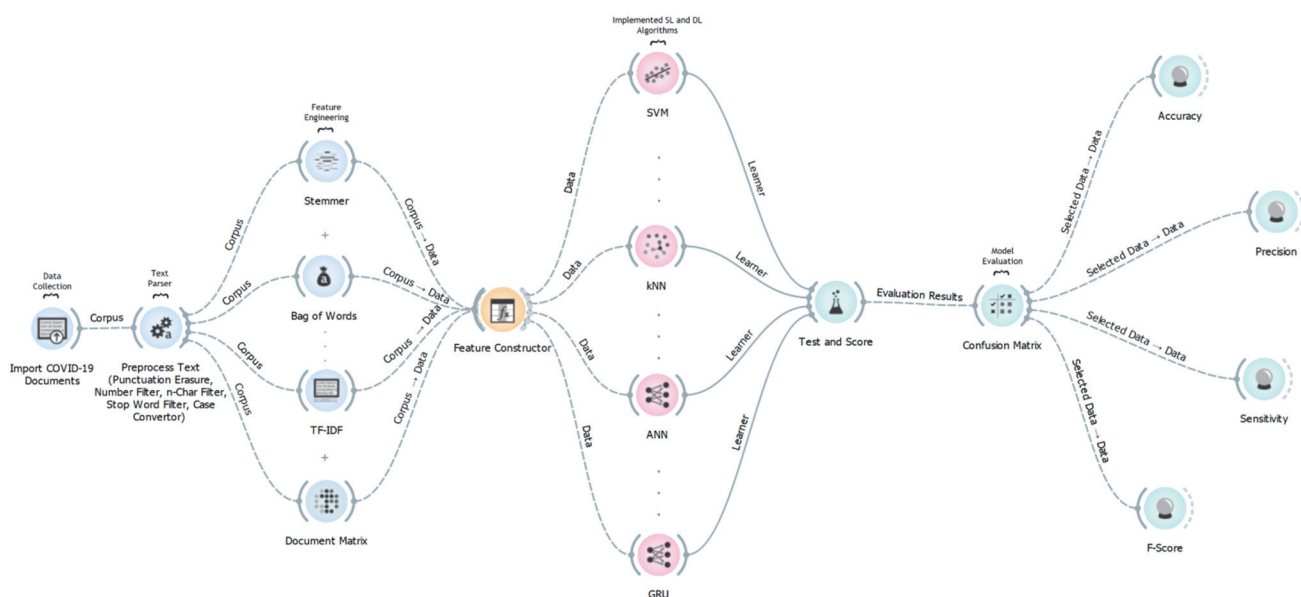


Figure 3 Workflow of the proposed approaches for the HSD system related to COVID-19

We listed these approaches in DL-based: artificial neural networks (ANN), convolutional neural network (CNN), recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent units (GRUs). DL-based approaches were already motivated in the literature. In order not to bore the reader, it was not motivated again. However, the parameters of these algorithms were also highlighted in section 4. These DL-based approaches were not applied together in the literature to solve the HSD problem during the COVID-19 process. The combination of these methods clearly demonstrated the power of this study and its contribution to the literature.

The proposed system for the HSD problem occurring in the COVID-19 crisis is summarized as in Fig. 3

When Fig. 3 was synthesized, the first step of the proposed system was initiated with the inclusion of the COVID-19 corpus files in the HSD system. The system was run separately for two different datasets. Then, parsing techniques were applied to clean the data in the text parser step process on these datasets. In the next step, the feature constructor process on the corpus was completed with feature engineering approaches. Stemmer, bag of words, TF-IDF, document matrix, and referenced row filter were some of the strategies in this process. Ten different SL and DL state-of-the-art algorithms were implemented to

increase the reliability of the proposed HSD system. While five of them were SL algorithms, the rest were DL algorithms. Finally, various evaluation criteria were calculated on the test data in the dataset for model evaluation whether the system works correctly. The evaluation findings of the proposed system were also discussed in detail.

## 4 EXPERIMENTS AND RESULTS FOR HSD

The COVID-19 study, which was carried out with different approaches, was completed on 2 different sets of experiments. In each performed experiment, 70% of the datasets were reserved for training and the rest for testing. All experiments were employed under equal conditions.

Evaluation metrics were operated to compare the performance of algorithms in the proposed system. The performance measurement criteria expressions calculated in the study are summarized in Tab. 1. This experimental COVID-19 study was completed using an Intel Core i5 10210U processor, 8 GB RAM, 256 GB SSD, and 4 GB MX110 graphics card.

**Table 1** Complexity metrics and evaluation criteria for HSD

Evaluation criteria	Equations	Complexity metrics
Accuracy	$\frac{(TP+TN)}{(TP+TN+FP+FN)}$	TP: True Positive TN: True Negative FP: False Positive FN: False Negative
Precision	$\frac{(TP)}{(TP+FP)}$	
Sensitivity	$\frac{(TP)}{(TP+FN)}$	
F-score	$2 \cdot \frac{(TP)}{(TP+FP+FN)}$	

**4.1 Results for COVID-19 Dataset-1**

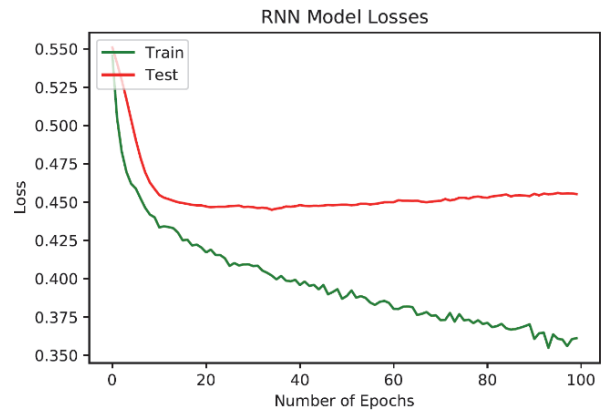
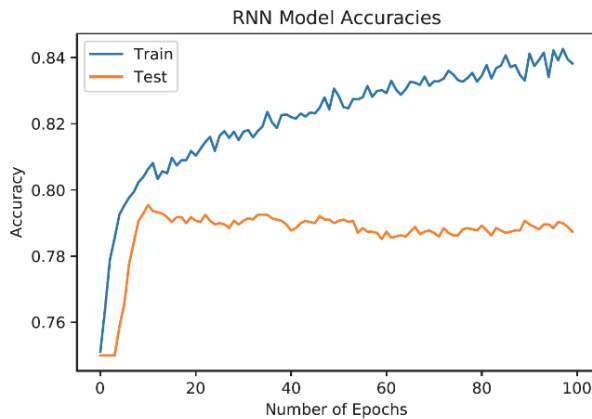
It was previously reported that the COVID-19 dataset-1 consisted of 2319 tweets. Ten different approaches preferred for this dataset were run under equal conditions in each experiment. For all algorithms, 70% of the COVID-19 dataset-1 was applied for training while the rest was used for testing. All DL-based approaches repeated 100 epochs for all set of experiments. Stochastic gradient descent (sgd) was chosen as the optimizer in all DL-based approaches. In addition, relu and soft max were employed as activation functions. Finally, hidden layers were added with an equal number of nodes (512) for each DL-based approach. Since the modelled networks were run on multi-class datasets, categorical cross entropy was set as the loss function. Default parameters were applied to the learning rate, decay steps, and momentum weights at the model compiler stage.

The performance measure criteria yielded with ten different algorithms operated for the COVID-19 dataset-1 are demonstrated in Tab. 2.

**Table 2** Evaluation criteria for COVID-19 dataset-1

Exp-1		Evaluation criteria			
		Accuracy	Precision	Sensitivity	F-Score
Applied Algorithms	ANN	0.7519	0.5097	0.4604	0.4818
	RNN	0.7871	0.6038	0.3718	0.4511
	LSTM	0.7540	0.2718	0.0442	0.0744
	CNN	0.7655	0.5475	0.2207	0.3079
	GRU	0.7517	0.2683	0.0256	0.0454
	SVM	0.5528	0.5400	0.5530	0.5010
	KNN	0.5147	0.4880	0.5150	0.4830
	NB	0.5176	0.5080	0.5180	0.5100
	RF	0.5220	0.4980	0.5220	0.5040
	LR	0.5308	0.5050	0.5310	0.4970

For the first experiment (Exp-1), SL-based approaches lagged far behind DL-based approaches. The highest accuracy and precision values were achieved by the RNN model. The second highest accuracy value was obtained by the CNN algorithm with 76.6%. ANN, LSTM, and GRU lagged behind RNN and CNN models with approximately 75.5% accuracy value. The worst accuracy value for Exp-1 was reached by KNN. Other SL-based algorithms again lagged far behind DL-based algorithms. In the other evaluation criteria for Exp-1, although striking superiority results were not observed in SL-based algorithms, ineffective results were also monitored in DL-based approaches. Although the highest sensitivity value was reached by the SL-based SVM algorithm, the worst sensitivity value was obtained by LSTM. Finally, the highest F-score value was obtained by NB, which is another SL-based algorithm, while the worst F-score value was also attained by LSTM. In this experiment, other DL-based models outperformed SL-based algorithms with higher than 75% accuracy results. The performance evaluation graph of the RNN model, which was the champion of Exp-1, is illustrated in Fig. 4.



a) RNN model accuracy for COVID-19 dataset-1

b) RNN model loss for COVID-19 dataset-1

**Figure 4** Performance evaluation graphs of the RNN model for the COVID-19 dataset-1

**4.2 Results for COVID-19 Dataset-2**

It was reported in section 3.1 that the COVID-19 dataset-2 consisted of approximately 20K tweets. Ten different SL and DL-based approaches preferred for COVID-19 dataset-2 were employed under the same conditions in each experiment. As in the first experiment, for all models, 70% of the COVID-19 dataset-2 was performed for training while the rest was used for testing.

**Table 3** Evaluation criteria for COVID-19 dataset-2

Exp-2		Evaluation criteria			
		Accuracy	Precision	Sensitivity	F-Score
Applied Algorithms	ANN	0.8892	0.7313	0.7058	0.7180
	RNN	0.9030	0.8085	0.6760	0.7348
	LSTM	0.8722	0.6959	0.6253	0.6577
	CNN	0.8908	0.7804	0.6330	0.6972
	GRU	0.8744	0.7034	0.6290	0.6634
	SVM	0.7440	?	0.7440	?
	KNN	0.7225	?	0.7230	?
	NB	0.6859	0.7010	0.6860	0.6910
	RF	0.7390	0.7030	0.7390	0.7140
	LR	0.7499	0.7030	0.7500	0.7130

All DL-based models executed 100 epochs in all set of experiments for COVID-19 dataset- 2.sgd was chosen again as the optimizer in all DL-based models. Same as previous experiment, relu and softmax were employed as activation functions. Moreover, hidden layers were designed with the same number of nodes (512) for each DL-based model in Exp-2. For the Exp-2,

categorical\_crossentropy was set as the loss function. Likewise, default coefficients were applied to the learning rate, decay steps, and momentum weights at the model compiler stage for Exp-2. The performance measure criteria yielded with ten different SL and DL algorithms laboured for the COVID-19 dataset-2 are indicated in Tab. 3.

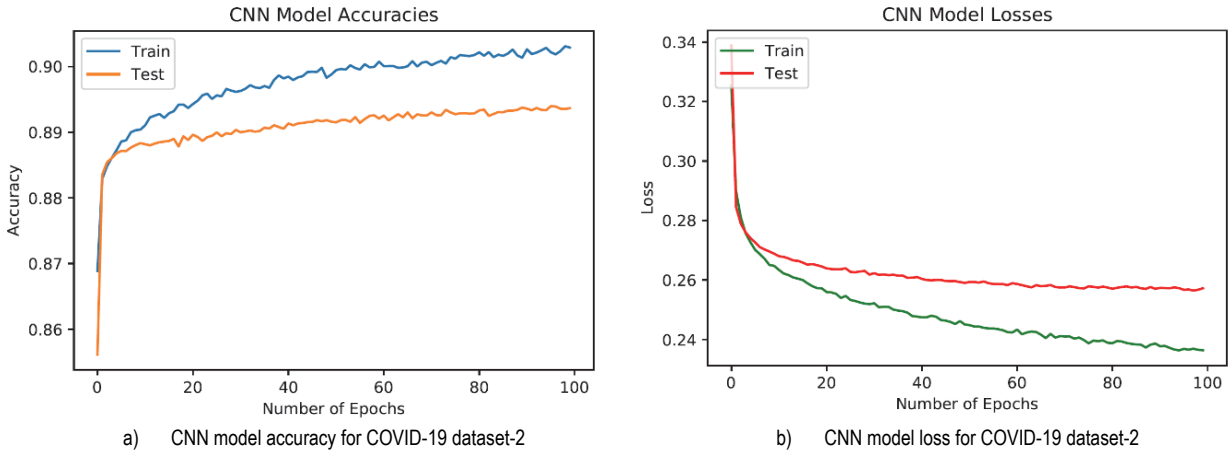


Figure 5 Performance evaluation graphs of the CNN model for the COVID-19 dataset-2

When the applied Exp-2 was analyzed, the highest accuracy rate was accomplished by the RNN algorithm. This 90% accuracy turned out to be an appreciated result for this problem with 20K tweets on a five-class COVID-19 dataset-2. With an accuracy of 89.1%, the CNN model overcame the second highest accuracy for the Exp-2. The LSTM, GRU, and ANN models again outperformed SL-based algorithms. The worst accuracy value was output by the NB algorithm. In turn, other SL-based algorithms reached the worst accuracy. Consequently, DL-based models achieved superior success against SL-based algorithms, except for the sensitivity value. The highest sensitivity value for the COVID-19 dataset-2 was attained by the SL-based LR algorithm. While DL-based approaches showed superiority for Exp-2 in general, the worst sensitivity values were achieved by LSTM and GRU. The highest precision and F-score values were also yielded by RNN. The second highest accuracy and precision values

for this challenge were gained by the CNN approach. While ANN and GRU outperformed SL-based algorithms in precision, LSTM produced approximately similar results. SVM and KNN algorithms could not yield the precision and F-score values at all. They could not be calculated because they were too small. ANN approach achieved the second highest F-score with 72%, outpacing SL-based algorithms. Other DL-based approaches obtained approximately similar results. As a result, in the synthesis of the two experiments, it was seen that DL-based approaches were overwhelmingly superior to SL-based approaches when considering the evaluation criteria.

The epoch-dependent accuracy and loss graphs of the CNN architecture are exhibited in Fig. 5. Throughput of numerical information of all experimental studies and applied methods in this completed HSD system are summarized in Fig. 6 with the help of line graphs.

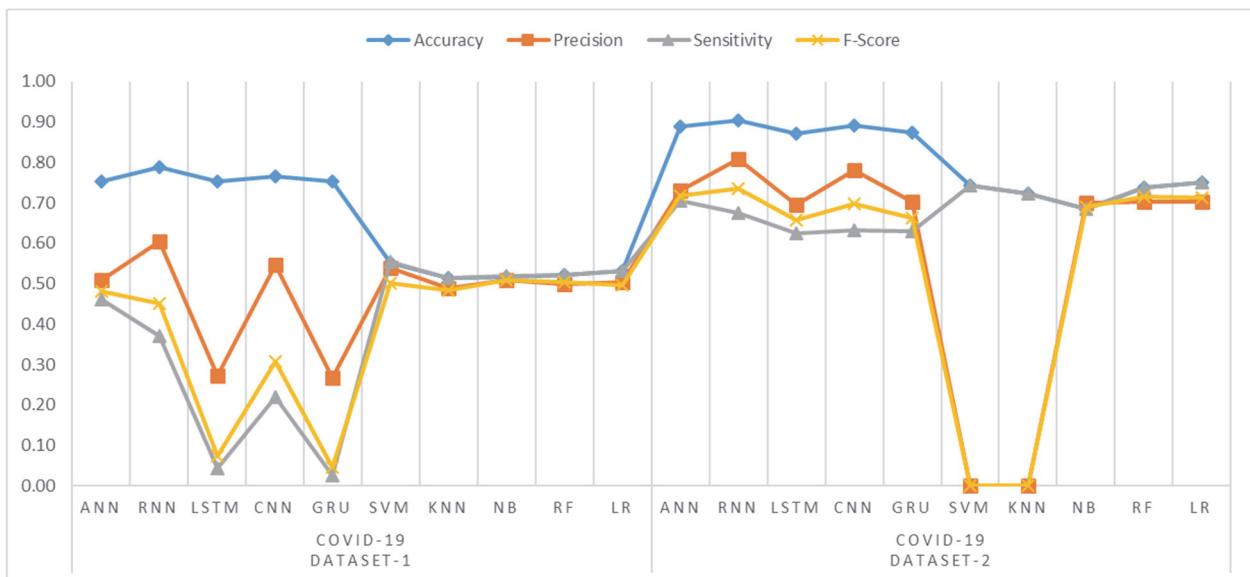


Figure 6 Ranking of performance evaluation for all COVID-19 experiments

## 5 CONCLUSION

The studies and the methods applied show that HSD is of importance that needs to be studied in more depth. In this study, we selected state-of-the-art SL and DL approaches to complement the two previous studies and support these studies. With this study, we succeeded in filling the gap in the field of HS, social network problem occurring in the COVID-19 unexpected bio-crisis.

In the synthesis of the study, it was observed that DL-based approaches were overwhelmingly outperformed by SL-based approaches. With the RNN model, an accuracy of 79% was obtained for the first dataset, while an accuracy value of over 90% was attained for the second dataset. The unbalanced distribution of the datasets used in the study constitutes the biggest limitation of the study. It is forecasted to obtain higher performances in datasets with different balanced distributions because this is an essential factor for training the proposed system. The absence of an operated study using the same datasets is thought to be a reference for future studies.

This study is a rare and powerful experimental original study proposed for HSD related to COVID-19 because many different approaches have been used together. In addition, the results obtained from these approaches were analyzed in depth. However, these proposed approaches and methods for HSD can be extended.

Recommendations for future work can be structured as follows. Hybrid DL approaches (e.g. CNN+Bi-LSTM) can be used for HSD. Metaheuristic methods with Pareto approaches can be applied to this problem. Moreover, different feature extraction models can be adapted to this HSD problem. For many unexpected events, such as the COVID-19 bio-crisis, automatic detection, monitoring, counteraction, and blocking techniques can be developed on social networks. Once and for all, by collecting different HS datasets, suggested approaches can be implemented. DL-based approaches can achieve higher performance if different loss functions, activation functions, optimizers, epoch number are well adjusted. In addition, the success of the algorithms can be increased by trying different variations instead of the default coefficients.

## 6 REFERENCES

- [1] Burnap, P. & Williams, M. L. (2015). Cyber hate speech on twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & internet*, 7(2), 223-242. <https://doi.org/10.1002/poi3.85>
- [2] Cao, R. & Lee, R. K. W. (2020, December). HateGAN: Adversarial Generative-Based Data Augmentation for Hate Speech Detection. *Proceedings of the 28th International Conference on Computational Linguistics*, 6327-6338.
- [3] Al-Makhadmeh, Z. & Tolba, A. (2020). Automatic hate speech detection using killer natural language processing optimizing ensemble deep learning approach. *Computing*, 102(2), 501-522. <https://doi.org/10.1007/s00607-019-00745-0>
- [4] Awal, M. R., Cao, R., Lee, R. K. W., & Mitrović, S. (2020). On analyzing annotation consistency in online abusive behavior datasets.
- [5] Shuja, J., Alanazi, E., Alasmay, W., & Alashaikh, A. (2020). Covid-19 open source data sets: A comprehensive survey. *Applied Intelligence*, 1-30. <https://doi.org/10.1007/s10489-020-01862-6>
- [6] Ziems, C., He, B., Soni, S., & Kumar, S. (2020). Racism is a virus: Anti-asian hate and counterhate in social media during the covid-19 crisis.
- [7] Xu, P., Ruan, Q., & Xu, W. (2021). Text Adversarial Examples Generation and Defense Based on Reinforcement Learning. *Tehnički vjesnik*, 28(4), 1306-1314. <https://doi.org/10.17559/TV-20200801053744>
- [8] Zhang, Y., Strauss, J., & Liu, L. (2021). An OLS and GMM Combined Algorithm for Text Analysis for Heterogeneous Impact in Online Health Communities. *Tehnički vjesnik*, 28(2), 587-597. <https://doi.org/10.17559/TV-20210121100916>
- [9] Alakus, T. B. & Turkoglu, I. (2020). Comparison of deep learning approaches to predict COVID-19 infection. *Chaos, Solitons & Fractals*, 140, 110120. <https://doi.org/10.1016/j.chaos.2020.110120>
- [10] Yu, Y. A. N. G. (2021). Lineage in Rural China: Social Network or Patriarchal Clan System? Evidence from Its Impact on Irrigation Infrastructure Management. *Technical Gazette*, 28(4), 1315-1320. <https://doi.org/10.17559/TV-20200725104840>
- [11] Xue, Y. (2021). A Dynamic Trust Relations-Based Friend Recommendation Algorithm in Social Network Systems. *Tehnički vjesnik*, 28(1), 185-192. <https://doi.org/10.17559/TV-20200825171016>
- [12] Gauthier, G. R., Smith, J. A., García, C., Garcia, M. A., & Thomas, P. A. (2021). Exacerbating inequalities: social networks, racial/ethnic disparities, and the COVID-19 pandemic in the United States. *The Journals of Gerontology: Series B*, 76(3), e88-e92. <https://doi.org/10.1093/geronb/gbaa117>
- [13] Ziems, C., He, B., Soni, S., & Kumar, S. (2020). Retrieved from <http://claws.cc.gatech.edu/covid/#dataset>
- [14] Vidgen, B., Botelho, A., Broniatowski, D., Guest, E., Hall, M., Margetts, H., Hale, S. et al. (2020). Detecting East Asian prejudice on social media.
- [15] Vidgen, B., Botelho, A., Broniatowski, D., Guest, E., Hall, M., Margetts, H., Hale, S. et al. (2020). Detecting East Asian Prejudice on Social Media [Dataset]. Zenodo.
- [16] Awal, M. R., Cao, R., Mitrovic, S., & Lee, R. K. W. (2020). On analyzing antisocial behaviors amid COVID-19 pandemic.
- [17] Hardage, D. & Najafirad, P. (2020). Hate and Toxic Speech Detection in the Context of Covid-19 Pandemic using XAI: Ongoing Applied Research.
- [18] Lee, R. K. W. & Li, Z. (2020). Online Xenophobic Behavior Amid the COVID-19 Pandemic: A Commentary. *Digital Government: Research and Practice*, 2(1), 1-5. <https://doi.org/10.1145/3428091>
- [19] Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Z. (2020). Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. *Journal of medical internet research*, 22(4), e19016. <https://doi.org/10.2196/19016>
- [20] Budhwani, H. & Sun, R. (2020). Creating COVID-19 stigma by referencing the novel coronavirus as the "Chinese virus" on Twitter: quantitative analysis of social media data. *Journal of Medical Internet Research*, 22(5), e19301. <https://doi.org/10.2196/19301>
- [21] Alshalan, R., Al-Khalifa, H., Alsaed, D., Al-Baity, H., & Alshalan, S. (2020). Detection of Hate Speech in COVID-19 - Related Tweets in the Arab Region: Deep Learning and Topic Modeling Approach. *Journal of Medical Internet Research*, 22(12), e22609. <https://doi.org/10.2196/22609>
- [22] Jo, W., Lee, J., Park, J., & Kim, Y. (2020). Online information exchange and anxiety spread in the early stage of the novel coronavirus (COVID-19) outbreak in South Korea: structural topic model and network analysis. *Journal of medical internet research*, 22(6), e19455. <https://doi.org/10.2196/19455>

- [23] Ahmed, W., Vidal-Alaball, J., Downing, J., & Seguí, F. L. (2020). COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. *Journal of medical internet research*, 22(5), e19458. <https://doi.org/10.2196/19458>
- [24] Dubey, A. D. (2020). The resurgence of cyber racism during the COVID-19 pandemic and its aftereffects: analysis of sentiments and emotions in tweets. *JMIR Public Health and Surveillance*, 6(4), e19833. <https://doi.org/10.2196/19833>
- [25] Dubey, A. D. (2020). Twitter Sentiment Analysis during COVID-19 Outbreak.
- [26] Uyheng, J. & Carley, K. M. (2021). Characterizing network dynamics of online hate communities around the COVID-19 pandemic. *Applied Network Science*, 6(1), 1-21. <https://doi.org/10.1007/s41109-021-00362-x>
- [27] Cotik, V., Debandi, N., Luque, F. M., Miguel, P., Moro, A., Pérez, J. M., Zayat, D. et al. (2020). A study of Hate Speech in Social Media during the COVID-19 outbreak.
- [28] Nemes, L. & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication*, 5(1), 1-15. <https://doi.org/10.1080/24751839.2020.1790793>

**Contact information:**

**Cem BAYDOGAN**, PhD candidate  
Department of Software Engineering,  
Faculty of Technology, Firat University, Elazig, 23119, Turkey  
E-mail: vcbaydogan@firat.edu.tr

**Bilal ALATAS**, PhD, Professor  
(Corresponding author)  
Department of Software Engineering,  
Faculty of Engineering, Firat University, Elazig, 23119, Turkey  
E-mail: balatas@firat.edu.tr