Selective Feature Sets Based Fake News Detection for COVID-19 to Manage Infodemic

MANIDEEP NARRA, MUHAMMAD UMER, SAIMA SADIQ, ALA’ ABDULMAJID ESHMAWI, HANEN KARAMTI, ABDULLAH MOHAMED, AND IMRAN ASHRAF

ABSTRACT During the COVID-19 pandemic, the spread of fake news became easy due to the wide use of social media platforms. Considering the problematic consequences of fake news, efforts have been made for the timely detection of fake news using machine learning and deep learning models. Such works focus on model optimization and feature engineering and the extraction part is under-explored area. Therefore, the primary objective of this study is to investigate the impact of features to obtain high performance. For this purpose, this study analyzes the impact of different subset feature selection techniques on the performance of models for fake news detection. Principal component analysis and Chi-square are investigated for feature selection using machine learning and pre-trained deep learning models. Additionally, the influence of different preprocessing steps is also analyzed regarding fake news detection. Results obtained from comprehensive experiments reveal that the extra tree classifier outperforms with a 0.9474 accuracy when trained on the combination of term frequency-inverse document frequency and bag of words features. Models tend to yield poor results if no preprocessing or partial processing is carried out. Convolutional neural network, long short term memory network, residual neural network (ResNet), and InceptionV3 show marginally lower performance than the extra tree classifier. Results reveal that using subset features also helps to achieve robustness for machine learning models.

INDEX TERMS Fake news detection, ResNet, inceptionV3, principal component analysis.

I. INTRODUCTION
Fake news has been in the limelight during the past few years due to the explosive use of social media platforms. Recently, the wide acceptance and use of social media platforms facilitated people for easy dissemination of views, information and opinions [1]. Social media plate forms have become the major source of information with wide and easy access [2]. With the rapid growth and easy access, sharing fake news has become very easy. Often, to get attention fake news is disseminated without realizing the widespread impact and panic that fake news creates. Fake news contains false information, altered information, or partial truth to deceive people and it spreads fast to reach a large audience [3]. It makes it difficult and non-trivial to detect fake news and stop it timely [4]. Therefore, fake news detection has become a topic of great interest recently.

Similar to other domains, coronavirus disease (COVID-19) related news suffers from fake and misinformation news. With COVID-19 declared as a pandemic, a large number of informative, propaganda, and misleading information exploded on social media platforms. Life-threatening COVID-19 led people to search and study online material to search for its symptoms, probable remedies, precautions to reduce the chance of the disease, and others. With a large audience, each looking for information regarding COVID-19,
the spread of false information and fake news became an easy task for adversaries. Consequently, a large number of fake news regarding COVID-19 have been circulating on different information-sharing platforms including social media and medical information-sharing platforms. It created a lot of stir and panic among common people. Especially, with the emergence of new mutations, news of deaths, and their impact on respiratory organs created panic among the general public [5].

Besides creating panic among the general public, fake news has the potential to negatively influence the economy of the country, reduce the confidence of the general public in their government and create an artificial collapse of daily products. The fake news about the lockdowns, health statistics, and vaccines also fuelled the panic in purchases like paper products, masks, sanitizers, and groceries [6]. This results in a shortage of supply and food security and exacerbated demand-supply gaps. Similarly, fake news rapidly affected the economy and fuel prices [7], [8]. Since the inception of COVID-19, the world health organization (WHO) issues data related to COVID-19 regularly which also contains warnings and directives related to COVID-19 and infodemic [9]. Despite the instructions from governments and WHO alike about fake news, it is very difficult to check the validity, correctness, and credibility of the news [10]. WHO collaborated with the world’s leading search engines to display the latest official reports on the top hits of search related to the COVID-19 [11].

A. MOTIVATION

Because of the problems faced due to fake news on online platforms, the quality of online news is very low as compared to traditional news sources. For example, handling compact and noisy online data is a challenging task [12]. It is easy to spread false information using online platforms that can mislead people easily toward negative thoughts. It also creates doubts and affects the legitimacy of online news. To avoid the negative impact of fake news on society, its timely detection is an important research problem [13]. With a large amount of data shared every day online, it is not possible to analyze it for detecting fake news. Data mining plays a vital role in extracting useful information from the data [14]. Fake news about COVID-19 has become one of the problems in our society today. Because of diverse fake news, people suffer in many ways on a physical, financial, and mental level. In this work, we analyze multiple data sets to identify the most useful machine learning classifier for detecting false news to avoid these awful circumstances [15]. Many machine learning algorithms have been used to explore the complexity and non-linear interaction between different factors by decreasing the error in prediction and factual results [16].

Since the performance of machine learning models depends on the feature set, this study adopts feature selection. Important features and features with high correlation tend to show better performance than using all the features in a dataset [17], [18]. Selecting features with higher weight and importance can be obtained using principal component analysis (PCA) and Chi-square. Similarly, the use of preprocessing steps removes redundant information and reduces the feature space size which improves the training process of the models. This study performs an in-depth analysis of feature selection and preprocessing for fake news detection.

B. CONTRIBUTIONS

To combat the fake news related to COVID-19, machine learning approaches can be leveraged to help analyze news related to COVID-19. Existing studies focus on enhancing the performance by optimizing the models or using appropriate feature engineering approaches. However, the use of a subset of features or selective features from the dataset remains an under-explored or ignored area. The data containing fake news of COVID-19 is large and has extensive symmetrical features [19]. This motivates the current research to develop a machine learning-based approach with the use of selective feature sets to detect the fake news related to COVID-19. In this regard, this study makes the following key contributions:

- An extensive investigation of various feature selection approaches is carried out for COVID-19 fake news detection. From this perspective, two approaches are analyzed including PCA and Chi-square.
- The efficiency of PCA and Chi-square is analyzed with different feature extraction approaches like term frequency-inverse document frequency (TF-IDF) and bag of words (BoW). The performance of TF-IDF and BoW is evaluated comprehensively with PCA and Chi-square separately.
- Several machine learning models are adopted for fake news detection like random forest (RF), extra tree (ET) classifier, gradient boosting machine (GBM), logistic regression (LR), Naive Bayes (NB), stochastic gradient (SG) and a voting classifier (VC) comprising LR and SG.
- The suitability of deep learning models is investigated for COVID-19 fake news detection using both custom-built models and pre-trained models. Convolutional neural networks (CNN) and long short-term memory (LSTM) models are custom designed for this purpose while two well-known pre-trained models residual neural network (ResNet) and InceptionV3 are adopted as well.

The rest of the paper is organized as follows: Section II discusses recent studies related to fake news. Section III gives the summary of the dataset and a brief description of the adopted methodology and models used for fake news detection. Section IV presents the discussions and analysis of the results. The conclusion and the future directions are given in Section V.

II. RELATED WORK

Fake news detection is the foundation of many tasks such as, claim validation [20] and argument search [21]. In many researches, fake news detection regarding a specific
target is performed from tweets [22], [23], [24] and online debates [23], [25], [26]. Such kind of target-oriented approaches are based on the lexical features [26], linguistics and structural features [25]. Traditionally, a two-step process is followed for fake news detection where preprocessing is carried out in the first step, followed by the feature extraction. Predominantly, existing studies follow the same procedure.

A. FAKE NEWS DETECTION USING MACHINE LEARNING APPROACHES

Several studies employed machine learning models to discover fake news on social media. The study [27] involves an investigation of a large number of machine learning models for fake news detection. The study follows a two-step method where several preprocessing steps are done followed by forming feature vectors using term frequency and document term matrix. The second step involves training and testing models using four evaluation metrics. Another study [5] used datasets from WHO with ten machine learning algorithms and seven feature extraction techniques.

Additionally, another study [28] analyzed data from Twitter to identify the key characteristics that had an impact on how well machine learning techniques classified fake news. The study included several brand-new Twitter features that may help identify bogus from genuine tweets. The research by [29] also suggested a graph-based semi-supervised learning algorithm to identify fraudulent Twitter users utilizing specific useful features.

Similar to this, by employing a two-layer classification, the researchers in [30] offered a text-based approach to fake news identification. The fake topics were discovered using the first layer, and the fake event was discovered using the second layer. The research paper [31] offered a hybrid strategy that used semantic analysis, naive Bayes, and support vector machines for identifying fake news on social networks.

B. FAKE NEWS DETECTION USING DEEP LEARNING APPROACHES

The majority of existing studies follow fake news detection using a single source or domain; fake news detection involving multiple domains is a complicated task and only a few studies follow this direction. For example, authors investigated fake news in cross-domain by applying a hybrid model [32]. They applied LSTM and depth LSTM with CNN and applied linguistic inquiry and n-grams to extract significant features from the news. A deep CNN model namely FNDNet has been proposed to detect fake news [33]. A CNN-based capsule network is proposed for fake news detection [34]. They used four capsule networks for long text and two capsules for short text. Authors proposed FakeBERT which is designed using deep-stacked layers of CNN for fake news detection [35].

A deep neural network was applied to categorize news content and context-based information separately as well as together, with the help of the best hyper-parameters [36]. Their proposed method’s effectiveness has been verified using a real-world dataset. The authors take into consideration the behavior of several Facebook account-related variables and use a deep learning-based analyzer to examine the activity of the account [37]. Authors applied CNN, LSTM, and Bi-LSTM on news article datasets for fake news detection [38].

C. FAKE NEWS DETECTION DURING COVID-19 PANDEMIC

Fake news became an attractive and important research area during the COVID-19 pandemic to fight the infodemic. Fake news prediction related to COVID-19 is quite important, similar to other domains, with a more real-time effect to cause rapid panic. Elhadad et al. [5] worked on the detection of misleading information related to COVID-19. Besides using twelve different performance metrics the study uses 5-fold cross-validation to validate the results. The best results are achieved by the neural network (NN), decision tree (DT), and LR classifiers.

The authors investigate the use of machine learning and deep learning techniques to detect COVID-19 fake news in [39]. The TF-IDF and word2vec word embedding techniques are also included in the work. Results indicate that the support vector machine (SVM) gives the highest F1 score of 93.39% with TF-IDF features. Raha et al. worked on automatic detection of fake news related to COVID-19 in [40]. RF gives a remarkable accuracy of 96.6% while NB gives an accuracy of 95.05%. Koirala [41] proposed a deep learning-based system for the classification of the COVID-19 fake news. The dataset used in this study is inconsistent which leads to deviations in the applied model where the accuracy of the models also deviates.

Besides using well-known machine learning models for fake news detection, the use of transfer learning and optimization models is reported to have better results. For example, the authors discussed machine learning and deep learning-based approaches for COVID-19 fake news detection in [42]. Ensemble of three transfer learning approaches including Bidirectional Encoder Representations from Transformers (BERT), ALBERT, and XLNET has been analyzed for fake news detection on social media comments [43].

Researchers investigated COVID-19-related misinformation using three feature selection approaches; particle swarm optimization, the genetic algorithm, and the salp swarm algorithm [19]. The genetic algorithm outperformed other approaches. Optimization approaches are leveraged in [44] where an optimized Salp swarm optimization approach is adopted for fake news detection. Experimental results show that the optimized model shows superior performance to standard models. Similarly, a metaheuristic optimization algorithm Grey Wolf optimization is adopted in [13] for fake news detection. Results show its better performance over SVM, NB, DT, and J48.

Other than using a fake news dataset for simple feature extraction, a few studies explore metadata where different important aspects related to fake and genuine news are analyzed. For example, Ibrishimova and Li [45] proposed a system based on factual accuracy and relative reliability of
a source. The authors also propose a fake news detection system that uses automated and manual knowledge verification and stylistic features. Gravanis et al. [46] proposed a system for news credibility validation. They used machine learning models and present an ensemble model to obtain good accuracy. For news credibility verification the study uses Lagrangian SVM and multinomial NB classification algorithms.

Following the use of machine learning model models and well-known features like TF-IDF and bag of words (BoW), existing studies focus on optimizing the performance of models with hyperparameter tuning. Shu et al. [3] proposed a system for fake news collection, detection, and visualization. The authors utilize fake news and genuine news articles from fact-checking websites and social media. For the development of the fake news detection model, they extracted news and social media interaction features. Comparative analysis of the discussed studies is presented in table 1 to provide an analytical overview. Approaches, datasets, findings, and future work of past studies are presented in this table.

The aforementioned studies did not use any combination of feature extraction and feature selection techniques for the detection of COVID-19-related fake news. Therefore, this study analyzes different combinations of feature extraction and feature selection techniques. Two feature extraction techniques include TF-IDF and BoW while feature selection techniques include PCA and Chi-square.

III. MATERIALS AND METHODS
This study works on the detection of COVID-19 fake news detection. Experiments in this work are carried out using feature extraction (TF-IDF and BoW) and feature selection techniques (PCA and Chi-square) in combination with several machine learning and deep learning models. This section describes the dataset, proposed methodology, and other approaches used for comparison for fake news detection.

A. HYPOTHESIS
This study formulates the null hypothesis \((H_0)\) and alternative hypothesis \((H_A)\) as follows

- \(H_0\): Selecting important features from PCA and Chi-square and preprocessing does not affect the performance of the machine learning models,
- \(H_A\): Use of selective features from PCA and Chi-square has a positive effect on models’ performance.

B. DATA COLLECTION
The COVID-19 fake news dataset is obtained from the IEEE data port [47]. On the Twitter platform, different keywords and hashtags are used for text extraction to ensure that textual data is related to COVID-19. Table 2 shows a list of keywords and hashtags used to collect the data. Topics are categorized and grouped using keywords, making it simple for users to use the search tool to find similar topics. The dataset contains several attributes like the ‘Date’, ‘Link’, ‘Text’, ‘Region’, ‘Country’, ‘Explanation’, ‘Origin’, ‘Origin_URL’, ‘Fact_checked_by’, ‘Poynter_Label’, and ‘Label’.

C. PROPOSED METHODOLOGY
This study investigates various feature engineering approaches in combination with machine learning classifiers. The proposed architecture for COVID-19 fake news detection is presented in Figure 1 which shows the sequence of steps performed in experiments. The dataset is obtained from the IEEE data port and preprocessed before the feature extraction. Data is split in the ratio of 70:30 for training and test set. For feature extraction, TF-IDF and BoW are used. To reduce the training time, subset feature selection, PCA and Chi-square are used before models’ training. To obtain more optimized results different Feature extraction and feature selection techniques are tested in various combinations to train the models. ML models include RF, ET, GMB, LR, NB, SG, and VC(LR+SGD) and are compared with deep learning models including CNN, LSTM, ResNet, and Inception V3. In the end, evaluation is carried out using accuracy, precision, recall, f1-score, specificity, and AUC.

D. PREPROCESSING
Tweets contain unstructured, short, and noisy data which needs to be cleaned before it can be used for classification. For removing noise and improving the performance of models, several preprocessing steps are carried out.

1) Social media posts contain hashtags to relate it to a specific topic such as #covid-19, #lockdown. These hashtags are unnecessary in terms of sentiments, so they must be removed.
2) To avoid confusion in recognizing the same word differently by a model because of capitalization. All capital letters are converted to lower case.
3) In order to reduce noise from the tweets, stopwords such as ‘a’, ‘for’, ‘the’, ‘an’, ‘is’, etc and numeric values are removed from the data as these values do not affect sentiment. Alphanumeric values such as COVID19 are not removed.
4) Stemming is applied to change the terms to their root forms. It reduces the feature complexity and improves the performance of models.
5) Punctuation and special characters have also been removed.

E. FEATURE EXTRACTION
Feature engineering aims at finding appropriate features from the data to obtain good results from the models. Feature engineering helps to enhance the consistency and accuracy of the learning algorithm because feature engineering extracts the meaningful feature from the raw data. In this research, Vectorization (TF-IDF), prediction-based (Bag of Word (BoW)), dimensionality reduction (PCA), and variance analysis (Chi-square) techniques are used.

TF-IDF can be used to find the similarity between documents easily. It counts the occurrence of the word in a
TABLE 1. Comparative analysis of the approaches from the literature. These approaches are used for fake news detection from recently published works and the approach, dataset, findings, and future works are discussed.

| Ref. | Year | Methods/Approach | Dataset | Findings | Future work |
|------|------|------------------|---------|----------|-------------|
| [27] | 2020 | 23 Algorithms    | BuzzFeed Political News Data set, Random Political News Data set, ISOT Fake News Data set | Supervised Machine learning model and text analysis are proposed for fake news detection. | Ensemble model and other feature extraction methods will be tested to improve the results. |
| [5]  | 2020 | Ten Machine learning Algorithms along with seven feature extraction techniques | COVID-19 related tweets | A voting ensemble is proposed and 5-fold cross-validation is applied to validate the performance of the proposed framework | Will cover the data from different other languages to detect misleading information. |
| [47] | 2019 | Machine learning models and content-based features | UNBiased dataset | Ensemble algorithm along with word embedding is proposed to classify fake news | Metadata to verify the source of news will be increased and will be tested using deep learning model. |
| [42] | 2020 | Deep learning model | COVID-19 related news articles | Deep learning model like LSTM is efficient in dealing with inconsistent data | Hybrid model will be used in future to improve the final classification. |
| [3]  | 2019 | FakeNewsTrackers using Social Article Fusion (SAF) model | Fake News dataset (PolitiFact, BuzzFeed) | Use of linguistic features, data visualization, and deep learning models for fake news detection has shown robust results. | Retweets, social networks, and favorites will be considered in future work. |
| [19] | 2021 | Particle swarm optimization, the genetic algorithm, and the salp swarm algorithm | Koalala dataset | By reducing the number of features, the proposed approach has obtained improved results | In the future, the proposed approach will be applied in educational and business sectors on large datasets. |
| [45] | 2021 | Salp Swarm Optimization, Grey Wolf Optimization | Political fake news datasets(BuzzFeed, DTM, LIAR, ISOT) | Proposed model combines optimization techniques with text mining methods to detect fake news on social media and achieved promising results. | Meta-heuristic approaches will be tested with hybrid models. |

FIGURE 1. Architecture of the proposed approach. The approach follows data collection, preprocessing, data splitting, feature extraction, model training, and evaluation.

document as well as in a given corpus. Weight is directly proportional to a document’s word frequency and inversely proportional to words’ frequency within documents. TF-IDF shows good results for text classification, however, is limited
by large vector size and sparsity problems. It does not con-consider the position and co-occurrence of terms and it also does not take into account the semantics and context. Additionally, it is unable to find similarities between synonyms and differentiate in polysemy words. BOW gives the frequency of terms in a given corpus by considering word similarities. It maps words to target vectors. Continuous BOW (CBOW) predicts the probability of words and skip-gram predicts the context of words. It is a very simple feature method that considers word similarities alone and with a large vocabulary, the feature vector can be sparse.

PCA is a versatile technique that works well for many practical applications. It has a very simple and fast implementa-mentation. PCA offers several extensions (i.e., kernel PCA, sparse PCA, etc.) to handle specific roadblocks. However, new principal components are not interpretable, which may be a deal-breaker in some settings. PCA must still manually set or tune a threshold for cumulative explained variance. Chi-square has the strength of being easier to compute than several other statistical approaches. It is mostly used with categorical data to see if there is a difference between two or more groups of participants. Chi-square requires participants to be independent, indicating that one sample cannot fit in more than one category. If a sample can fit into two catego-ries a chi-square analysis is not appropriate. Another limita-tion of using chi-square is that the data must be frequency data.

F. MACHINE LEARNING CLASSIFIERS

For COVID-19 fake news detection, many machine learn-ing classification algorithms are used in this study. A brief description of each classifier is provided here.

NB is a supervised algorithm that uses conditional proba-bility theorem to find the class of a new sample [48]. For a given class, NB determines the conditional probability of a vector by the training dataset. After calculating the condi-tional probability value for each vector, the new vector class is calculated based on its conditional probability. NB is widely used for text-based classification problems.

RF is a tree-based ensemble learning model which predicts by combining several weak learners. RF uses bagging to train decision trees by utilizing different bootstrap samples. These bootstrap samples are obtained from the sub-sampling of the training dataset with replacement in which the size of the sample is the same as the size of the training dataset [49]. While constructing the decision tree in an RF the major issue is the identification of attributes for the root node at every level. This process is known as attribute selection [50]. By subsampling the training dataset with a replacement boot-strap, the sample is derived in which the size of the sample is the same as that of the training dataset.

ET is an ensemble learning classifier that aggregates the outcomes of multiple de-correlated decision trees [51]. The ET works quite similar to the RF but varies for the con-struction of decision trees within a forest. Every tree is given a random sample with K-features from the feature set in which every decision tree selects the best feature for splitting data based on Gini Index. Multiple de-correlated decision trees are created by these random samples. ET clas-sifier generates multiple decision trees to learn the patterns in the training data. These trees help in the prediction of the test data and then voting is performed for the final prediction.

GBM is a group of machine learning classifiers that com-bine many weak learning classifiers to make a powerful learn-ing model [52]. When doing gradient boosting usually deci-sion trees are used. GBM develops every tree independently so, it is a time-consuming and costly choice. GBM enhances the learning algorithm strength which is termed as probabil-ity approximating correct learning (PAC). PAC gives notable results on the unprocessed data. GBM deals with missing values efficiently.

LR is a statistical method that is used to analyze the data where one or more than one variable is used to find the final result. LR is used to estimate the probability of the class members. So, LR is the best choice when the target class is categorical [53]. It processes the connection between the categorical dependent variables and one or more independent variables by estimating probabilities using a logistic func-tion. LR gives promising results for binary classification. The sigmoid function is used to predict the probability values. It maps the values between 0 and 1.

The concept of SG is based on the working principle of SVM and logistic regression convex loss functions [54]. Due to its quality of combining multiple binary classifiers in One-vs-All (OvA), SG is a powerful algorithm to deal with multi-class classification problems. SG is the best choice for large datasets as it takes only a single example per iteration. SG is based on the simple regression technique so, it is easy to implement and easy to understand. Contrarily, SG is a nois-y choice because the examples selected from the batch are ran-dom as well as the hyperparameters of SG need to be correctly valued to get the best results. SG has a high sensitivity value in terms of feature scaling.

A VC is an ensemble learning technique and this study combines LR and SG for the ensemble. The voting classifier is constituted of multiple classifiers and a single regression model is used to compute the voting results. Each classifier in the voting classifier predicts its target label and voting is done between the classifiers to predict the final results. This study employs the hard or majority voting where the final class is predicted based on the majority votes. The class with the highest weighted probability wins. Details of hyperpara-meters used for experiments are presented in Table 3. The performance of the model being trained is greatly impacted.

| Keywords             | Hash tags                    |
|----------------------|-----------------------------|
| corona, coronavirus, covid, covid19, sarcov2, ncov, ncov2019, 2019-ncov, pandemic, quarantine, flatten the curve, lockdown, social distancing, work from home, wuhan virus, corona vaccine | #corona, #sarcov2, #covid_19, #ncov, #flatteningthecurve, #w05, #workfromhome, #healthworker, #vaccine, #homeschooling, #hometasking, #stayhome, #selfisolating |
by the hyperparameters, which directly determine how the training algorithm behaves.

### TABLE 3. Hyperparameter setting of ML models. These hyperparameters are obtained using the GridsearchCV methods and models obtain optimized results using these hyperparameters.

| Classifiers | Parameters |
|-------------|------------|
| RF          | n_estimators=200, max_depth=30, random_state=52 |
| ET          | n_estimators=200, max_depth=30, random_state=52 |
| GBM         | n_estimators=200, max_depth=30, random_state=52, learning_rate=0.1 |
| LR          | penalty='l2', solver='lbfgs' |
| NB          | alpha=1.0, binaryize=0.0 |
| SG          | penalty='l2', loss='log' |
| VC (LR+SG)  | voting='soft' |

### G. DEEP LEARNING CLASSIFIERS

The deep learning classifiers are also used for comparison in this study for fake news detection. The classifier used in experiments are CNN, LSTM, Residual Network (ResNet), and Inception v3.

CNN [55] is a popular deep learning model for image processing. The data’s best features are provided by CNN, which also has a low computational cost. It is made up of a flattening, pooling, and convolutional layer. With the use of various kernels, the convolutional layer, which is the bottom layer, convolvs the features to produce high-level features. The pooling layer helps in lowering feature size, which shortens the model’s calculation time. The three most popular pooling procedures are maximum, minimum, and average. In the end, the output for the required number of classes is obtained using the fully connected layer.

An artificial neural network called LSTM [56] is employed in deep learning and artificial intelligence. LSTM features feedback connections as opposed to typical feedforward neural networks. Such a recurrent neural network (RNN) may analyze whole data sequences (speech or video) in addition to single data points (such as photos). Since there may be unexpected delays between significant occurrences in a time series, LSTM networks are well-suited to categorizing, processing, and generating predictions. The vanishing gradient issue that might arise when training conventional RNNs is the reason why LSTMs were created. A cell, an input gate, an output gate, and a forget gate make up a typical LSTM unit. The three gates control the information flow (in and out of the cell), and the cell remembers values across arbitrary time periods.

ResNet [57] is a robust backbone model often employed in various computer vision tasks. ResNet employs skip connections to transfer output from one layer to another. This aids in reducing the issue of disappearing gradients. ResNets operate on the premise of building deeper networks than conventional simple networks while also determining the optimal amount of layers. It is simple to train networks with several layers (even thousands), without raising the training error percentage.

Inception-v3 proposed by Szegedy et al. [58] is a 48-layer convolutional neural network. A pre-trained version of the network that has been trained on more than a million photos is available for loading from the ImageNet database. A CNN’s Inception Module is a block of the image model that attempts to simulate an ideal local sparse structure. To put it simply, it enables us to employ numerous filter sizes in a single picture block rather than being limited to single filter size.

### H. EVALUATION PARAMETERS

The effectiveness of a machine learning model is measured using evaluation metrics. To evaluate the performance of machine learning models, this study utilized the following evaluation measures: accuracy, precision, recall, F1 score, specificity, and area under the curve (AUC). Following equations are used to calculate these measures.

\[
\text{Accuracy} = \frac{\text{Number of correctly classified predictions}}{\text{Total Predictions}}
\]

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

\[
\text{Recall} = \frac{TP}{TP + FN}
\]

\[
F1 - \text{Score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}
\]

\[
\text{Specificity} = \frac{TN}{FP + TN}
\]

where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively.

### IV. RESULTS AND DISCUSSION

Extensive experiments are performed using various machine learning models for fake news detection related to COVID-19. Experiments are performed covering several aspects in this regard. For example, the performance of models is tested with TF-IDF and BOW using each PCA and Chi-square for analysis. For deep learning models, Global Vectors (GloVe) and FastText word embedding approaches are used. Additionally, the performance of various preprocessing types is evaluated including full preprocessing, partial preprocessing, and no preprocessing. For partial preprocessing, the step of case conversion is not carried out as some research works to point out that capital letter words may be an indicator of fake news. For experiments, the data split ratio is 0.7 to 0.3 for training and testing. For performance, specificity and area under the curve (AUC) are used beside traditional parameters of accuracy, precision, recall, and F1 score.

### A. COMPARISON OF ML-BASED MODELS USING TF-IDF AND BOW

The results of ML-based models using TF-IDF and BOW are presented in Table 4. Results indicate that the ET classifier gives the best results when using TF-IDF and PCA. It achieves a score of 0.9474 for accuracy while 0.95 each for precision, recall, and F1 score and 0.94 for specificity.
and AUC. SG also shows good performance, however, it is slightly lower than ET with an accuracy of 0.9401 and 0.94 value each for precision, recall, and F1 score. RF and VC achieve accuracies higher than 0.90 for all evaluation measures. LR and NB show similar results with accuracy values of 0.8991 and 0.8892, respectively which are the lowest among all classifiers for fake news detection.

**TABLE 4.** Performance comparison of ML models using TF-IDF and BoW. Acc., Pre. and Sp. indicates accuracy, precision, and specificity, respectively.

| Model | Acc. | Pre. | Recall | F1 | Sp. | AUC |
|-------|------|------|--------|----|-----|-----|
| RF    | 0.9199 | 0.92 | 0.92  | 0.92 | 0.91 | 0.91 |
| ET    | 0.9474 | 0.95 | 0.95  | 0.94 | 0.94 | 0.94 |
| GBM   | 0.8653 | 0.88 | 0.87  | 0.86 | 0.85 | 0.85 |
| LR    | 0.8991 | 0.91 | 0.90  | 0.89 | 0.90 | 0.91 |
| NB    | 0.8892 | 0.89 | 0.89  | 0.88 | 0.87 | 0.90 |
| SG    | 0.9401 | 0.94 | 0.94  | 0.94 | 0.95 | 0.93 |
| VC (LR+SG) | 0.9002 | 0.91 | 0.90  | 0.90 | 0.92 | 0.91 |

**B. COMPARISON OF ML MODELS WITH TF-IDF AND PCA**

Similar to the use of models with TF-IDF and BOW, separate experiments are done using TF-IDF with PCA for fake news detection. PCA is a dimensionality reduction tool and is mostly used to reduce the dimensions of a large-sized dataset and transform large variable sets into smaller ones. It retains important information by analyzing interrelation among variables. Table 5 presents the comparison of models using TF-IDF and PCA for fake news detection about COVID-19. It can be observed that the performance of models using TF-IDF and PCA is poor than the performance achieved with TF-IDF and BOW. However, ET still achieves the highest results with a 0.9305 accuracy and a score of 0.93 each for precision, recall, and F1 score. ET shows 0.91 score for specificity and 0.92 for AUC. Other models like RF, GBM, LR, NB, SG, and VC do not show any improvement in the results.

**TABLE 5.** Performance comparison of ML models using TF-IDF and PCA. Results are obtained using PCA-based TF-IDF features. Acc., Pre. and Sp. indicates accuracy, precision, and specificity, respectively.

| Model | Acc. | Pre. | Recall | F1 | Sp. | AUC |
|-------|------|------|--------|----|-----|-----|
| RF    | 0.8974 | 0.90 | 0.90  | 0.89 | 0.87 | 0.89 |
| ET    | 0.9305 | 0.93 | 0.93  | 0.93 | 0.91 | 0.92 |
| GBM   | 0.8527 | 0.86 | 0.85  | 0.83 | 0.84 | 0.83 |
| LR    | 0.8839 | 0.89 | 0.88  | 0.87 | 0.86 | 0.86 |
| NB    | 0.8875 | 0.89 | 0.89  | 0.88 | 0.88 | 0.90 |
| SG    | 0.9243 | 0.92 | 0.92  | 0.92 | 0.91 | 0.93 |
| VC (LR+SG) | 0.9146 | 0.89 | 0.88  | 0.87 | 0.86 | 0.88 |

**C. RESULTS OF MODELS USING BOW AND PCA**

The performance of ML-based models has been evaluated and compared using BOW and PCA for fake news detection. Results presented in Table 6 indicate that ML-based models achieve better results with the combination of BOW and PCA as compared to the combination of TF-IDF and PCA. The highest accuracy of 0.9407 is achieved with the combination of BOW and PCA. SG achieves the second-highest value of accuracy which is 0.9379. ET and SG show similar results in terms of precision, recall, and F1 score with a value of 0.94 which is the highest among all models. RF and VC show similar results for fake news detection. However, NB shows the worst results among all models even lower than those achieved with TF-IDF and PCA and TF-IDF and BOW.

**TABLE 6.** Performance comparison of ML models using BOW and PCA. Results are obtained using PCA-based BOW features. Acc., Pre. and Sp. indicates accuracy, precision, and specificity, respectively.

| Model | Acc. | Pre. | Recall | F1 | Sp. | AUC |
|-------|------|------|--------|----|-----|-----|
| RF    | 0.9232 | 0.93 | 0.93  | 0.92 | 0.91 | 0.92 |
| ET    | 0.9407 | 0.94 | 0.94  | 0.94 | 0.92 | 0.93 |
| GBM   | 0.8603 | 0.88 | 0.86  | 0.85 | 0.87 | 0.88 |
| LR    | 0.9249 | 0.93 | 0.92  | 0.92 | 0.88 | 0.85 |
| NB    | 0.8788 | 0.88 | 0.88  | 0.87 | 0.84 | 0.87 |
| SG    | 0.9379 | 0.94 | 0.94  | 0.94 | 0.93 | 0.95 |
| VC (LR+SG) | 0.9241 | 0.93 | 0.92  | 0.92 | 0.91 | 0.94 |

**D. MODELS’ COMPARISON USING BOW AND CHI-SQUARE**

This section presents another combination of features including BOW and Chi-square for fake news detection. Chi-square is a statistical method and focuses on highly dependent features for the target variable. It can be seen from the results given in Table 7 that results obtained from the combination of BOW and Chi-square are very similar to the results obtained from the combination of TF-IDF and PCA. In this scenario, ET outperforms with a 0.9274 accuracy. Its values for precision, recall, and F1 score are also the best among all models, however, its specificity and AUC values are comparatively lower than that of LR.

**TABLE 7.** Performance comparison of ML models using BOW and Chi-square. Results are obtained using Chi-square extracted BOW features. Acc., Pre. and Sp. indicates accuracy, precision, and specificity, respectively.

| Model | Acc. | Pre. | Recall | F1 | Sp. | AUC |
|-------|------|------|--------|----|-----|-----|
| RF    | 0.9131 | 0.91 | 0.91  | 0.91 | 0.91 | 0.90 |
| ET    | 0.9274 | 0.93 | 0.93  | 0.93 | 0.91 | 0.90 |
| GBM   | 0.8532 | 0.86 | 0.85  | 0.83 | 0.84 | 0.83 |
| LR    | 0.9168 | 0.92 | 0.92  | 0.91 | 0.93 | 0.93 |
| NB    | 0.8831 | 0.88 | 0.88  | 0.87 | 0.86 | 0.87 |
| SG    | 0.9238 | 0.92 | 0.92  | 0.92 | 0.90 | 0.93 |
| VC (LR+SG) | 0.8921 | 0.92 | 0.92  | 0.92 | 0.90 | 0.89 |

**E. RESULTS COMPARISON USING PCA AND CHI-SQUARE**

Supervised ML-based models have been evaluated using a combination of feature subset of PCA and Chi-square. From Table 8, it can be seen that the combination of PCA and Chi-square substantially degraded the performance of all classifiers for fake news detection. The highest performing model is still ET in this combination but it achieves a 0.8864 value of accuracy, 0.90 value of Precision, 0.89 value of Recall, and 0.88 value of F1 score. GBM, LR, SG, and VC models show results lower than the 0.85 value for each evaluation parameter. NB shows the worst performance with a 0.6894 accuracy, 0.72 precision, 0.69 recall, and 0.70 F1 scores. Existing studies report that using Chi-square and PCA
together improves the computations, however, it reduces the performance [59]. Results in this study corroborate the same.

**TABLE 8. Performance comparison of ML models using PCA and Chi-square. Acc., Pre. and Sp. indicates accuracy, precision, and specificity, respectively.**

| Model          | Acc. | Pre. | Recall | F1    | Sp. | AUC  |
|----------------|------|------|--------|-------|-----|------|
| RF             | 0.8743 | 0.88 | 0.87   | 0.87  | 0.85 | 0.89 |
| ET             | 0.8864 | 0.90 | 0.89   | 0.88  | 0.89 | 0.85 |
| GBM            | 0.8446 | 0.85 | 0.84   | 0.84  | 0.83 | 0.88 |
| LR             | 0.8221 | 0.82 | 0.82   | 0.82  | 0.83 | 0.80 |
| NB             | 0.6894 | 0.72 | 0.70   | 0.71  | 0.74 |     |
| SG             | 0.8336 | 0.83 | 0.83   | 0.83  | 0.82 | 0.87 |
| VC(LR+SG)      | 0.8241 | 0.82 | 0.82   | 0.82  | 0.85 | 0.88 |

**F. PERFORMANCE COMPARISON OF MODELS USING TF-IDF AND CHI-SQUARE**

Finally, another combination of feature subsets has been evaluated with models for fake news detection and it includes TF-IDF and Chi-square. Results presented in Table 9 indicate that TF-IDF and Chi-square show inferior performance as compared to the performance of the combination of PCA and Chi-square. ET achieves a 0.8817 accuracy, 0.89 precision, 0.88 recall, 0.87 F1 score, 0.90 specificity and 0.87 AUC. The height precision, recall, F1 score, specificity, and AUC are obtained by LR despite its low accuracy of 0.8209. It is clear from the results that the combination of the feature subset of TF-IDF and Chi-square is not contributing to a better performance of models for fake news detection.

Figure 2 shows the comparison of all models with a different subset of features such as TF-IDF with BOW, TF-IDF with PCA, BOW with PCA, BOW with Chi-square, PCA with Chi-square, and TF-IDF with Chi-square. It shows the superiority of ET over all other classifiers for each combination of feature subsets. ET surpasses every combination of features subset when it is applied with TF-IDF and BOW where it obtains a 0.9474 accuracy, 0.95 each for precision, recall, and F1 score. AUC value is 0.94 which is also the highest among all models. Only its specificity of 0.94 is second to SG which obtains a specificity of 0.95. Finally, results reveal that the effectiveness of models depends upon the appropriate combination of feature subsets. For the analysis of text data feature reduction techniques such as PCA and Chi-square, no improvements are observed. It seems that classifiers are not trained well when features are reduced.

The tree-based algorithm, ET shows robust results when trained on preprocessed data using TF-IDF and BOW for fake news prediction. ET is an extremely randomized tree classifier and selects a random cut-point from a random subspace. If the randomization level in ETC is adjusted properly then variance vanishes while bias increases according to the trees. The tree-based model generalizes better and avoids overfitting and outperformed other models when combined with an appropriate combination of feature subsets.

**TABLE 9. Performance comparison of ML models using TF-IDF and Chi-square. Results are obtained using TF-IDF features extracted using Chi-square. Acc., Pre. and Sp. indicates accuracy, precision, and specificity, respectively.**

| Model       | Acc.     | Pre. | Recall | F1    | Sp. | AUC  |
|-------------|----------|------|--------|-------|-----|------|
| RF          | 0.8783   | 0.88 | 0.88   | 0.87  | 0.86 | 0.89 |
| ET          | 0.8817   | 0.89 | 0.88   | 0.87  | 0.90 | 0.87 |
| GBM         | 0.8521   | 0.85 | 0.85   | 0.84  | 0.81 | 0.82 |
| LR          | 0.8209   | 0.92 | 0.92   | 0.91  | 0.90 | 0.94 |
| NB          | 0.6616   | 0.81 | 0.82   | 0.81  | 0.84 | 0.79 |
| SG          | 0.8288   | 0.82 | 0.82   | 0.81  | 0.80 | 0.83 |
| VC(LR+SG)   | 0.8224   | 0.81 | 0.82   | 0.81  | 0.84 | 0.79 |

**TABLE 10. Performance comparison of deep learning models using Glove features. Acc., Pre. and Sp. indicates accuracy, precision, and specificity, respectively.**

| Model        | Acc.     | Pre. | Recall | F1     | Sp.  | AUC  |
|--------------|----------|------|--------|--------|------|------|
| CNN          | 0.9231   | 0.90 | 0.91   | 0.90   | 0.92 | 0.92 |
| LSTM         | 0.8965   | 0.85 | 0.89   | 0.86   | 0.87 | 0.90 |
| ResNet       | 0.8535   | 0.85 | 0.87   | 0.86   | 0.89 | 0.84 |
| InceptionV3  | 0.9124   | 0.90 | 0.89   | 0.89   | 0.87 | 0.87 |

**TABLE 11. Performance comparison of deep learning models using FastText. Acc., Pre. and Sp. indicates accuracy, precision, and specificity, respectively.**

| Model        | Acc.     | Pre. | Recall | F1     | Sp.  | AUC  |
|--------------|----------|------|--------|--------|------|------|
| CNN          | 0.9341   | 0.92 | 0.93   | 0.92   | 0.90 | 0.91 |
| LSTM         | 0.9188   | 0.90 | 0.90   | 0.90   | 0.93 | 0.94 |
| ResNet       | 0.9215   | 0.93 | 0.90   | 0.91   | 0.90 | 0.90 |
| InceptionV3  | 0.9241   | 0.92 | 0.90   | 0.91   | 0.94 | 0.91 |

**G. PERFORMANCE OF DEEP LEARNING MODELS**

This study also uses several state-of-the-art deep learning models for performance comparison. In this regard, two kinds of deep learning models are considered. Two custom-built models CNN and LSTM are used while two pre-trained deep learning models ResNet and InceptionV3 are also included in the experiments. Table 10 shows the results obtained for fake news detection using the GloVe features. Results indicate that the custom-built CNN models can detect fake news with a 0.92 accuracy score followed by the Inception V3 which has a 0.91 accuracy score. Other parameters like precision, recall, and F1 scores are also better than LSTM and ResNet. Comparatively, the performance of deep learning models is low as compared to the best performing ET and SG which show the best performance using selective features from Chi-square.

Similar to using the GloVe features, deep learning models are used with FastText as well and results are provided in Table 11. Results indicate that the performance of models has been improved using the FastText features. For example, the accuracy score of CNN has increased from 0.92 to 0.93 and InceptionV3 shows an accuracy score of 0.92 which is better than its performance with GloVe. Similarly, the performance of LSTM and ResNet has also been improved with FastText. Despite that, the performance of CNN is relatively less than machine learning models with selective features.

**H. INFLUENCE OF PREPROCESSING STEPS ON MODELS’ PERFORMANCE**

Existing studies suggest that the performance of machine learning models is greatly influenced by the use of differ-
ent preprocessing steps. For example, using partial preprocessing where few steps are not included like removing #, emoticons, or case conversion. In this perspective, the use of upper case letters has been reported important for fake news detection. Given this information, the current study utilizes three different preprocessing techniques where models are used with full preprocessing, without preprocessing, and partial preprocessing. For partial preprocessing, all steps are performed but the case conversion. Several studies suggest that words with capital words may be used as a characteristic for fake news detection [60], [61], as a result, case conversion is not carried out to analyze the accuracy of models. Experimental results suggest that the performance of models is best when complete preprocessing is used before feature extraction. Without preprocessing the accuracy of the best performing ET model is reduced to 0.7991 indicating degradation of 15.65%. Similarly, using partial preprocessing, the accuracy degrades to 0.8741 for fake news detection which is a decrease of 7.73%. The same pattern is found for other models as well which show poor performance when trained on the data without preprocessing or partial preprocessing.

I. RESULTS OF K-FOLD CROSS-VALIDATION

K-fold cross-validation is carried out to validate the performance of models. For this purpose, the best performing model ET is used with the best subset feature set of TF-IDF and BOW. Results of 10-fold cross-validation are given in Table 13. It indicates that the performance of the ET model is superior regarding the accuracy, precision, recall, and F1 score with a small standard deviation.

J. COMPARISON WITH EXISTING STUDIES

For a fair comparison, studies that use the same dataset for experiments have been selected. Three studies use the CoVID19-FNIR dataset for experiments involving both machine and deep learning models. Reference [62] uses SVM, K nearest neighbor (KNN), decision tree (DT), RF,
and NB for fake news detection, with the highest accuracy of 0.90 by SVM. Similarly, the authors utilize several machine learning models for fake news detection in [63] including LR, linear SVM (LSVM), DT, RF, KNN, NB, and SG. In addition, deep learning-based recurrent neural networks (RNN), LSTM, and gated recurrent unit are adopted. The best performance is observed using LSVM which obtains 0.9411 accuracy while LSTM provides an accuracy of 0.92. The third study does not perform fake news detection, rather it focuses on finding the class and subclass similarity of various articles/news and provide the similarity score [64]. The current study, however, provides better performance concerning the accuracy and superior performance concerning precision, recall, and F1 score than existing studies. Results reveal that using selective features from both BoW and TF-IDF to train machine learning models yield better results than using either of these features alone. This accuracy can further be enhanced provided an appropriate preprocessing phase is carried out.

### TABLE 14. Performance comparison with existing studies. Models from existing studies are implemented on the current dataset for a fair comparison.

| Ref. | Model | Accuracy | Precision | Recall | F1  |
|------|-------|----------|-----------|--------|-----|
| [62] | SVM   | 0.9001   | 0.90      | 0.90   | 0.90|
| [63] | LSVM  | 0.9411   | 0.94      | 0.94   | 0.94|
| [63] | LSTM  | 0.9210   | 0.92      | 0.91   | 0.92|
| Current study | ET | 0.9474   | 0.95      | 0.95   | 0.95|

Table 15 shows the performance of machine learning and deep learning models regarding the execution time. It shows the average execution time for all features used with the machine learning models.

### TABLE 15. Average execution time of ML and DL models using all features.

| Model       | Average execution time (seconds) |
|-------------|----------------------------------|
| RF          | 80-85s                           |
| ET          | 68-72s                           |
| GBM         | 110-113s                         |
| LR          | 50-54s                           |
| NB          | 64-69s                           |
| SG          | 51-55s                           |
| VC(LR+SG)   | 61-67s                           |
| CNN         | 153-157s                         |
| LSTM        | 225-247s                         |
| ResNet      | 177-182s                         |
| InceptionV3 | 259-277s                         |

### K. STATISTICAL T-TEST

The statistical t-Test has also been performed to show the significance of the proposed approach (Extra Tree Classifier with TF-IDF and BOW feature combination). In the null hypothesis of the t-test, $H_0$ shows that the accuracy difference of methods is not significant while alternate hypothesis $H_a$ shows that the accuracy difference is significant. We have performed a t-test of the proposed model and the second-best performing feature combination (BOW and PCA) with Extra Tree Classifier (ET) as the performing model. This test shows a 9.28251 value for test statistics and 0.001439 $p$-value. It concludes that the proposed model has improved the performance. Secondly, the test is performed on the proposed approach with the previous best study LSTM approach [63]. The results show a 5.5629 value for test statistics and 0.007410 $p$-value. It also proves that the proposed model has improved the performance. Results prove that the difference is statistically significant with $p < 0.05$. The proposed model obtained the highest mean rank for accuracy.

### TABLE 16. Performance comparison of ML models using TF-IDF and BOW on new dataset.

| Model       | Acc. | Prec. | Recall | F1  | Sp. | AUC   |
|-------------|------|-------|--------|-----|-----|-------|
| RF          | 0.8794 | 0.91  | 0.91   | 0.89 | 0.89| 0.87  |
| ET          | 0.9603 | 0.92  | 0.92   | 0.92 | 0.91| 0.93  |
| GBM         | 0.8752 | 0.87  | 0.86   | 0.86 | 0.88| 0.85  |
| LR          | 0.8938 | 0.90  | 0.89   | 0.89 | 0.87| 0.86  |
| NB          | 0.8678 | 0.91  | 0.88   | 0.89 | 0.84| 0.87  |
| SG          | 0.9424 | 0.90  | 0.94   | 0.95 | 0.89| 0.90  |
| VC(LR+SG)   | 0.8138 | 0.90  | 0.90   | 0.90 | 0.86| 0.85  |

### V. CONCLUSION

Fake news presents a challenging problem and its importance has been elevated during the COVID-19 outbreak. Despite several existing approaches, studies investigating the importance of subset feature selection for fake news detection are very few. Hence, a feature-based approach is presented in this study where different combinations of feature engineering and feature selection approaches are investigated. The impact of selecting TF-IDF, BoW, PCA, and Chi-square is analyzed regarding fake news detection. The authors find three observations from experimental results. First, selective features tend to yield better results than using all features. The use of TF-IDF and BoW features combined produce better results than PCA and Chi-square selected features. Secondly, the performance of pre-trained deep learning models ResNet and InceptionV3 is marginally lower than machine learning models. For parameter optimization, such models require larger datasets to show better performance. Thirdly, preprocessing is very important to obtain high accuracy. Full preprocessing produces better results than no preprocessing or partial preprocessing for text analysis. So, the best results are obtained using the full preprocessing where ET obtains a 0.9474 accuracy score for fake news detection. This study does not consider the distribution of classes in the dataset. It infers that an imbalanced class distribution can influence the performance of models. Authors intend to perform fake news detection by combining textual and stylometric features in the future.

### ACKNOWLEDGEMENT

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R192), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

### REFERENCES

[1] X. Zhang and A. A. Ghorbani, “An overview of online fake news: Characterization, detection, and discussion,” *Inf. Process. Manage.*, vol. 57, no. 2, Mar. 2020, Art. no. 102025.
[2] N. Ruchansky, S. Seo, and Y. Liu, “CSI: A hybrid deep model for fake news detection,” in Proc. ACM Conf. Inf. Knowl. Manage., Nov. 2017, pp. 797–806.

[3] K. Shu, D. Mahudeswaran, and H. Liu, “FakeNewsTracker: A tool for fake news collection, detection, and visualization,” Comput. Math. Org. Theory, vol. 25, no. 1, pp. 60–71, Mar. 2019.

[4] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi, and B.-W. On, “Fake news stance detection using deep learning architecture (CNN-LSTM),” IEEE Access, vol. 8, pp. 156695–156706, 2020.

[5] M. K. Elhadad, K. F. Li, and F. Gebali, “Detecting misleading information on COVID-19,” IEEE Access, vol. 8, pp. 165201–165215, 2020.

[6] M. Mhalla, “The impact of novel coronavirus (COVID-19) on the global oil and aviation markets,” J. Asian Sci. Res., vol. 10, no. 2, p. 96, 2020.

[7] C. Albulescu, “Coronavirus and oil price crash,” SSRN Electron. J., vol. 4, pp. 1–13, Apr. 2020.

[8] N. J. Gormsen and R. S. J. Koijen, “Coronavirus: Impact on stock prices and growth expectations,” Rev. Asset Pricing Stud., vol. 10, no. 4, pp. 574–597, Dec. 2020.

[9] M. Richtel, “Who fights a pandemic besides coronavirus: An ‘infoemic,’” The New York Times, New York, NY, USA, Tech. Rep., vol. 9, p. 6, 2020.

[10] M. Cinielli, W. Quattrocchi, A. Galeazzi, C. MicheleValensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, and A. A. Scala, “The COVID-19 social media infodemic,” Sci. Rep., vol. 10, no. 1, pp. 1–10, 2020.

[11] C. Shu and J. Schieber. (2020). Facebook, Reddit, Google, LinkedIn, Microsoft, Twitter and Youtube Issue Joint Statement

[12] M. M. Narra

[13] [VOLUME 10, 2022] 98735

[14] [9] M. Richtel, “WHO fights a pandemic besides coronavirus: An ‘infoemic,’” The New York Times, New York, NY, USA, Tech. Rep., vol. 9, p. 6, 2020.

[15] M. M. Narra

[16] [VOLUME 10, 2022] 98735

[17] [9] M. Richtel, “WHO fights a pandemic besides coronavirus: An ‘infoemic,’” The New York Times, New York, NY, USA, Tech. Rep., vol. 9, p. 6, 2020.

[18] M. M. Narra

[19] [VOLUME 10, 2022] 98735

[20] [9] M. Richtel, “WHO fights a pandemic besides coronavirus: An ‘infoemic,’” The New York Times, New York, NY, USA, Tech. Rep., vol. 9, p. 6, 2020.

[21] M. M. Narra

[22] [VOLUME 10, 2022] 98735

[23] [9] M. Richtel, “WHO fights a pandemic besides coronavirus: An ‘infoemic,’” The New York Times, New York, NY, USA, Tech. Rep., vol. 9, p. 6, 2020.
[50] L. Breiman, “Bagging predictors,” Mach. Learn., vol. 24, no. 2, pp. 123–140, 1996.

[51] A. Sharafi and H. Gupta, “Extra-tree classifier with metaheuristics approach for email classification,” in Advances in Computer Communication and Computational Sciences. Singapore: Springer, 2019, pp. 189–197.

[52] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” Ann. Statist., vol. 29, no. 5, pp. 1189–1232, Oct. 2001.

[53] R. E. Wright, Logistic Regression. Washington, DC, USA: Amer. Psychol. Assoc., 1995.

[54] W. A. Gardner, “Learning characteristics of stochastic-gradient-descent algorithms: A general study, analysis, and critique,” Signal Process., vol. 6, no. 2, pp. 113–133, Apr. 1984.

[55] J. Gu, Z. Wang, J. Kua, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cui, and T. Chen, “Recent advances in convolutional neural networks,” Pattern Recognit., vol. 77, pp. 354–377, May 2018.

[56] Y. Yu, X. Si, C. Hu, and Z. Jianxun, “A review of recurrent neural networks: LSTM cells and network architectures,” Neural Comput., vol. 31, no. 7, pp. 1235–1270, 2019.

[57] V. Feng, “An overview of resnet and its variants,” Towards Data Sci., vol. 2, Jul. 2017. Accessed: Mar. 2, 2022. [Online]. Available: https://towardsdatascience.com/an-overview-of-resnet-and-its-variants-5281e2f56035

[58] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 2818–2826.

[59] A. K. Gärke-Escamila, A. H. E. Hassani, and E. András, “Classification models for heart disease prediction using feature selection and PCA,” Informat. Med. Unlocked, vol. 19, May 2020, Art. no. 100330.

[60] M. Aldwairi and A. Alwahedi, “Detecting fake news in social media networks,” Proc. Comput. Sci., vol. 141, pp. 215–222, Jun. 2018.

[61] Abdullah, A. Yasin, M. Awan, F. H. Shehzad, and M. Ashraf, “Fake news classification bimodal using convolutional neural network and long short-term memory,” Int. J. Emerg. Technol. Learn., vol. 11, pp. 209–212, Aug. 2020.

[62] S. Almatarneh, P. Gamallo, B. Alshargabi, Y. Al-Khassawneh, and R. Alzubi, “Comparing traditional machine learning methods for COVID-19 fake news,” in Proc. 22nd Int. Arab Conf. Inf. Technol. (ACIT), Dec. 2021, pp. 1–4.

[63] R. Malhotra and A. Mahur, “COVID-19 fake news detection system,” in Proc. 12th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence), Jan. 2022, pp. 428–433.

[64] G. De Magistris, S. Russo, P. Roma, J. T. Starczewski, and C. Napoli, “An explainable fake news detector based on named entity recognition and stance classification applied to COVID-19,” Information, vol. 13, no. 3, p. 137, Mar. 2022.

**Manideep Narra** received the master’s degree in engineering management with information technology concentration from the Indiana Institute of Technology, Fort Wayne, IN, USA. He has worked in different domains including marketing services and healthcare. His accomplishments included designing, high quality, and deliverability with security. His current research interests include machine learning, artificial intelligence, and data mining.

**Muhammad Umer** received the B.S. degree from the Department of Computer Science, Khwaja Fareed University of Engineering and IT (KFUEIT), Pakistan, in October 2018. He is currently pursuing the Ph.D. degree in computer science with KFUEIT. He was a Research Assistant at the Fareed Computing and Research Center, KFUEIT. He is also working as a Lecturer at The Islamia University of Bahawalpur. His current research interests include data mining, mainly working on machine learning and deep learning-based IoT, text mining, and computer vision tasks.

**Saima Sadiq** is currently pursuing the Ph.D. degree in computer science with the Khwaja Fareed University of Engineering and IT (KFUEIT). She is working as an Assistant Professor at the Department of Computer Science, Government Degree College for Women. Her current research interests include data mining, machine learning, and deep learning-based text mining.

**Haneen Karamti** is an Associate Professor at Princess Nourah Bint Abdullah University. Her current research interests include data mining, mainly working in text information retrieval, text mining, and computer vision tasks.

**Abdullah Mohamad** is working as an Assistant Professor and a Research Head at Future University in Egypt. His current research interests include data mining, machine learning, and deep learning-based text mining.

**Imran Ashraf** received the M.S. degree in computer science from the Blekinge Institute of Technology, Karlskrona, Sweden, in 2010, and the Ph.D. degree in information and communication engineering from Yeungnam University, Gyeongsan, South Korea, in 2018. He has worked as a Postdoctoral Fellow at Yeungnam University. He is currently working as an Assistant Professor at the Information and Communication Engineering Department, Yeungnam University. His research interests include indoor positioning and localization, advanced location-based services in wireless communication, smart sensors (LiDAR) for smart cars, and data mining.