I**NTRODUCTION:**

)e advent of the World Wide Web and the rapid adoption of social media platforms (such as Facebook and Twitter) paved the way for information dissemination that has never been witnessed in the human history before. Besides other use cases, news outlets benefitted from the widespread use of social media platforms by providing updated news in near real time to its subscribers. )e news media evolved from newspapers, tabloids, and magazines to a digital form such as online news platforms, blogs, social media feeds, and other digital media formats [1]. It became easier for consumers to acquire the latest news at their fingertips. Facebook referrals account for 70% of traffic to news websites [2]. )ese social media platforms in their current state are extremely powerful and useful for their ability to allow users to discuss and share ideas and debate over issues such as democracy, education, and health. However, such platforms are also used with a negative perspective by certain entities commonly for monetary gain [3, 4] and in other cases for creating biased opinions, manipulating mindsets, and spreading satire or absurdity. )e phenomenon is commonly known as fake news

)ere has been a rapid increase in the spread of fake news in the last decade, most prominently observed in the 2016 US elections [5]. Such proliferation of sharing articles online that do not conform to facts has led to many problems not just limited to politics but covering various other domains such as sports, health, and also science [3]. One such area affected by fake news is the financial markets [6], where a rumor can have disastrous consequences and may bring the market to a halt. Our ability to take a decision relies mostly on the type of information we consume; our world view is shaped on the basis of information we digest. )ere is increasing evidence that consumers have reacted absurdly to news that later proved to be fake [7, 8]. One recent case is the spread of novel corona virus, where fake reports spread over the Internet about the origin, nature, and behavior of the virus [9]. )e situation worsened as more people read about the fake contents online. Identifying such news online is a daunting task

Fortunately, there are a number of computational techniques that can be used to mark certain articles as fake on the basis of their textual content [10]. Majority of these techniques use fact checking websites such as “PolitiFact” and “Snopes.” )ere are a number of repositories maintained by researchers that contain lists of websites that are identified as ambiguous and fake [11]. However, the problem with these resources is that human expertise is required to identify articles/websites as fake. More importantly, the fact checking websites contain articles from particular domains such as politics and are not generalized to identify fake news articles from multiple domains such as entertainment, sports, and technology

)e World Wide Web contains data in diverse formats such as documents, videos, and audios. News published online in an unstructured format (such as news, articles, videos, and audios) is relatively difficult to detect and classify as this strictly requires human expertise. However, computational techniques such as natural language processing (NLP) can be used to detect anomalies that separate a text article that is deceptive in nature from articles that are based on facts [12]. Other techniques involve the analysis of propagation of fake news in contrast with real news [13]. More specifically, the approach analyzes how a fake news article propagates differently on a network relative to a true article. )e response that an article gets can be differentiated at a theoretical level to classify the article as real or fake. A more hybrid approach can also be used to analyze the social response of an article along with exploring the textual features to examine whether an article is deceptive in nature or not.

A number of studies have primarily focused on detection and classification of fake news on social media platforms such as Facebook and Twitter [13, 14]. At conceptual level, fake news has been classified into different types; the knowledge is then expanded to generalize machine learning (ML) models for multiple domains [10, 15, 16]. )e study by Ahmed et al. [17] included extracting linguistic features such as n-grams from textual articles and training multiple ML models including K-nearest neighbor (KNN), support vector machine (SVM), logistic regression (LR), linear support vector machine (LSVM), decision tree (DT), and stochastic gradient descent (SGD), achieving the highest accuracy (92%) with SVM and logistic regression. According to the research, as the number of n increased in n-grams calculated for a particular article, the overall accuracy decreased. )e phenomenon has been observed for learning models that are used for classification. Shu et al. [12] achieved better accuracies with different models by combining textual features with auxiliary information such as user social engagements on social media. )e authors also discussed the social and psychological theories and how they can be used to detect false information online. Further, the authors discussed different data mining algorithms for model constructions and techniques shared for features extraction. )ese models are based on knowledge such as writing style, and social context such as stance and propagation. A different approach is followed by Wang [18]. )e author used textual features and metadata for training various ML models. )e author focused mainly on using convolutional neural network (CNN). A convolutional layer is used to capture the dependency between the metadata vectors, followed by a bidirectional LSTM layer. )e maxpooled text representations were concatenated with the metadata representation from the bidirectional LSTM, which was fed to fully connected layer with a softmax activation function to generate the final prediction. )e research is conducted on a dataset from political domain which contains statements from two different parties. Along with that, some metadata such as subject, speaker, job, state, party, context, and history are also included as a feature set. Accuracy of 27.7% was achieved with combination of features such as text and speaker, whereas 27.4% accuracy was achieved by combining all the different metadata elements with text. A competitive solution is provided by Riedel et al. [19], which is a stance detection system that assigns one of four labels to an article, “agree,” “disagree,” “discuss,” or “unrelated,” depending on the conformity of article headline with article text. )e authors used linguistic properties of text such as term frequency (TF) and term frequency-inverse document frequency (TF-IDF) as a feature set, and a multilayer perceptron (MLP) classifier is used with one hidden layer and a softmax function on the output of the final layer. )e dataset contained articles with a headline, body, and label. )e system’s accuracy on the “disagree” label on test examples was poor, whereas it performs best with respect to the “agree” label. )e authors used a simple MLP with some fine-tuned hyperparameters to achieve an overall accuracy of 88.46%. Shu et al. [12] also discussed several varieties of veracity assessment methods to detect fake news online. Two major categories of assessment methods are explored: one is linguistic cue approaches and the other is network analyses approaches. A combination of both creates a more robust hybrid approach for fake news detection online. Linguistic approaches involve deep syntax, rhetorical structure, and discourse analysis. )ese linguistic approaches are used to train classifiers such as SVM or na¨ıve Bayes models. Network-based approaches included analyzing and processing social network behavior and linked data. A unique approach is followed by Vosoughi et al. [13] to explore the properties of news spread on social media; i.e., the authors discussed the spread of news (rumors) on social media such as Twitter and analyzed how the spread of fake news differs from real news in terms of its diffusion on Twitter. Multiple analysis techniques are discussed in the paper to explore the spread of fake news online, such as the depth, the size, the maximum breadth, the structural virality, the mean breadth of true and false rumor cascades at various depths, the number of unique Twitter users reached at any depth, and the number of minutes it takes for true and false rumor cascades to reach depth and number of Twitter user