Stance detection with BERT embeddings for credibility analysis of information on social media

Hema Karande [1] , Rahee Walambe [Corresp., 2, 3] , Victor Benjamin [4] , Ketan Kotecha [Corresp., 1, 2] , Ts Raghu [4]

[1] Computer Science, Symbiosis Institute of Technology, Pune, Pune, Maharashtra, India

[2] Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed University), Pune, India., Pune, Maharashtra, India

[3] Electronics and Telecommunication, Symbiosis Institute of Technology, Pune, Pune, Maharashtra, India

[4] Department of Information Systems, Arizona State University, Phoenix, Arizona, United States

Corresponding Authors: Rahee Walambe, Ketan Kotecha
Email address: rahee.walambe@scaai.siu.edu.in, head@scaai.siu.edu.in

The evolution of electronic media is a mixed blessing. Due to the easy access, low cost, and faster reach of the information, people search out and devour news from online social networks. In contrast, the increasing acceptance of social media reporting leads to the spread of fake news. This is a minacious problem that causes disputes and endangers the societal stability and harmony. Fake news spread has gained attention from researchers due to its vicious nature. proliferation of misinformation in all media, from the internet to cable news, paid advertising and local news outlets, has made it essential for people to identify the misinformation and sort through the facts. Researchers are trying to analyze the credibility of information and curtail false information on such platforms. Credibility is the believability of the piece of information at hand. Analyzing the credibility of fake news is challenging due to the intent of its creation and the polychromatic nature of the news. In this work, we propose a model for detecting fake news. Our method investigates the content of the news at the early stage i.e. when the news is published but is yet to be disseminated through social media. Our work interprets the content with automatic feature extraction and the relevance of the text pieces. In summary, we introduce stance as one of the features along with the content of the article and employ the pre-trained contextualized word embeddings BERT to obtain the state-of-art results for fake news detection. The experiment conducted on the real-world dataset indicates that our model outperforms the previous work and enables fake news detection with an accuracy of 95.32%.

1 **Stance Detection with BERT Embeddings for**
2 **Credibility Analysis of Information on Social Media**

4 Hema Karande[1], Rahee Walambe[2,3]*, Victor Benjamin[4], Ketan Kotecha[1,2]* T. S. Raghu[4]

6 [1]Computer Science, Symbiosis Institute of Technology, Symbiosis International (Deemed
7 University), Pune, India 412115
8 [2]Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed
9 University), Pune, India 412115
10 [3]Electronics and Telecommunications, Symbiosis Institute of Technology, Symbiosis
11 International (Deemed University), Pune, India 412115
12 [4]Arizona State University, Tempe, Phoenix, USA
13 *corresponding author

15 Corresponding Author:
16 Ketan Kotecha[1,2]
17 [1]Computer Science, Symbiosis Institute of Technology, Symbiosis International (Deemed
18 University), Pune, India 412115
19 [3]Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed
20 University), Pune, India 412115
21 Email address: head@scaai.siu.edu.in

23 Rahee Walambe[2,3]
24 [2]Electronics and Telecommunications, Symbiosis Institute of Technology, Symbiosis
25 International (Deemed University), Pune, India 412115
26 [3]Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed
27 University), Pune, India 412115
28 Email address: rahee.walambe@scaai.siu.edu.in

Abstract

31 The evolution of electronic media is a mixed blessing. Due to the easy access, low cost, and faster
32 reach of the information, people search out and devour news from online social networks. In
33 contrast, the increasing acceptance of social media reporting leads to the spread of fake news. This
34 is a minacious problem that causes disputes and endangers societal stability and harmony. Fake
35 news spread has gained attention from researchers due to its vicious nature. proliferation of
36 misinformation in all media, from the internet to cable news, paid advertising and local news
37 outlets, has made it essential for people to identify the misinformation and sort through the facts.
38 Researchers are trying to analyze the credibility of information and curtail false information on
39 such platforms. Credibility is the believability of the piece of information at hand. Analyzing the
40 credibility of fake news is challenging due to the intent of its creation and the polychromatic nature
41 of the news. In this work, we propose a model for detecting fake news. Our method investigates
42 the content of the news at the early stage i.e. when the news is published but is yet to be
43 disseminated through social media. Our work interprets the content with automatic feature
44 extraction and the relevance of the text pieces. In summary, we introduce stance as one of the
45 features along with the content of the article and employ the pre-trained contextualized word

46 embeddings BERT to obtain the state-of-art results for fake news detection. The experiment
47 conducted on the real-world dataset indicates that our model outperforms the previous work and
48 enables fake news detection with an accuracy of 95.32%.
49
50 **Keywords: Fake News, Credibility, Misinformation, LSTM, BERT**
51

Introduction

53 In the information age, social networking sites have become a hotbed for spreading
54 misinformation. Misinformation (Soll, 2016) as a phenomenon is as old as true or factual ones.
55 The scale and scope of misinformation, however, have assumed alarming levels as social media
56 platforms and networks can spread misinformation rapidly. With the substantial increase in the
57 spread of misinformation, adverse impacts on individuals and society at large have also become
58 significant (Figueira & Oliveira, 2017). In this work, we propose a framework for identifying the
59 misinformation by employing state-of-the-art artificial intelligence algorithms. The first step in the
60 identification of misinformation is to understand what constitutes the misinformation. Fake news,
61 misinformation, disinformation all are various forms of non-factual information with variations in
62 the intent of the creator/spreader. Ethical Journalism Network[1] (EJN) defines fake news as
63 "information deliberately fabricated and published to deceive and mislead others into believing
64 falsehoods or doubting verifiable facts." Misinformation, disinformation, and mal-information are
65 specific subsets of information disorder. Disinformation is false and is deliberately designed to
66 harm an individual, organization, social group, or country. Mal-information is reality-based, used
67 to harm a person, social group, organization, or country[2] (UNESCO). Misinformation is not
68 created to cause harm and is false information that is disseminated by people who believe that it is
69 true. In comparison, disinformation is false information intentionally and knowingly propagated
70 by individuals to gain political, financial, social, or psychological benefits. Disinformation via
71 advertising can be used to discredit or falsely credit a product or a specific manufacturer for
72 financial gain[3] (Banner flow). In the political domain, disinformation could manifest from using
73 false information to discredit opposition parties or creating false narratives to aid one specific party
74 or candidate (Allcott & Gentzkow, 2017). Socially, one typical example could be the spread of
75 certain medical myths that are prevalent in specific communities and spreading them without much
76 thought (Waszak, Kasprzycka-Waszak, & Kubanek, 2018).
77 Misinformation or Information Disorder is usually in the form of false or out of context
78 information, photographs, or videos that are either intentionally created and distributed.
79 Sometimes, they are taken out of context to mislead, deceive, confuse or misinform the reader
80 (Pérez-Rosas, Kleinberg, Lefevre, & Mihalcea, 2017). Although there is news created for fun and
81 circulated as a part of a joke they have seriously impacted society. Researchers (Friggeri, Adamic,
82 Eckles, & Cheng, 2014) surveyed different aspects of false information and answered the question
83 "what can be termed as false?". The primary points considered are who is spreading the false
84 information, what are the reasons behind the reader's belief, and what is impact this false news
85 creates. The effects of dis /misinformation on society can prove detrimental. Misinformation has
86 caused a serious impact on various activities such as affecting the stock market (Bollen, Mao, &
87 Zeng, 2011), hampering the responses during natural disasters (Gupta, Lamba, Kumaraguru, &

---

[1] https://ethicaljournalismnetwork.org/tag/fake-news/page/2
[2] https://en.unesco.org/fightfakenews
[3] https://www.bannerflow.com/blog/marketing-fake-news-dangerous-game/

88   Joshi, 2013), instigating terrorism activity (Starbird, Maddock, Orand, Achterman, & Mason,
89   2014), kindling cyber-troop activity (Bradshaw & Howard, 2018), hampering the decision-making
90   ability during elections(News18) and creating panic, bringing about the economic crisis(herald)
91   and inciting religion-based attacks(Indianexpress) during Covid-19 pandemic.
92   Looking at the huge outburst of fake news around the coronavirus, the World Health
93   Organization(WHO)announced the new coronavirus pandemic was accompanied by a
94   'Misinformation Infodemic'. Various aspects of misinformation and its identification using AI
95   tools for COVID 19 data is reported in a recent article(Jyoti Choudrie, 2020). Fact Checkers report
96   fake news from general elections and attacks at Pulwama to the scrapping of Article 370 and the
97   ongoing protests against the Citizenship Amendment Act, which triggered a wide distribution of
98   misinformation across social media platforms(Economictimes). Misinformation affects
99   communities and their susceptibility in various ways; for instance, mob lynching and communal
100   poison.
101   The dependability of mass on social media news items has rapidly grown. It is reported
102   that out of the English-speaking news seekers in India 84 percent rely on Online news whereas 16
103   percent on the outpaced print media(Reuters). The urban, semi-urban teen-agers are the primary
104   consumers of social media news (Pérez-Rosas et al., 2017). Due to such tremendous use of online
105   platforms, the spread of information disorder is overwhelming and immense, causing harm to
106   society. In the year 2014, the World Economic Forum declared misinformation as one of the 10
107   global risks(W.E.Forum, 2014). Governments have taken some anti-misinformation efforts aiming
108   to curb the spread of unlawful content and misinformation spanning from the laws, Media literacy
109   campaigns, government task force, bills, platform agreements, to arrests(Poynter).
110   From the social media platforms available, Facebook and WhatsApp are particularly
111   widely used for news updates. As reported by Reuters, 75% use Facebook, 82% use WhatsApp,
112   26% use Instagram, 18% use Twitter. Hence it becomes the responsibility of these platforms to
113   help mitigate the spread of misinformation. Facebook[4] states that they are working in different
114   ways – e.g. most false news is motivated due to financial aspects, hence undermining the economic
115   incentives may prove useful. The International Fact-Checking Network and the fact-checkers are
116   working hard to investigate the facts behind a piece of information likely to be fake. Several
117   experiments were carried out to assess the effect of hoaxes, false reviews, and fake news. To create
118   a misinformation detection system, we need to consider the various aspects of the knowledge and
119   categorization of different features. Several researchers performed research and submitted it. We
120   present the literature in parts that concentrate on social and cognitive dimensions, categorization
121   strategies, and AI-based detection systems using different acceptable algorithms.
122   Mis- and disinformation can easily be disseminated – wittingly or unwittingly – through
123   all types of media. The ease of access to such quick information without any validation, has put a
124   responsibility on the reader to decide the correctness of the information at hand. Correctness,
125   Trustworthiness, or Credibility is the quality of the information to be trusted and believed in. In
126   the context of news, it encompasses the broader aspects of trustworthiness, reliability,
127   dependability, integrity, and reputation. When people are unable to debunk the information and
128   act accordingly, that makes poor decisions impacting their lives. It is essential to check the
129   credibility of the source, author, check your biases, check the date and supporting sources to
130   determine the reliability via comparison with reliable sources[5] (EMIC).

---

[4] https://www.facebook.com/facebookmedia/blog/working-to-stop-misinformation-and-false-news
[5] https://guides.emich.edu/newseval

131
132 While performing the credibility analysis we need to first examine how misinformation and
133 disinformation are being represented, spread understood, and acted upon. The role and motivation
134 of an individual behind resharing the original content is an important aspect while devising a policy
135 to curtail the spread and also for developing some technical solutions to mitigate it.
136 The most powerful of the information disorder content is that which harms people by influencing
137 their emotions. Since the social platforms are designed to express emotions through likes,
138 comments, and shares, all the efforts towards fact-checking and debunking false information are
139 ineffective since the emotional aspect of the sharing of information is impossible to control.
140 Detection and mitigation of information disorder are challenging due to the psychological aspects
141 of Motivation for Dissemination and the Proliferation of misinformation. The two primary
142 channels for spreading the misinformation are employed namely, Echo Chamber(Shu, Bernard, &
143 Liu, 2019) which is a situation where beliefs are reinforced or enhanced by contact and repetition
144 within a closed structure and Filter Bubble (Shu et al., 2019) is the state of intellectual isolation
145 that can result from custom searches when a website algorithm selectively estimates what
146 information a user wants to see based on user information, such as location, past click history, and
147 search. The concept of the filter bubble is used to target a specific group of people to spread the
148 specific misinformation. As per (Kumar & Geethakumari, 2014) cognitive psychology plays an
149 important role in the spread of misinformation.
150   As stated earlier, there are political, financial, and social aspects that play a role as a
151 motivation behind the creation of fake news items. These diverse angles, namely, the dynamic and
152 ubiquitous nature of information, difficulty in verification, and homophily prove to be some of the
153 primary challenges in finding the credibility of the information.
154
155 **Previous Work**
156 Misinformation detection is studied in different ways, starting with how it is created, spread, and
157 eventually affects the community. (Shu, Sliva, Wang, Tang, & Liu, 2017) surveys the literature
158 from two distinct phases: characterization and detection. Characterization is concerned with
159 understanding the basic concepts and principles of fake news in traditional and social media
160 whereas data mining with feature extraction and model construction is included in detection. (Shu
161 et al., 2017) in his paper presents the characteristics of Fake News on traditional and social media
162 that include Psychological and social foundations as well as fake accounts and echo chamber
163 creation on social media. The author also puts forward the detection approaches including News
164 Content and Social Context.
165 Various approaches towards fake social media news identification are proposed, including Data
166 Orientation, Feature Orientation, Model Orientation, and Application Orientation. Depending on
167 these approaches multiple systems have developed that concentrate on temporal features,
168 psychology, or the data for a data-oriented approach. Much explored approaches are Feature
169 Orientation that considers the content or the social context of the news. Depending on the dataset
170 the Model is selected either to be Supervised, Unsupervised, or Semi-supervised(Shu et al., 2017).
171 Feature Selection is an important step while approaching fake news detection. Features are broadly
172 categorized into content features and social context features by (Cao et al., 2018). The content
173 features include lexical, syntactic, and topic features whereas social context features include user
174 features, propagation features, and temporal features.
175 There is vast work done in detecting misinformation with various approaches, traditionally some
176 classification methods used were Decision Tree & Bayesian Networks (Castillo, Mendoza, &

177 Poblete, 2011), Random Forest & SVM (Kwon, Cha, Jung, Chen, & Wang, 2013), Logistic
178 Regression (Castillo et al., 2011) for the handcrafted features. The features like author, context,
179 and writing style (Potthast, Kiesel, Reinartz, Bevendorff, & Stein, 2017) of the news can help in
180 identifying the fake news, although writing style alone cannot be a good option. Linguistic signs
181 may be used to identify language characteristics such as n-grams, punctuation, psycholinguistic
182 characteristics, readability, etc. Classification based on the credibility of the person who liked it is
183 an approach taken in some cases (Shu et al., 2017). The conventional techniques of machine
184 learning have often resulted in a high-dimensional representation of linguistic information leading
185 to the curse of dimensionality where enormous sparse matrices need to be treated. This issue can
186 be solved with the use of word embeddings, which gives us low dimensional distributed
187 representations. Misinformation specifically a news item may constitute Words, Sentences,
188 Paragraphs, and Images. For applying any AI technique on text firstly we need to format the input
189 data into a proper representation that can be understood by the model we are designing. Different
190 state-of-art representation techniques like one-hot encoding, word embeddings like Continuous
191 Bag of Words and Skip-gram (Mikolov, Chen, Corrado, & Dean, 2013) that compute continuous
192 vector representations of very big datasets of terms, GloVe is Global word representation vectors
193 (Pennington, Socher, & Manning, 2014) global corpus statistics that train just on non-zero
194 elements in a word-word matrix, and not on the entire sparse matrix or single background windows
195 in a large corpus. BERT (Devlin, Chang, Lee, & Toutanova, 2018) bi-directional pre-training
196 setup, using the transformer encoder. Open-AI GPT pre-training model internally using the
197 transformer decoder concept. Pre-trained embeddings can be adapted to build a neural network-
198 based fake news detection model. Text data is a sequential time series data which has some
199 dependencies between the previous and later part of the sentence. Recurrent Neural Networks has
200 been widely used to solve NLP problems, traditionally encoder decoders architecture in Recurrent
201 Neural Network was a good option where an input sequence is fed into the encoder to get the
202 hidden representation which is further fed to the decoder and produce the output sequence. One
203 step towards fake news detection was to detect stance (Davis & Proctor, 2017) that involves
204 estimating the relative perspective (or stance) of two texts on a subject, argument, or problem.
205 This can help in identifying the authenticity of the news article based on whether the headline
206 agrees with, disagrees with, or is unrelated to the body of the article. Recurrent Neural Networks
207 (Shu et al., 2017), (Ma et al., 2016) to capture the variation of contextual information, CSI model
208 composed of three modules (Ruchansky, Seo, & Liu, 2017) implements Recurrent Neural Network
209 for capturing user activity's temporal pattern, learning the source characteristic based on user
210 behavior, and classifying the article. Researchers have also investigated the Rumor form of the
211 news that is circulated without confirmation or certainty to facts (DiFonzo & Bordia, 2007). A
212 rumor detection system (Cao et al., 2018) for Facebook (notify with a warning alert),
213 Twitter(credibility rating is provided and the user is allowed to give feedback on it) and
214 Weibo(users report fake tweets and elite users scrutinize and judge them) are in function.
215 	As the news articles usually contain a huge amount of text, this makes the input sequence
216 long enough. In such cases, the old information gets washed off and scattered focus over the
217 sequences which is due to a lack of explicit word alignment during decoding. Theirs raised a need
218 to solve these issues and the attention mechanism has done a good job. There are different flavors
219 of attention mechanisms that came up depending on their use cases, first and very basic version
220 i.e. the basic attention which extracts important elements from a sequence. Multi-dimensional
221 attention captures the various types of interactions between different terms. Hierarchical attention
222 extracts globally and locally important information. Self-attention (Vaswani et al., 2017) captures

223 the deep contextual information within a sequence. Memory-based attention discovers the latent
224 dependencies. Task-specific attention captures the important information specified by the task.
225   Singhania et.al. implemented a 3HAN hierarchical attention model (Singhania, Fernandez,
226 & Rao, 2017) that has 3 layers for words, sentences, and headlines each using bi-directional GRUs
227 of the network encoding and attention. Wang et.al. implemented Attention-based LSTM (Wang,
228 Huang, Zhu, & Zhao, 2016) for aspect-level sentiment analysis that finds the aspects and their
229 polarity for the sentence. Goldberg applied a novel design for the NLP task that incorporates an
230 attention-like mechanism in a Convolutional Network (Goldberg, 2016). Further enhancement
231 with a deep attention model with RNN given by (Chen, Li, Yin, & Zhang, 2018) learns selective
232 temporal hidden representations of the news item that bring together distinct features with a
233 specific emphasis at the same time and generate hidden representation.

234   Convolutional Neural Networks were used in computer vision tasks but recently they have
235 gained popularity in Natural Language processing tasks as well. CAMI (Yu, Liu, Wu, Wang, &
236 Tan, 2017) tries to early detect the misinformation. It is carried out by dividing the events into
237 phases and representing them using a paragraph vector (Le & Mikolov, 2014). Automatic,
238 identification of fake news based on geometric deep learning (Monti et al.,2019) generalizing
239 classical CNNs to graphs. FNDNet (Kaliyar, Goswami, Narang, & Sinha, 2020) deep
240 convolutional neural network. DMFN (Nguyen, Do, Calderbank, & Deligiannis, 2019) model
241 capturing dependencies among random variables using a Markov random field. Pattern driven
242 approach (Zhou & Zafarani, 2019) capturing the relation between news spreader and relation
243 between the spreaders of that news item. A mutual evaluation model (Ishida & Kuraya, 2018) that
244 dynamically builds a relational network model to identify credibility taking into consideration the
245 consistency of the content. Without a dynamic relation network, the content dependent model
246 would lead to a different score of the same article, since a different individual will have different
247 perspectives. Several researchers have proposed various approaches to the detection of fake news
248 as discussed in this section. Various classifiers and representation techniques are proposed. The
249 reported accuracy for these models ranges from 85 to 90%. However, there is a scope for
250 improving the accuracy of the fake news detection model.

251   From the above work, it can be observed that the a number of researchers have carried out
252 a detailed study in finding out various linguistic features and using different combinations of
253 features in classification tasks. Deep learning automatically learns features from data, instead of
254 adopting handcrafted features, it makes the task easier and can handle a huge amount of data. We
255 have used the deep learning approach for feature extraction and added a new feature that helps
256 improve the understanding of the text under consideration namely the stance which estimates the
257 relative perspective of two pieces of texts. The major contribution over the previous works lies in
258 the addition of stance as a feature along with the state of art BERT Model.
259 Recently BERT-based models are applied in NLP tasks, that is the hybrid of BERT and Artificial
260 Intelligence techniques like RNN (Kula, Choraś, & Kozik, 2020),CNN (Kaliyar, Goswami, &
261 Narang, 2021) and both (Ding, Hu, & Chang, 2020). BERT models have also proved its importance
262 to deal with multi-modal news articles (Zhang et al., 2020).
263
264   We have considered cosine distance between the vectors of the news title and the news
265 body as the similarity measure. In traditional literature, the stance is defined in several ways. E.g.
266 stance is detected towards a particular topic (Sun, Wang, Zhu, & Zhou, 2018), agreement or

267   disagreement towards a particular claim (Mohtarami et al., 2018) and even attitude expressed in a
268   text towards a target (Augenstein, Rocktäschel, Vlachos, & Bontcheva, 2016). All of these have
269   predefined categories like a negative, positive, neutral, agree, disagree, unrelated, for or against.
270   Our intention here is to find the relation/similarity between the two text pieces(namely the title and
271   the body of the text). Hence, we do not find the score towards a particular topic but the measure
272   of similarity between the title and the body of the news. The reason we have made such a choice
273   is that for the unseen data we are not already aware of the topic it is focusing on. And will make
274   our system more generalized. Such an approach can identify how close or farther the title is to the
275   text in the body of an article under consideration

276   Due to this additional feature, our training data is better equipped for more accurate
277   predictions. Also, with the help of state-of-art language model BERT, our model captures the
278   semantics of the text efficiently with a multi-layer bidirectional transformer encoder which helps
279   to learn deep bi-directional representations of text(article) and finetuning it on our training data to
280   classify an article into fake or real, using the probability score our model assigns to it. In summary,
281   the main contributions of this article are:

- Introducing stance as one of the features along with the content of the article to obtain state-of-the-art performance when detecting fake news using an AI model
- Develop a model that captures the semantics of information using the pre-trained contextualized word embeddings BERT(Language Modelling Technique)
- Experimentation and validation of the above approach on the benchmark dataset.

287   The remaining article is structured as follows: Section 2 outlines the methodology we have adopted
288   to design our system, Section 3 describes the experimental setup and parameters we are using,
289   Section 4 describes our findings and discussion on them, Section 5 concludes our work.

Materials & Methods

291   We are looking in this paper to build a classifier that detects fake news with better accuracy than
292   already reported. We have experimented with multiple AI models and evaluated accuracies. Our
293   model for fake news detection is based on the content features which use pre-trained embedding
294   to better capture the contextual characteristics. The complete process is explained in the below
295   subsections.

Pre-processing

297   It is observed that text data is not clean. The data is extracted from different sources hence they
298   have different characteristics, to bring such data on a common ground text pre-processing is an
299   important step. Our pre-processing pipeline depends on the embeddings we will be using for our
300   task. Since most of the embeddings do not provide vector values for the special characters and
301   punctuations we need to clean the data by removing such words. Our preprocessing step involves
302   the removal of punctuations, special characters, extra spaces, and lowering the case of the text.

Architecture

304   Fig. 1 depicts the complete pipeline of our model.

Dataset

306   Manual fact-checking is a tedious and lengthy task and such fact-checked articles are very less to
307   train deep learning models. Researchers have extracted news articles from the websites they
308   believe to be authentic as real or genuine and similarly the fake articles. McIntire Fake and Real

309 News Dataset[6], a quite large dataset and a balanced dataset that contains both fake stories and true
310 news. McIntire used the part from Kaggle's fake news collection for the Fake news class, while
311 he got articles from credible journalism organizations like the New York Times, Wall Street
312 Journal, Bloomberg, National Public Radio, and The Guardian for the real news class. The dataset
313 contains 6335 news articles out of the 3171 that are genuine and 3164 that are fake. The dataset is
314 equally distributed with ~50-50% of real labeled articles and fake labeled articles (Refer Fig. 2).
315 In real-world applications the data may be skewed, however, for experimentation and dataset
316 building, we have considered a balanced dataset which leads to a more robust and generalized
317 model.
318
319 **Data filtering and feature Selection**
320
321 The data filtering/cleaning is the task of preparing the data for encoding and training purposes.
322 Here firstly we remove the spaces and punctuations from the text as they do not play an important
323 role in the understanding of the context. This module of our pipeline ends up with feature selection
324 for our training. We perform stance detection on the title and text to get the similarity score which
325 gives us an additional feature to make our classifier perform better. We then encode our selected
326 features
327
328 **Stance Calculation**
329 We calculate the stance estimating the relative perspective of two pieces of the text concerning a
330 subject, argument, or issue. It is the probability score that will be used as an additional feature to
331 detect the misinformation. From Sensationalism detection by (Dong, Yao, Wang, Benatallah, &
332 Huang, 2019) we consider that the similarity between the article body and article headline can be
333 correlated with articles' credibility. The similarity is captured by first embedding the article body
334 and article headline onto the same space and computing the cosine distance as the similarity
335 measure. To achieve this we first tokenize the article's body and headline, then calculate the cosine
336 distance between each article's body and headline pair. Cosine distance is a good similarity
337 measure as it determines the document similarity irrespective of the size; this is because the cosine
338 similarity captures the angle of the documents and not the magnitude. Mathematically, it measures
339 the cosine of the angle between two vectors projected in a multi-dimensional space.

$$Cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \, \|\vec{b}\|} = \frac{\sum_{1}^{n} a_i \, b_i}{\sqrt{\sum_{1}^{n} a_i^2} \cdot \sqrt{\sum_{1}^{n} b_i^2}}$$

340

341
342 Here, $\vec{a} \cdot \vec{b} = \sum_{1}^{n} a_i \, b_i = a_1 b_1 + a_2 b_2 + ... + a_n b_n$ is the dot product of the two vectors.
343
344 **Feature Encoding**
345
346 Any AI model requires the input data to be in number format. Different state-of-art representation
347 techniques like one-hot encoding, word embeddings like Continuous Bag of Words and Skip-

---

[6] https://github.com/GeorgeMcIntire/fake_real_news_dataset

348   gram, GloVe (Global vectors for word representation), Open AI GPT pre-training model internally
349   using the transformer decoder concept, BERT (Bidirectional Encoder Representations from
350   Transformers) bi-directional pre-training model that uses transformer encoder within are available
351   to perform such encoding tasks. Our model uses language model BERT's smaller version namely
352   Distil-BERT (Sanh, Debut, Chaumond, & Wolf, 2019) which has about half the total number of
353   parameters of the BERT base. In Distil-BERT, a triple loss combining language modeling,
354   distillation, and loss of cosine distance was included to exploit the inductive biases gained during
355   training by these types of models. We are further fine-tuning the model for our text classification
356   task on our dataset.
357
358   **Classification Models**
359   We have evaluated three AI models namely simple Artificial Neural Network, Recurrent Neural
360   Network, Long Short-Term Memory (LSTM), Bidirectional LSTM, and Convolutional Neural
361   Network.
362   1. Long Short-Term Memory (LSTM): LSTMs is chosen over RNN as it succeeds in overcoming
363   the issue of long-term dependencies by keeping a memory at all times. The output of the
364   embedding layer is fed into the LSTM layer that consists of 64 units and further passed through
365   the dense layers with sigmoid activation and binary cross-entropy for computation. The next model
366   implemented is the bidirectional LSTM model, which is an extension of traditional LSTM that
367   trains two LSTMs in opposite directions to enhance model efficiency in capturing all the necessary
368   information.
369   2. Bidirectional LSTM: Similar to the LSTM model it takes input from the embedding layers, our
370   Bi-LSTM layer contains 64 units that are fed to the dense layers for computation. For comparison
371   purposes, the activation and loss values used are the same as the previous model.
372   3. Convolutional Neural Network: Convolutional neural networks although designed for computer
373   vision tasks recently it has given excellent results with NLP tasks as well. In the case of computer
374   vision tasks, the filters slide over the patches of an image, whereas in NLP tasks the filter slides
375   few words at a time over the sentence matrices. This makes Convolutional Neural Networks work
376   well for classification tasks. So we have also implemented a CNN model that consists of a single
377   Conv-1D with the kernel of size 5 and the Max Pooling layer. Further, a flattened layer and a fully
378   connected layer with 64 nodes are used for computation. In all of the above models, we have used
379   a Dense layer that operates linearly where every input is related by some weight to each output.
380   The Loss function used is the cross-entropy loss that measures classification model performance
381   whose output is a probability value between 0 and 1.Loss of cross-entropy increases as the
382   predicted likelihood diverges from the actual mark. In the case of binary classification, where the
383   number of groups equals 2, cross-entropy can be determined as:
384    $$L = -(ylog(p) + (1-y)loglog(1-p))$$
385   These models classify the news articles as fake or real using the sigmoid activation function.
386   **Experimental Setup**
387
388   The framework is developed in the Python backend and uses the python libraries namely Keras,
389   NLTK, Pandas, Numpy, Sklearn, sci-kit, etc. The dataset was divided into training, validation and
390   testing sets with train_test_split functionality of sci-kit learn. The scale of the training set size was
391   70 percent, the validation set scale of 15 percent, and 15 percent of the test set size. Data pre-
392   processing involved the use of the NLTK tool for the removal of HTML tags, punctuations,
393   multiple spaces, etc. The distil-Bert-base-uncased model from the hugging face is used to obtain

394   the embeddings that are later fine-tuned on our dataset. We encode our data with the max length
395   of 512, dimension to be 768, dropout of 0.1, and the number of layers as 6 to gain the input ids and
396   the mask ids. For the classification purpose, we have used LSTM, Bidirectional, Conv1D, Dense
397   layers from Keras. The number of units chosen for each of the layers was based on the
398   experimentations carried out. For the GloVe model, different vector dimensions were tried 100
399   and 300 and the vector dimension 100 gave good accuracy results. The loss function and activation
400   function used were cross-entropy loss and sigmoid activation for all the models.

401   **Performance Metrics**

402

403   We have used the Confusion matrix, Accuracy, Precision, Recall, F1, and ROC to evaluate our
404   model's efficiency (Hossin & Sulaiman, 2015).

405   1. Confusion Matrix: The information performed by a classifier regarding actual and
406   predicted classifications is studied by a confusion matrix.
407   2. Accuracy: Accuracy is the proportion of true outcomes within the total number of
408   examined cases.
409   3. precision: Precision tells about what proportion of predicted Positives is truly Positive.
410   4. recall: It tells us what proportion of real positives is graded correctly.
411   5. F1 Score: It gives the harmonic mean of precision and recall.
412   6. ROC: ROC demonstrates how well the percentage of the positives are isolated from the
413   negative groups.
414   These metrics helped us analyze the results we gained from our model. Table 1. Depicts the values
415   for accuracies during training, validation, and testing, along with the precision, recall, F1, and the
416   ROC.

Results and Discussion

418   In this work, we have proposed an approach to upgrade the fake news identification system with
419   the inclusion of an additional feature termed "stance". Stance helps us in understanding the
420   relevance of the article title (i.e. the headline of the news) to the article body(i.e. the text of the
421   news). We add this feature to our content features that are obtained from the pre-trained BERT
422   model, which provides additional insight into the article. The AI models we have experimented
423   with are ANN, LSTM, Bidirectional LSTM, and CNN. Results are obtained by training and testing
424   these models with different vector representation techniques and even including the attention layer
425   to some models. The results are presented in Table 1. Table 2. shows the classification results for
426   our proposed model.
427   The best results we obtained are with the usage of a pre-trained language model. Table 1 shows
428   the accuracies with different settings as with GloVe Embeddings we attain an accuracy of 92.1%
429   for the LSTM model. With the inclusion of the Attention layer to these models the accuracy shows
430   some improvement of about 1%. Our proposed model adapting the BERT embedding to capture
431   the contextual information from the articles have proved to perform better with an accuracy of
432   95.32%. We have experimented and obtained the results using the pre-trained BERT model and
433   fine-tuned BERT model. The results we obtained via both these models demonstrate a negligible
434   difference. One of the possible explanations for this could be that the BERT is trained on the
435   English language corpus. The dataset we have used for experimentation is also in English and has
436   a similar structure and features.
437   To show the effect of stance as a feature we have experimented with BERT encoding, which builds
438   a representation of the news title and news body along with the similarity between them, and

439   demonstrated that it outperforms encoding the news title and news body alone. From the results,
440   it can be observed that there is a considerable increase in the testing accuracy (Refer Table 3.).
441   We have dealt only with the content part of the article, the reason being when a news article is
442   published and not much-circulated yet, the metadata such as reposts, likes, shares, etc. are not
443   available. Then content can be the only parameter considered for fake news detection. The below
444   plots give a clear view of the results obtained from our model (Refer Fig.3). We have carried out
445   a 5-fold Cross-validation resampling procedure to evaluate our model and make the results
446   comparable with the other models on the same dataset (Refer Table. 4). We implemented a
447   stratified k-fold cross validation, however observe a few misclassified samples in the test results.
448   This is primarily due to the overlapping of features in the two classes and having unclear distinction
449   due to that.
450   To validate our approach, we carried out a comparison with pre-existing approaches by
451   considering a benchmark dataset by G. McIntire Fake and Real News Dataset. Many have
452   proposed a solution to the fake news detection problem using hand-crafted feature engineering and
453   applying Machine Learning models like Naïve Bayes, Random Forest, Boosting, Support Vector
454   Machine, Decision Tree, Logistic Regression. (Reddy, Raj, Gala, & Basava, 2020) reported an
455   accuracy of 95% with a gradient boosting algorithm on the combination of stylometric and CBOW
456   (Word2Vec) features. (Bali, Fernandes, Choubey, & Goel, 2019) used Sentiment polarity,
457   Readability, Count of words, Word Embedding, and Cosine similarity as the features to
458   discriminate fake news with machine learning models. He reported the highest accuracy of 87.3%
459   with the XGBoost model. (Esmaeilzadeh, Peh, & Xu, 2019) uses the LSTM-encoder-decoder with
460   attention to the text, headline, and self-generated summary of the text to obtain the features. He
461   proved that the summary of the text gives better accuracy of 93.1%. Embedding methods such as
462   LSTM, depth LSTM, LIWC CNN, and N-gram CNN were used (Huang & Chen, 2020) and
463   weights were optimized using the Self-Adaptive Harmony Search (SAHS) algorithm with an initial
464   accuracy of 87.7% that was increased to 91%. (Khan, Khondaker, Islam, Iqbal, & Afroz, 2019) in
465   his research he used word count, average word length, article length, count of numbers, count of
466   parts of speech(adjective), count of exclamation along with the sentiment feature, and n-gram
467   GloVe embedding encoded features. His work reported accuracy of 90% with the Naïve Bayes
468   classifier and 95% with the LSTM deep neural network. (Bhutani, Rastogi, Sehgal, & Purwar,
469   2019) reported an accuracy of 84.3% with the Naïve Bayes model and TF-IDF scores, sentiments,
470   and Cosine similarity scores as features. (Gravanis, Vakali, Diamantaras, & Karadais, 2019)
471   developed a content-based model and reported an accuracy of 89% with the Support Vector
472   Machine. (George, Skariah, & Xavier, 2020) in his work used CNN for Linguistic features and
473   Multi Headed self-Attention for the contextual features. Accuracies of 84.3% and 83.9% were
474   achieved with machine learning models such as the Naïve Bayes classifier and Random forest
475   respectively. With deep learning models like LSTM and FastText embedding accuracy reported
476   was 94.3%. With the increasing creation and availability of misinformation, automatic feature
477   extraction with the help of Deep Learning algorithms has also been experimented. A comparison
478   of other work with our model is shown in Table 5.
479
480   ## Conclusion
481
482   Information disorder and the spread of fake news is a highly prevalent and challenging problem in
483   this digital era. The highly harmful potential of fake news on social stability and the psychological
484   health of the masses is undisputed and proven time and again. To identify and detect the

485 misinformation is the primary step in controlling its spread and combat its harmful effects on
486 society. To this end, in this paper, we have proposed a methodology to identify the fake news
487 based on stance detection in addition to other standard features. The content of an article is the
488 basic informative piece that does play a significant role in assigning credibility to the article. We
489 propose a model that uses language features based on the article's content to discriminate against
490 fake news from real news. Our model tries to detect the fake articles at a stage when they are yet
491 to propagate in the social network. To make the detection more precise we have added a stance
492 feature to the dataset. This stance value has helped us understand the content by finding the
493 relevance between the article headline and article body. Along with this additional feature, we
494 have learned the language representation with the help of the BERT technique. The transfer
495 learning concept that injects external knowledge gives a better understanding of our content. It is
496 observed that our system has shown improvement in the results as compared to the other models.
497 Previous work is conducted with handcrafted linguistic features, stylometric features, TF-IDF
498 scores, n-gram features as well as automatic feature extraction from the content of the articles.
499 Evaluation of our model is done on Kaggle's open source dataset for news articles. The results
500 presented demonstrate the performance of various deep learning architectures for misinformation
501 detection, wherein the BERT embeddings based CNN architecture provides the highest
502 performance of 95% accuracy and comparable precision and recall. Besides, we have also
503 compared our approach with other pre-existing approaches and shown that our proposed model
504 provides superior performance as compared to the other methods. One of the limitations is that it
505 is essential to have required features in the dataset. Without this, this approach will not work
506 effectively. We will extend this work in the future to overcome this limitation.
507 On the social media platform Images, Videos and Audio have become a part of a news
508 article to make it more informative and sensational at times. The fake news articles are usually
509 intertwined with the factual information that makes interpretation of the contextual information
510 difficult. As an extension of this work Image, Video, and Audio can also be used in the content-
511 based detection systems to make them near to perfect.

Acknowledgment:

513 This work has been supported by the Scheme for Promotion of Academic and Research
514 Collaboration (SPARC) 2018-19, MHRD (project no. P571).

References

516
517 Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of*
518 *economic perspectives, 31*(2), 211-236.
519 Augenstein, I., Rocktäschel, T., Vlachos, A., & Bontcheva, K. (2016). Stance detection with
520 bidirectional conditional encoding. *arXiv preprint arXiv:1606.05464*.
521 Bali, A. P. S., Fernandes, M., Choubey, S., & Goel, M. (2019). *Comparative performance of machine*
522 *learning algorithms for fake news detection.* Paper presented at the International
523 Conference on Advances in Computing and Data Sciences.
524 Bhutani, B., Rastogi, N., Sehgal, P., & Purwar, A. (2019). *Fake news detection using sentiment*
525 *analysis.* Paper presented at the 2019 Twelfth International Conference on Contemporary
526 Computing (IC3).
527 Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of*
528 *computational science, 2*(1), 1-8.

529   Bradshaw, S., & Howard, P. N. (2018). Challenging truth and trust: A global inventory of organized
530   social media manipulation. *The Computational Propaganda Project, 1*.

531   Cao, J., Guo, J., Li, X., Jin, Z., Guo, H., & Li, J. (2018). Automatic rumor detection on microblogs: A
532   survey. *arXiv preprint arXiv:1807.03505*.

533   Castillo, C., Mendoza, M., & Poblete, B. (2011). *Information credibility on twitter.* Paper presented
534   at the Proceedings of the 20th international conference on World wide web.

535   Chen, T., Li, X., Yin, H., & Zhang, J. (2018). *Call attention to rumors: Deep attention based recurrent*
536   *neural networks for early rumor detection.* Paper presented at the Pacific-Asia conference
537   on knowledge discovery and data mining.

538   Davis, R., & Proctor, C. (2017). Fake news, real consequences: Recruiting neural networks for the
539   fight against fake news. In: Stanford CS224d Deep Learning for NLP final project.

540   Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional
541   transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

542   DiFonzo, N., & Bordia, P. (2007). *Rumor psychology: Social and organizational approaches*:
543   American Psychological Association.

544   Dong, M., Yao, L., Wang, X., Benatallah, B., & Huang, C. (2019). *Similarity-aware deep attentive*
545   *model for clickbait detection.* Paper presented at the Pacific-Asia Conference on
546   Knowledge Discovery and Data Mining.

547   Economictimes. Retrieved from https://economictimes.indiatimes.com/news/politics-and-
548   nation/fake-news-still-a-menace-despite-government-crackdown-fact-
549   checkers/articleshow/72895472.cms?from=mdr

550   Esmaeilzadeh, S., Peh, G. X., & Xu, A. (2019). Neural abstractive text summarization and fake news
551   detection. *arXiv preprint arXiv:1904.00788*.

552   Figueira, Á., & Oliveira, L. (2017). The current state of fake news: challenges and opportunities.
553   *Procedia Computer Science, 121*, 817-825.

554   Friggeri, A., Adamic, L., Eckles, D., & Cheng, J. (2014). *Rumor cascades.* Paper presented at the
555   Eighth International AAAI Conference on Weblogs and Social Media.

556   George, J., Skariah, S. M., & Xavier, T. A. (2020). *Role of Contextual Features in Fake News*
557   *Detection: A Review.* Paper presented at the 2020 International Conference on Innovative
558   Trends in Information Technology (ICITIIT).

559   Goldberg, Y. (2016). A primer on neural network models for natural language processing. *Journal*
560   *of Artificial Intelligence Research, 57*, 345-420.

561   Gravanis, G., Vakali, A., Diamantaras, K., & Karadais, P. (2019). Behind the cues: A benchmarking
562   study for fake news detection. *Expert Systems with Applications, 128*, 201-213.

563   Gupta, A., Lamba, H., Kumaraguru, P., & Joshi, A. (2013). *Faking sandy: characterizing and*
564   *identifying fake images on twitter during hurricane sandy.* Paper presented at the
565   Proceedings of the 22nd international conference on World Wide Web.

566   herald, D. Retrieved from https://www.deccanherald.com/national/covid-19-infodemic-can-
567   impact-economy-warn-experts-830024.html

568   Hossin, M., & Sulaiman, M. (2015). A review on evaluation metrics for data classification
569   evaluations. *International Journal of Data Mining & Knowledge Management Process,*
570   *5*(2), 1.

Huang, Y.-F., & Chen, P.-H. (2020). Fake News Detection Using an Ensemble Learning Model Based on Self-adaptive Harmony Search Algorithms. *Expert Systems with Applications*, 113584.

Indianexpress. Retrieved from https://indianexpress.com/article/india/fake-news-in-the-time-of-coronavirus-boom-analyses-178-fact-checks-on-covid-19-related-misinformation-since-january-6401441/

Ishida, Y., & Kuraya, S. (2018). Fake news and its credibility evaluation by dynamic relational networks: A bottom up approach. *Procedia Computer Science, 126*, 2228-2237.

Jyoti Choudrie, K. K. a. R. W. (2020). Developing an identification and understanding of COVID-19 information: an AI based study. Retrieved from https://www.bcs.org/content-hub/developing-an-identification-and-understanding-of-covid-19-information-an-ai-based-study/

Kaliyar, R. K., Goswami, A., Narang, P., & Sinha, S. (2020). FNDNet–A deep convolutional neural network for fake news detection. *Cognitive Systems Research, 61*, 32-44.

Khan, J. Y., Khondaker, M., Islam, T., Iqbal, A., & Afroz, S. (2019). A benchmark study on machine learning methods for fake news detection. *arXiv preprint arXiv:1905.04749*.

Kumar, K. K., & Geethakumari, G. (2014). Detecting misinformation in online social networks using cognitive psychology. *Human-centric Computing and Information Sciences, 4*(1), 1-22.

Le, Q., & Mikolov, T. (2014). *Distributed representations of sentences and documents.* Paper presented at the International conference on machine learning.

Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B. J., Wong, K.-F., & Cha, M. (2016). Detecting rumors from microblogs with recurrent neural networks.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mohtarami, M., Baly, R., Glass, J., Nakov, P., Màrquez, L., & Moschitti, A. (2018). Automatic stance detection using end-to-end memory networks. *arXiv preprint arXiv:1804.07581*.

News18. Retrieved from https://www.news18.com/news/politics/over-150-cases-of-fake-news-reported-during-lok-sabha-elections-says-union-minister-2244463.html

Nguyen, D. M., Do, T. H., Calderbank, R., & Deligiannis, N. (2019). *Fake news detection using deep markov random fields.* Paper presented at the Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers).

Pennington, J., Socher, R., & Manning, C. D. (2014). *Glove: Global vectors for word representation.* Paper presented at the Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP).

Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2017). Automatic detection of fake news. *arXiv preprint arXiv:1708.07104*.

Potthast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2017). A stylometric inquiry into hyperpartisan and fake news. *arXiv preprint arXiv:1702.05638*.

Poynter. Retrieved from https://www.poynter.org/ifcn/anti-misinformation-actions/

Reddy, H., Raj, N., Gala, M., & Basava, A. (2020). Text-mining-based Fake News Detection Using Ensemble Methods. *International Journal of Automation and Computing*, 1-12.

614   Reuters. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-
615   03/India_DNR_FINAL.pdf

616   Ruchansky, N., Seo, S., & Liu, Y. (2017). *Csi: A hybrid deep model for fake news detection.* Paper
617   presented at the Proceedings of the 2017 ACM on Conference on Information and
618   Knowledge Management.

619   Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT:
620   smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

621   Shu, K., Bernard, H. R., & Liu, H. (2019). Studying fake news via network analysis: detection and
622   mitigation. In *Emerging Research Challenges and Opportunities in Computational Social
623   Network Analysis and Mining* (pp. 43-65): Springer.

624   Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data
625   mining perspective. *ACM SIGKDD explorations newsletter, 19*(1), 22-36.

626   Singhania, S., Fernandez, N., & Rao, S. (2017). *3han: A deep neural network for fake news
627   detection.* Paper presented at the International Conference on Neural Information
628   Processing.

629   Soll, J. (2016). The Long and Brutal History of Fake News.

630   Starbird, K., Maddock, J., Orand, M., Achterman, P., & Mason, R. M. (2014). Rumors, false flags,
631   and digital vigilantes: Misinformation on twitter after the 2013 boston marathon
632   bombing. *IConference 2014 Proceedings*.

633   Sun, Q., Wang, Z., Zhu, Q., & Zhou, G. (2018). *Stance detection with hierarchical attention
634   network.* Paper presented at the Proceedings of the 27th international conference on
635   computational linguistics.

636   Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017).
637   *Attention is all you need.* Paper presented at the Advances in neural information
638   processing systems.

639   W.E.Forum. (2014). Retrieved from http://reports.weforum.org/outlook-14/top-ten-trends-
640   category-page/10-the-rapid-spreadof-
641   misinformation-online/
642

643   Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2016). *Attention-based LSTM for aspect-level sentiment
644   classification.* Paper presented at the Proceedings of the 2016 conference on empirical
645   methods in natural language processing.

646   Waszak, P. M., Kasprzycka-Waszak, W., & Kubanek, A. (2018). The spread of medical fake news
647   in social media–the pilot quantitative study. *Health policy and technology, 7*(2), 115-118.

648   WHO. Retrieved from https://www.who.int/docs/default-source/coronaviruse/situation-
649   reports/20200202-sitrep-13-ncov-v3.pdf

650   Yu, F., Liu, Q., Wu, S., Wang, L., & Tan, T. (2017). A convolutional approach for misinformation
651   identification.

652   Zhou, X., & Zafarani, R. (2019). Network-based fake news detection: A pattern-driven approach.
653   *ACM SIGKDD explorations newsletter, 21*(2), 48-60.
654

655   Ding, J., Hu, Y., & Chang, H. (2020). *BERT-Based Mental Model, a Better Fake News Detector.*
656   Paper presented at the Proceedings of the 2020 6th International Conference on Computing
657   and Artificial Intelligence.

658 Kaliyar, R. K., Goswami, A., & Narang, P. (2021). FakeBERT: Fake news detection in social
659   media with a BERT-based deep learning approach. *Multimedia Tools and Applications*, 1-
660   24.
661 Kula, S., Choraś, M., & Kozik, R. (2020). *Application of the BERT-Based Architecture in Fake
662   News Detection.* Paper presented at the Conference on Complex, Intelligent, and Software
663   Intensive Systems.
664 Zhang, T., Wang, D., Chen, H., Zeng, Z., Guo, W., Miao, C., & Cui, L. (2020). *BDANN: BERT-
665   Based Domain Adaptation Neural Network for Multi-Modal Fake News Detection.* Paper
666   presented at the 2020 International Joint Conference on Neural Networks (IJCNN).
667

Figure 1

Pipeline architecture of the system

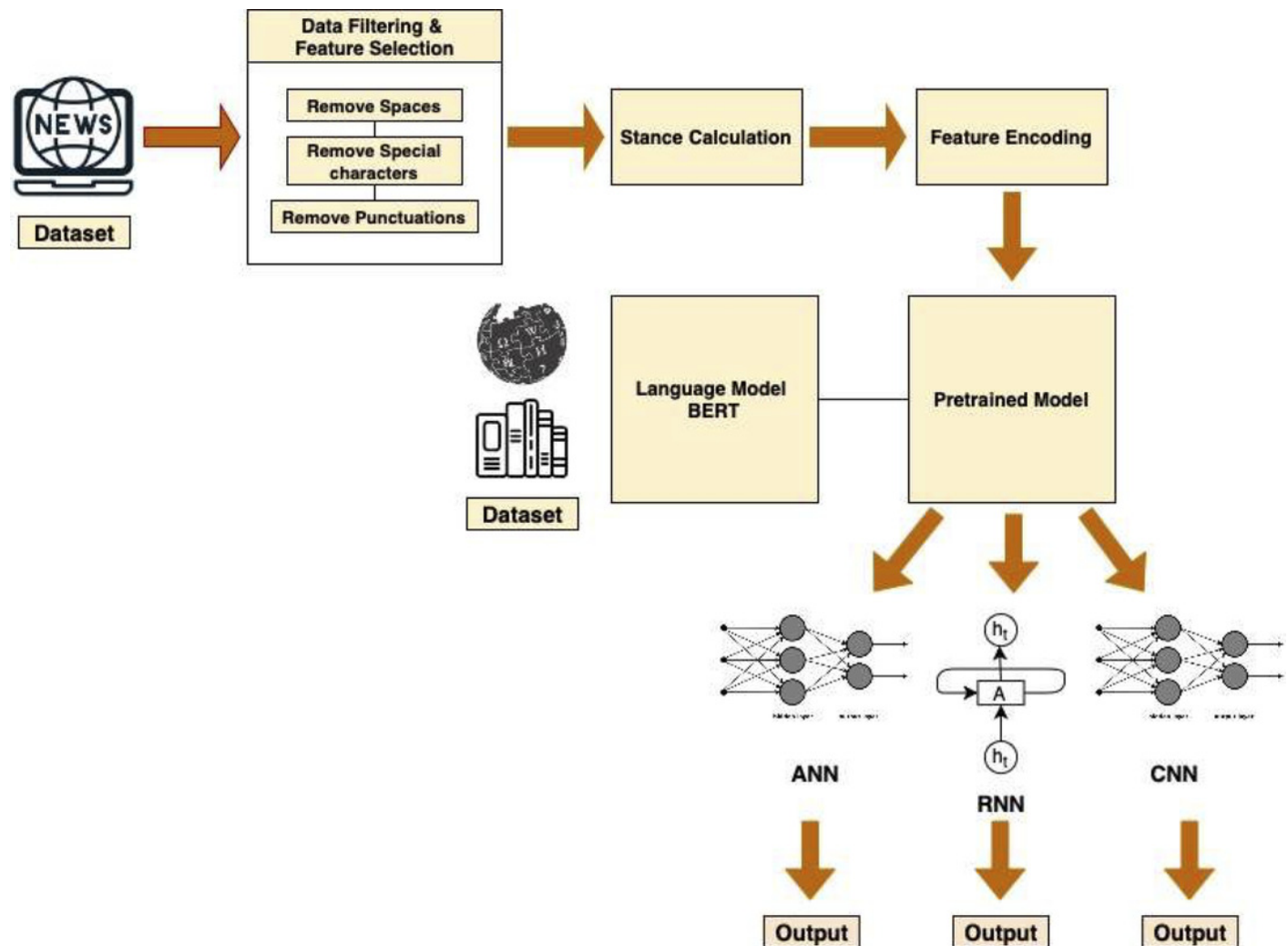



Figure 2

A part of a dataset consisting of true and false news instances

	title	text	label
0	You Can Smell Hillary's Fear	Daniel Greenfield, a Shillman Journalism Fello...	FAKE
1	Watch The Exact Moment Paul Ryan Committed Pol...	Google Pinterest Digg Linkedin Reddit Stumbleu...	FAKE
2	Kerry to go to Paris in gesture of sympathy	U.S. Secretary of State John F. Kerry said Mon...	REAL
3	Bernie supporters on Twitter erupt in anger ag...	— Kaydee King (@KaydeeKing) November 9, 2016 T...	FAKE
4	The Battle of New York: Why This Primary Matters	It's primary day in New York and front-runners...	REAL
5	Tehran, USA	\nI'm not an immigrant, but my grandparents ...	FAKE
6	Girl Horrified At What She Watches Boyfriend D...	Share This Baylee Luciani (left), Screenshot o...	FAKE
7	'Britain's Schindler' Dies at 106	A Czech stockbroker who saved more than 650 Je...	REAL
8	Fact check: Trump and Clinton at the 'commande...	Hillary Clinton and Donald Trump made some ina...	REAL
9	Iran reportedly makes new push for uranium con...	Iranian negotiators reportedly have made a las...	REAL

Figure 3

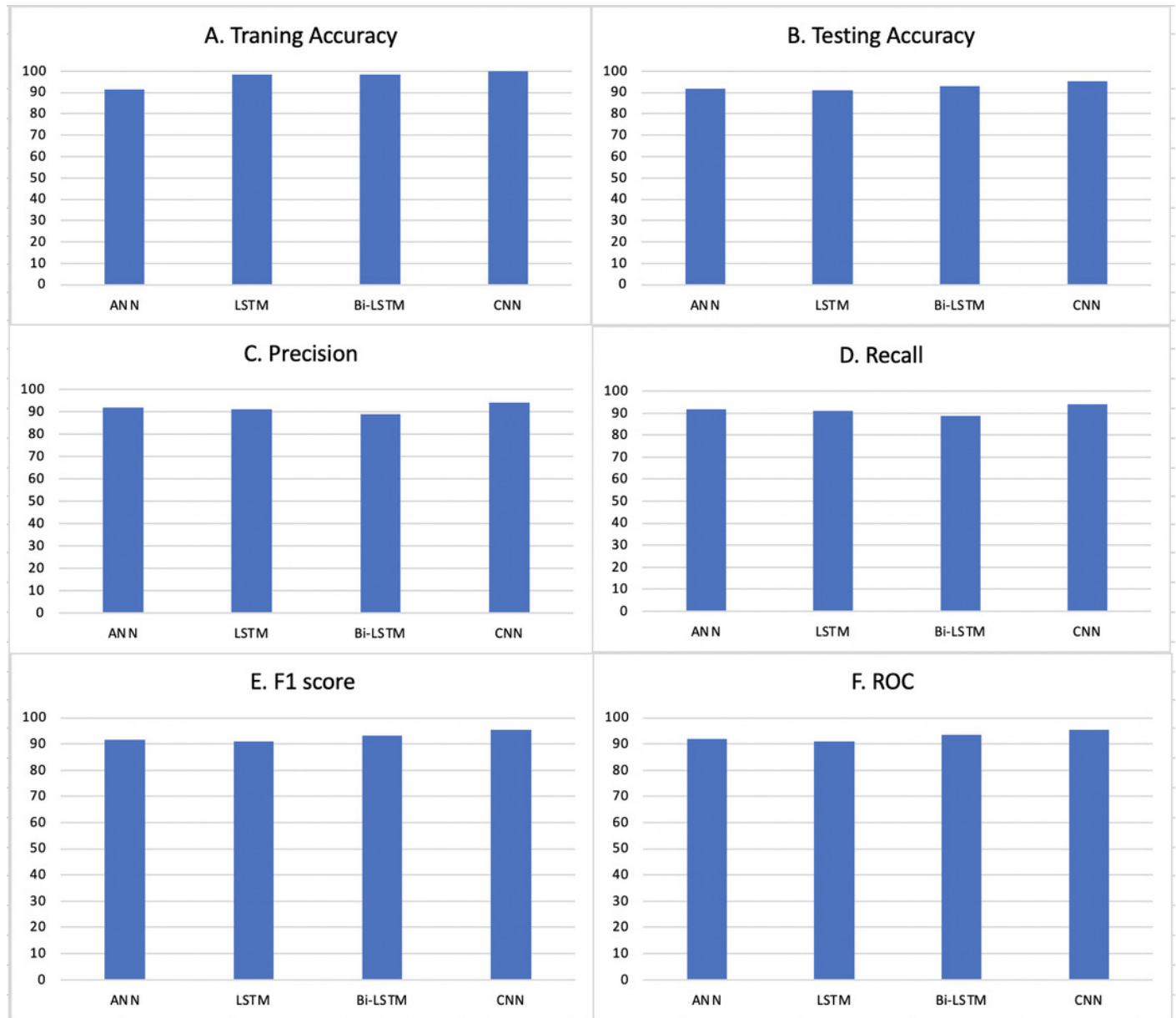
Figure 3(a) Training accuracies for different models

Performance of different AI models

Table 1: Performance of different AI models

	Models	Accuracy(%)	Precision (%)	Recall (%)	F1 (%)
Tokenizer	LSTM	86.6	85.1	88.7	86.9
	Bi-LSTM	85.4	84.9	86.2	85.5
	CNN	**93.0**	93.0	93.0	93.01
GloVe embeddings	LSTM	**92.1**	91.7	92.7	92.2
	Bi-LSTM	91.9	90.2	93.9	92.1
	CNN	91	91.6	90.2	90.9
GloVe embeddings and attention mechanism	LSTM	**92.1**	91.7	92.7	92.2
	Bi-LSTM	91.9	90.2	93.9	92.1
	CNN	91	91.6	90.2	90.9
BERT embeddings	LSTM	91.16	91.01	91.01	91.01
	Bi-LSTM	93.05	88.76	88.76	93.3
	CNN	**95.32**	94.11	94.11	95.31

Classification results for the proposed model

Table 2: Classification results for the proposed model

Models	Training Accuracy(%)	Validation Accuracy(%)	Testing Accuracy(%)	Precision (%)	Recall (%)	F1 (%)	ROC (%)
ANN	91.52	91.86	91.85	91.98	91.98	91.69	91.85
LSTM	98.51	91.16	91.16	91.01	91.01	91.01	91.15
Bi-LSTM	98.48	93.06	93.05	88.76	88.76	93.3	93.47
CNN	99.96	95.33	**95.32**	94.11	94.11	95.31	95.33

Effectiveness of stance feature in the classification of news articles.

Table 3. Effectiveness of stance feature in the classification of news articles.

Features	Models	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)	Precision (%)	Recall (%)	F1 (%)	ROC (%)
News Title, News Body	ANN	89.2	88.0	88.33	86.57	86.57	88.61	88.41
	LSTM	95.29	88.8	90.64	87.42	87.42	91.0	90.9
	Bi-LSTM	97.99	89.79	92.21	93.6	93.6	92.0	92.26
	CNN	99.1	93.68	93.90	91.3	91.3	94.0	94.0
News Title, News Body, Similarity between them (Stance)	ANN	89.31	89.37	89.38	86.4	86.4	89.8	89.4
	LSTM	94.36	89.05	91.06	87.8	87.8	91.44	91.37
	Bi-LSTM	98.6	92.11	92.6	93.5	93.5	92.5	92.6
	CNN	99.3	92.9	94.42	94.33	94.33	94.43	94.42

5-fold cross validation results for the proposed model

Table 4: 5-fold cross-validation results for the proposed model

Features	Models	Training Accuracy(%)	Validation accuracy(%)	Testing accuracy(%)	Precision (%)	Recall (%)	F1 (%)	ROC (%)
News Title, News Body, Similarity between them (Stance)	ANN	91.31	88.72	90.73	87.35	87.35	91.14	91.17
	LSTM	97.08	87.61	88.60	84.98	84.98	89.29	89.40
	Bi-LSTM	99.23	92.72	93.24	92.03	92.03	93.34	93.44
	CNN	99.92	95.25	95.85	94.81	94.81	95.89	95.90

Performance of various Fake News Identification models.

1   Table: 5 Performance of various Fake News Identification models.

Models	NB	RF	SVM	MLP	Boost	LSTM	Bi-LSTM	CNN
(Reddy et al., 2020)	86	82.5			95			
(Bali et al., 2019)	73.2	82.6	62	72.8	87.3			
(Bhutani et al., 2019)	84.3	83.9						
(Gravanis et al., 2019)	70		89					
(Bharadwaj & Shao, 2019)	90.7	94.7				92.7		
(George et al., 2020)	84.3	83.9				94.3		
(Esmaeilzadeh et al., 2019)						92.1	93.1	
(Huang & Chen, 2020)						84.9	87.7	91
(Khan et al., 2019) Char-LSTM	90					95	85	86
Our Model						91.16	93.05	**95.32**