Fake News Prediction*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Mandava SVPJ Sathvik
dept. name of organization (of Aff.)
name of organization (of Aff.)
Hyderabad, India
email address or ORCID
20bec024@iiitdwd.ac.in

2nd Prashant Kumar Singh
dept. name of organization (of Aff.)
name of organization (of Aff.)
Varanasi, UP, India
email address or ORCID
20bec033@iiitdwd.ac.in

3rd Boddu Moses Vijay Kumar
dept. name of organization (of Aff.)
name of organization (of Aff.)
Hyderabad, India
email address or ORCID
20bec010@iiitdwd.ac.in

4th Kanishk Bajpai
dept. name of organization (of Aff.)
name of organization (of Aff.)
Kanpur Nagar, UP, India
email address or ORCID
20bec020@iiitdwd.ac.in

4th Mukesh Kumar Mishra
dept. name of organization (of Aff.)
name of organization (of Aff.)
Jhansi, India
email address or ORCID
mukeshkumar@iiitdwd.ac.in

***Abstract*—The Internet and social media is open to everyone, it has become easy to acquire information and news. On the other hand we find news articles which are fake. It became a very tough task for the public to confirm them. It misleads too. So building intelligent systems for differentiating between news is important. We have implemented several deep learning techniques to detect false news. In this paper we used BERT, LSTM, Bi-LSTM, CNN Bi-LSTM deep learning models to detect the fake news on large datasets which includes fake news from social media and news articles. The text is classified into two, fake and real. We have trained 10 different machine learning algorithms, 4 deep learning algorithms and compared. The 10 machine learning models are Passive Aggressive Classifier,MLP Classifier, Logistic Regression, Naive Bayes, Decision Tree Classifier, Random Forest Classifier, KNeighbors Classifier, Support Vector Machine, Boosting, Adaptive Boosting and Gradient Boosting. We trained our models on three different datasets Covid-19, Gossipcop and politifact. The results obtained from the deep learning algorithms are highly accurate. The results from deep learning models were much better compared to machine learning models.**

***Index Terms*—fake news prediction, BERT, deep learning, deep learning**

I. Introduction

In this modern but still developing world, several great things have been achieved with the new advancement of technology updating everyday. One of those is Social media being extremely powerful in the current state. Nowadays, News plays an important role in our life, keeps us informed and updated about the incidents happening nearby us and even beyond our immediate environment. Due to Social media majorly and Internet, News is at everyone's fingertips. But not every information shared these days should be trusted blindly, at least not without checking properly. Some precaution can be done before as follows : - Consider the source Read beyond the headline for getting a whole picture Check the authors whether they are real and credible Check the supporting sources Check the date of publication to see if the story is relevant and up to date Observer whether your own biases are affecting your judgment Ask from independent people with knowledge to get confirmation But Mostly people don't take any of the above, leading them to face several consequences like affecting their decision making in several fields like financial (share markets) -¿in 2013, 130 billion dollars were wiped out in stock value after a false news spread on twitter that Barack Obama was injured in an explosion. Election -¿ According to a research poll, 64% of US citizens reported that fake news has caused a great deal of confusion" about the factual content of reported events. In the US presidential campaign of 2016, fake news has been accused of being foremost contributing factor of the increasing political polarization and partisan conflict as well as affecting the outcome Hence, reachability and spreading speed, therefore, strength of the news as well of fake news raised up enormously. In 2017, the inventor of the World Wide Web, Tim Berners-Lee claimed that fake news was one of the three most significant new disturbing Internet trends that must first be resolved, if the Internet is to be capable of truly "serving humanity." People rarely prefer traditional news over the one received through social media due to easy accessibility.

Identify applicable funding agency here. If none, delete this.

II. Methodology

+The training of the models is categorized into two. The first category is training traditional machine learning models. The machine learning models are Passive Aggressive Classifier, MLPClassifier, Logistic Regression, Multinomial Naive Bayes, Decision Tree Classifier, Random Forest Classifier, KNeighborsClassifier, Support Vector Machine, Adaptive Boosting and Gradient Boosting. The second category is training deep learning techniques LSTM, Bi LSTM, CNN-BiLSTM and BERT. The training requires several steps. The first step is the preprocessing step, in which removal of stop words, unnecessary data, lemmatization, tokenizing, lowercasing. The second step involves splitting the data into training data and testing data. The third step involves the feature extraction, for traditional machine learning models we have used TF-IDF for extracting features, for deep learning models we have used word embedding for extracting features. The fourth step involves the designing of the architecture of neural networks. The evaluation of the models is done by the metrics accuracy, precision, recall, f1-score, support and confusion matrix.

A. Data Collection

The training of the models are conducted with three fake news datasets. Those are gossipcop, politifact and CoAID (COVID-19 heAlthcare mIsinformation Dataset). Gossipcop and politifact are collected from FakeNewsNet.

1) **The Gossipcop dataset:** Gossipcop dataset has two files gossipcop real.csv, which has tweets of correct information over 5328 tweets. The second file gossipcop fake.csv has tweets of fake news over 5322 tweets. We merged both the files and labeled real news as 1 and fake news as 0. The final dataset has the columns including id, URL,title, tweet-id, and true. We have used the title as text.

2) **The politifact dataset:** The politifact dataset has two files politifact real.csv contains 432 tweets of real news, the politifact fake.csv contains 618 tweets of fake news. We merged both the files and labeled real news as 1 and fake news as 0. The final dataset has the columns including id, URL, title, tweet-id and true.

3) **CoAID (COVID-19 heAlthcare mIsinformation Dataset):** is a dataset of tweets and articles circulated over covid-19. The dataset contains 5486 tweets and news articles regarding covid-19.

B. Data Processing

Data preprocessing is crucial for training in neural networks and machine learning models. The tweets have data which is unstructured. Deep learning and machine learning algorithms require structured data for the best performance. The raw data has to be structured and refined to make it machine readable.

1) **Lower casing:** It is converting capital letters into small letters. Example: ‘NLP’ to ‘nlp’

2) **Removing unimportant data:** removing unimportant data which are not required for machine learning and deep learning like full stops, question marks, commas etc.

3) **Tokenization:** is the process of segregating words from sentences. Example: ‘I am a student’ to ‘I’ ‘am’ ‘a’ ‘student’

4) **Removal of Stop Word:** Stop words are most repetitive words and have no significance like conjunctions, prepositions etc. These types of words are removed from all the three datasets.

5) **Lemmatization:** is used to reduce the actual word to root word for avoiding the excess of word types. For example: ‘ate’, ‘eaten’, ’eating’ to eat.

C. Data Splitting

In data splitting the 90% data used for training and 10% data used for testing. The training data is used to train for the input and output for the machine learning model. The training data is the unseen data for the trained model and evaluation is done on the training data.

D. Applying optimization and machine learning models

Machine learning feature extraction: TF-IDF(Term Frequency – Inverse Document Frequency): is used for extracting features of machine learning models. The TF-IDF converts each word in the form of vectors. The value for each word depends on the number of times the word is repeating in the data. TF(t, d) = (Number of occurrences of term t in document d)/ (Total number of terms in the document d) IDF(t) = loge (Total number of documents in the corpus) (Number of documents with term t in them)

2)Machine Learning models: We have used ten machine learning models PassiveAggressiveClassifier,MLPClassifier, LogisticRegression,MultinomialNB, DecisionTreeClassifier, RandomForestClassifier, KNeighborsClassifier, Support Vector Machine, Adaptive Boosting and Gradient Boosting. Passive Aggressive Classifier: is an online machine learning algorithm. Once the algorithm is predicted and receives the feedback on the prediction and then modifies the mechanism going through several rounds to improve the prediction. Logistic Regression: Logistic Regression is a machine learning model used for classification. It works on the concept of probability and it uses sigmoid function. The output is the probability varies from zero to one. The input can be varied from -infinity to +infinity. In predicting binary classification if the predicted probability is greater than 0.5 it’s true if lesser than 0.5 then false. Decision Tree: Decision Tree is a machine learning algorithm that can be used for binary classification. It has a hierarchical tree structure having nodes and branches. The branches contain the comparison of features and pattern whereas leaf nodes contain the decision whether the news is fake or real.

Random Forest: The Random Forest is a supervised machine learning combination of decision trees. It predicts them by averaging the decision trees. As it uses the combination of decision trees it’s performance is significantly improved. It is one of the best machine learning algorithms for classification. It uses simple decision rules to predict the output in our experiment, fake and real. Adaptive Boosting: Adaptive Boosting is a machine learning algorithm that performs well on binary classification. The boosting algorithms make the weaker predictions and adding more weights simply gives more preference and in the next step the wrong predictions will have higher weights. The training is done till the loss becomes very low and eventually develops better predictions. Gradient Boosting: Gradient implies “Gradient descent”. It can be used with a differentiable loss function. The Adaptive Boosting uses an exponential loss function. This makes Gradient Boosting perform better than Adaptive Boosting. Extreme Gradient Boosting: is an ensemble training technique. It is a group of decision trees in sequential order. The falsely predicted predictions are given more weights and passed to the next decision tree. The individual decision trees then are grouped to give strong predictions. KNeighbors: is a supervised machine learning model that classifies the algorithms by grouping the similar data points. It groups the data which are similar in nature. It uses euclidean distance for grouping and finding similar data points.

Support Vector Machine: The support vector machine is a supervised machine learning model that is capable of performing nonlinear classifications. The data points are plotted on an n-dimensional plane and are converted into high dimensional feature spaces. The prediction is done by the hyperplane best dividing categories.

Naive Bayes: The Naive Bayes is a probabilistic machine learning algorithm that uses naives theorem for the predictions. It makes classification using the Maximum Posterior decision rules in a Bayesian setting.

Deep Learning: We applied four deep learning models. These are done in two steps. The first step is to extract features from the text using a deep learning feature extraction method. Feature extraction: The neural networks used in deep learning models are trained only on numbers and float values. The neuron cannot take direct text as input. Therefore it is important to convert them in the form of vectors. Simply we are convert the natural language to machine understandable language. We used Gensim library in this experiment.

Deep neural network: The figure shows the working architecture of our experiment. That is used for the detection of fake news. The embedded matrix represents the features of the news in the dataset. The Word embedding matrix is embedded into the embedding layer and then to the hidden layer. The hidden layer is the core for the entire experiment. The most important operations are done in the hidden layers in our experiment the LSTM, BiLSTM, CNN-BiLSTM and BERT. LSTM: Long Short Term Memory is a deep learning algorithm. LSTM has the memory block and three multiplicative gating units. It is capable of memorizing the previous sentences and previous text. The neural network which is built on dense neurons cannot memorize the previous sentences. Forget Gate Layer: gt as shown in figure decides which information to exclude from cell state. gt = (Wg · [Ht1, xt] + bg) The new information will be added at input gate layer nt as shown in the equation nt = (Wn · [Ht1, xt] + bn) Then decides which values to get updated S˜t = tanh (Ws · [Ht1, xt] + bs) Now, the cell has to be updated, Updating the cell follows the below mathematical equation St = gt St1 + it S˜ t Then the weights are forwarded to output gate layer. The output ot follows the equation ot = (Wo [Ht1, xt] + bo) Then it is forwarded to tanh layer so that the values range from -1 to +1. Ht = ot tanh (St) Peephole: Let the gate layer look at the cell state as shown in the following Equations gt = Wg · [St1, Ht1, xt] + bg) nt = (Wn · [St1, Ht1, xt] + bn) ot = (Wo · [St, Ht1, xt] + bo) Coupling forgot and input gates follow the equation St = gt St1 + (1 gt) S˜

BiLSTM is a deep learning neural network that can make utilize of past and future features of the text. Simply, it has ability to understand the previous words and future words and can differentiate accordingly. The BiLSTM contains two LSTM layers and are interconnected to each other to extract features of the upcoming words and previous words. One LSTM layer is for future features and the other for extracting past features as shown in the figure.

CNN-BiLSTM The CNN-BiLSTM is a deep neural network. The convolutional neural network is popular for extracting features from text and images. The convolutional neural networks has significant importance in computer vision. They are used for most of the computer vision applications from simple application to GAN(Generative Adversial Neural networks). Researches also show that convolutional neural networks perform well on text as well. In this neural network the text is passed into convolutional neural networks to extract the features and then passed into BiLSTM layer. Simply, the input of BiLSTM is extracted features from CNN layers.

III. Prepare Your Paper Before Styling

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections III-A–III-E below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads—LaTeX will do that for you.

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as "3.5-inch disk drive".
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: "Wb/m^2" or "webers per square meter", not "webers/m^2". Spell out units when they appear in text: ". . . a few henries", not ". . . a few H".
- Use a zero before decimal points: "0.25", not ".25". Use "cm^3", not "cc".)

C. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \tag{1}$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(1)", not "Eq. (1)" or "equation (1)", except at the beginning of a sentence: "Equation (1) is . . ."

D. LaTeX-Specific Advice

Please use "soft" (e.g., `\eqref{Eq}`) cross references instead of "hard" references (e.g., `(1)`). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

Please don't use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Please note that the `{subequations}` environment in LaTeX will increment the main equation counter even when there are no equation numbers displayed. If you forget that, you might write an article in which the equation numbers skip from (17) to (20), causing the copy editors to wonder if you've discovered a new method of counting.

BibTeX does not work by magic. It doesn't get the bibliographic data from thin air but from .bib files. If you use BibTeX to produce a bibliography you must send the .bib files.

LaTeX can't read your mind. If you assign the same label to a subsubsection and a table, you might find that Table I has been cross referenced as Table IV-B3.

LaTeX does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it's supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` (there won't be any anyway) and it might stop a wanted equation number in the surrounding equation.

E. Some Common Mistakes

- The word "data" is plural, not singular.
- The subscript for the permeability of vacuum μ_0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter "o".
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an "inset", not an "insert". The word alternatively is preferred to the word "alternately" (unless you really mean something that alternates).
- Do not use the word "essentially" to mean "approximately" or "effectively".
- In your paper title, if the words "that uses" can accurately replace the word "using", capitalize the "u"; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones "affect" and "effect", "complement" and "compliment", "discreet" and "discrete", "principal" and "principle".
- Do not confuse "imply" and "infer".
- The prefix "non" is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the "et" in the Latin abbreviation "et al.".
- The abbreviation "i.e." means "that is", and the abbreviation "e.g." means "for example".

An excellent style manual for science writers is [7].

F. Authors and Affiliations

The class file is designed for, but not limited to, six authors. A minimum of one author is required for all conference articles. Author names should be listed starting from left to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

G. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is "Heading 5". Use "figure caption" for your Figure captions, and "table head" for your table title. Run-in heads, such as "Abstract", will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced.

H. Figures and Tables

a) Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

TABLE I
TABLE TYPE STYLES

Table Head	**Table Column Head**		
	Table column subhead	***Subhead***	***Subhead***
copy	More table copy[a]		

[a] Sample of a Table footnote.

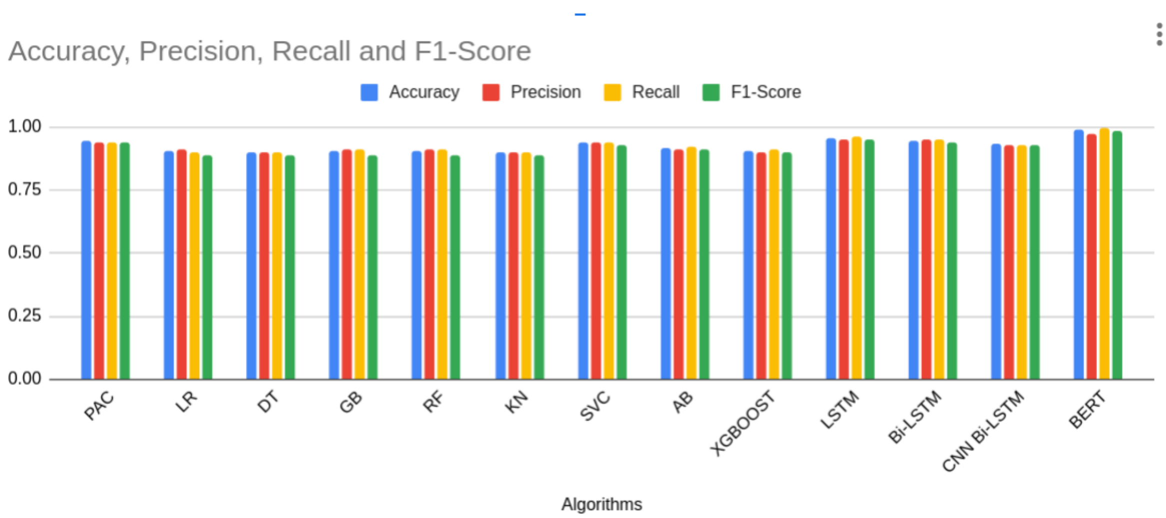

Fig. 1. Example of a figure caption.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity "Magnetization", or "Magnetization, M", not just "M". If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write "Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in America is without an "e" after the "g". Avoid the stilted expression "one of us (R. B. G.) thanks ...". Instead, try "R. B. G. thanks...". Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

Please number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use "Ref. [3]" or "reference [3]" except at the beginning of a sentence: "Reference [3] was the first ..."

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published, even if they have been submitted for publication, should be cited as "unpublished" [4]. Papers that have been accepted for publication should be cited as "in press" [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

REFERENCES

[1] KaiDMML, K. S. (2019). KaiDMML/FakeNewsNet: This is a dataset for fake news detection research. GitHub. Retrieved November 24, 2022, from https://github.com/KaiDMML/FakeNewsNet

[2] DIAA SALAMA ABDELMINAAM, FATMA HELMY ISMAIL, MOHAMED TAHA1, AHMED TAHA1, ESSAM H. HOUSSEIN, AND AYMAN NABIL "CoAID-DEEP: An Optimized Intelligent Framework for Automated Detecting COVID-19 Misleading Information on Twitter ," IEEE Access Volume: 9, pp. 27840 - 27867, Feb 9, 2021. doi: 10.1109/ACCESS.2021.3058066 [Online]. Available: https://ieeexplore.ieee.org/document/9350542/authorsfull-text-header

[3])Sammut, C., Webb, G. I. (2011). Encyclopedia of Machine Learning. Springer.

[4] Yu, T., Wei, JM., Li, J. (2012). PAC Learnability of Rough Hypercuboid Classifier. In: Huang, DS., Ma, J., Jo, KH., Gromiha, M.M. (eds) Intelligent Computing Theories and Applications. ICIC 2012. Lecture Notes in Computer Science(), vol 7390. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-31576-3$_8$2

[5] Cramer, J. S. (2002, November). Cramer, J.S. (2002) The Origins of Logistic Regression, vol. 119. - references - scientific research publishing. Retrieved November 24, 2022, from https://scirp.org/reference/referencespapers.aspx?referenceid=3247626

[6] Quinlan, J. R. (1986, March). Induction of decision trees - machine learning. SpringerLink. Retrieved November 20, 2022, [online]. Available : https://link.springer.com/article/10.1007/BF00116251

[7] L. Cui and D. Lee, "CoAID: COVID-19 healthcare misinformation dataset," 2020, arXiv:2006.00885. [Online]. Available:http://arxiv.org/abs/2006.00885

[8] Breiman, L. (2001, October). Random forests - machine learning. SpringerLink. Retrieved November 20, 2022, from https://link.springer.com/article/10.1023/A:1010933404324 [Online]. Available: https://link.springer.com/article/10.1023/A:1010933404324

[9] Freund, Y., amp; Sachapire, R. E. (1996). Experiments with a new boosting algorithm - University of California ... boosting experiments.pdf. Retrieved November 24, 2022, from https://cseweb.ucsd.edu/ yfreund/papers/boostingexperiments.pdf

[10] Friedman, J. H. (2001, October). Greedy Function Approximation: A Gradient Boosting Machine. JSTOR. Retrieved November 24, 2022, from https://www.jstor.org/stable/2699986

[11] Chen, T., Guestrin, C. (2016, June 10). XGBoost: A scalable tree boosting system. arXiv.org. Retrieved November 20, 2022, [online]. Available : https://arxiv.org/abs/1603.02754

[12] J. Laaksonen and E. Oja, "Classification with learning k-nearest neighbors," Proceedings of International Conference on Neural Networks (ICNN'96), 1996, pp. 1480-1483 vol.3, doi: 10.1109/ICNN.1996.549118. [Online]. Available: https://ieeexplore.ieee.org/document/549118

[13] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt and B. Scholkopf, "Support vector machines," in IEEE Intelligent Systems and their Applications, vol. 13, no. 4, pp. 18-28, July-Aug. 1998, doi: 10.1109/5254.708428. [online]. Available: https://ieeexplore.ieee.org/document/708428/

[14] Kibriya, A.M., Frank, E., Pfahringer, B., Holmes, G. (2004). Multinomial Naive Bayes for Text Categorization Revisited. In: Webb, G.I., Yu, X. (eds) AI 2004: Advances in Artificial Intelligence. AI 2004. Lecture Notes in Computer Science(), vol 3339. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30549-$1_4 3 Chiu, J.P.C., Nichols, E. (2016, July 19). Namedentityrecognitionwithbidirect$ $CNNS.arXiv.org.RetrievedNovember 20, 2022, [online].Available : https : //arxiv.org/abs/1511.08308v5$

[16] Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. (2019, May 24). Bert: Pre-training of deep bidirectional Transformers for language understanding. arXiv.org. Retrieved November 20, 2022, [Online], Available: https://arxiv.org/abs/1810.04805

[17] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. (2017, December 6). Attention is all you need. arXiv.org. Retrieved November 20, 2022, from https://arxiv.org/abs/1706.03762

[18] Goutte, C., Gaussier, E. (2005). A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In: Losada, D.E., Fernández-Luna, J.M. (eds) Advances in Information Retrieval. ECIR 2005. Lecture Notes in Computer Science, vol 3408. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-31865-$1_2 5 Allsymbolsintensorflow2nbsp; : nbsp; tensorflowV2.11.0.TensorFlow.(n.d.).RetrievedNovember25, 2022, fromh$ $//www.tensorflow.org/api_docs/python/tf/all_symbols$

[19] Documentation of scikit-learn 0.21.3¶. learn. (n.d.). Retrieved November 25, 2022, from https://scikit-learn.org/0.21/documentation.html

[20] Cuilimeng. (2021, January 12). Cuilimeng/CoAID. GitHub. Retrieved November 25, 2022, from https://github.com/cuilimeng/CoAID

[21] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186. [Online] Available: https://ieeexplore.ieee.org/document/8308186

[22] Okoro, Efeosasere Abara, Benjamin Alex, Umagba Ajonye, Anyalewa Isa, Zayyad. (2018). A hybrid approach to fake news detection on social media. Nigerian Journal of Technology. 37. 10.4314/njt.v37i2.22. [online] Available: https://www.researchgate.net/publication/324151369$_{A_h}ybrid_approach_to_fake_news_detect$ $W.On," FakeNewsStanceDetectionUsingDeepLearningArchitecture(CNN- LSTM)," inIEEEAccess, vol.8, pp.156695 - 156706, 2020, doi : 10.1109/ACCESS.2020.3019735.[Online]Available : https : //ieeexplore.ieee.org/document/9178321$

[23] Chiu, J. P. C., amp; Nichols, E. (2016, July). Papers with code - CNN BILSTM explained. CNN BiLSTM Explained — Papers With Code. Retrieved November 25, 2022, from https://paperswithcode.com/method/cnn-bilstm

[24] A. Aziz Sharfuddin, M. Nafis Tihami and M. Saiful Islam, "A Deep Recurrent Neural Network with BiLSTM model for Sentiment Classification," 2018 International Conference on Bangla Speech and Language Processing (ICB-

SLP), 2018, pp. 1-4, doi: 10.1109/ICBSLP.2018.8554396. [Online] Available : https://ieeexplore.ieee.org/document/8554396
[25] G. Xu, Y. Meng, X. Qiu, Z. Yu and X. Wu, "Sentiment Analysis of Comment Texts Based on BiLSTM," in IEEE Access, vol. 7, pp. 51522-51532, 2019, doi: 10.1109/ACCESS.2019.2909919. [Online] Available : https://ieeexplore.ieee.org/document/8684825
[26] Vaijala, S., Majumder, B., Gupta, A., amp; Surana, H. (2020). Practical natural language processing a comprehensive guide to building real-world Nlp systems. O'Reilly.
[27] Tunstall, L., Werra, L. van, amp; Wolf, T. (2022). Natural language processing with transformers: Building language applications with hugging face. O'Reilly.
[28] Bird, S., Klein, E., amp; Loper, E. (2009). Natural language processing with python. O'Reilly Media.
[29] C. Rajalakshmi, T. Subika, K. Vaishali and J. Shana, "Fake News Prediction On COVID Dataset Using Machine Learning," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 1-7, doi: 10.1109/ICCCNT51525.2021.9579543. [Online] Available : https://ieeexplore.ieee.org/document/9579543
[30] Wutonytt, T. H. Y. (2021, July 18). WUTONYTT/Fake-news-detection: Fake news detection using a BERT-based deep learning approach. Wutonytt/Fake-News-Detection' . Retrieved November 25, 2022, from https://github.com/wutonytt/Fake-News-Detection
[31] R. K. Kaliyar, M. Agrawal and A. Goswami, "FndIP: Fake News Detection on Social Media Using Incompatible Probabilistic Method," 2022 International Conference on Inventive Computation Technologies (ICICT), 2022, pp. 417-422, doi: 10.1109/ICICT54344.2022.9850787. [Online] Available : https://ieeexplore.ieee.org/document/9850787
[32] S. S S, K. N, K. M N, A. K S, B. K. S and R. K K, "Authentic News Prediction in Machine Learning using Passive AggressiveAlgorithm," 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS), 2022, pp. 372-376, doi: 10.1109/ICAIS53314.2022.9743010. [Online] Available : https://ieeexplore.ieee.org/document/9743010
[33] S. M. N, K. M. V, S. Verma and S. Rajagopal, "NLP Based Fake News Detection Using Hybrid Machine Learning Techniques," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), 2022, pp. 818-822, doi: 10.1109/ICESC54411.2022.9885679. [Online] Available : https://ieeexplore.ieee.org/document/9885679
[34] S. S. Reddy, S. Mandal, V. L. V. S. K. B Kasyap and A. R K, "A Novel Approach to Detect Fake News Using eXtreme Gradient Boosting," 2022 10th International Symposium on Digital Forensics and Security (ISDFS), 2022, pp. 1-4, doi: 10.1109/ISDFS55398.2022.9800777. [Online] Available : https://ieeexplore.ieee.org/document/9800777
[35] A. Shete, H. Soni, Z. Sajnani and A. Shete, "Fake News Detection Using Natural Language Processing and Logistic Regression," 2021 2nd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS), 2021, pp. 136-140, doi: 10.1109/ACCESS51619.2021.9563292. [Online] Available : https://ieeexplore.ieee.org/document/9563292/
[36] S. A. Alameri and M. Mohd, "Comparison of Fake News Detection using Machine Learning and Deep Learning Techniques," 2021 3rd International Cyber Resilience Conference (CRC), 2021, pp. 1-6, doi: 10.1109/CRC50527.2021.9392458. [Online] Available : https://ieeexplore.ieee.org/document/9392458
[37] Patterson, J., amp; Gibson, A. (2017). Deep learning: A practitioner's approach. O'Reilly.
[38] J. U. Rahman, S. Ali, A. Khan and M. Fayaz, "A Deep Ensemble Model for News Classification on Social Media," 2021 International Conference on Cyber Warfare and Security (ICCWS), 2021, pp. 57-62, doi: 10.1109/ICCWS53234.2021.9702934. [online] Available : https://ieeexplore.ieee.org/document/9702934
[39] S. M and K. K. P, "Estimation of Precision in Fake News Detection Using Novel Bert Algorithm and Comparison with Random Forest," 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT), 2022, pp. 217-223, doi: 10.1109/ICICICT54557.2022.9917629. [Online] Available : https://ieeexplore.ieee.org/document/9917629
[40] V. Tiwari, R. G. Lennon and T. Dowling, "Not Everything You Read Is True! Fake News Detection using Machine learning Algorithms," 2020 31st Irish Signals and Systems Conference (ISSC), 2020, pp. 1-4, doi: 10.1109/ISSC49989.2020.9180206.[Online] Available : https://ieeexplore.ieee.org/document/9180206
[41] H. Patel, I. Dave, D. Furiya and V. Hirlekar, "Fake Tweet Detection Using Machine Learning," 2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), 2021, pp. 1-5, doi: 10.1109/SMARTGENCON51891.2021.9645809. [Online] Available : https://ieeexplore.ieee.org/document/9645809
[42] S. Tyagi, A. Pai, J. Pegado and A. Kamath, "A Proposed Model for Preventing the spread of misinformation on Online Social Media using Machine Learning," 2019 Amity International Conference on Artificial Intelligence (AICAI), 2019, pp. 678-683, doi: 10.1109/AICAI.2019.8701408. [Online] Available : https://ieeexplore.ieee.org/document/8701408
[43] GERON, A. (2017). Hands-on machine learning with scikit-learn and tensorflow: Concepts, tools and techniques to build Intelligent Systems. O'Reilly.
[44] GERON, A. (2019). Hands-on machine learning with scikit-learn, Keras, and tensorflow: Concepts, tools and techniques to build Intelligent Systems (2nd ed.). O'Reilly.
[45] Muller Andreas Christian, amp; Guido, S. (2018). Introduction to machine learning with python a guide for Data scientists. O'Reilly.
[46] Z. A. Guven, "Comparison of BERT Models and Machine Learning Methods for Sentiment Analysis on Turkish Tweets," 2021 6th International Conference on Computer Science and Engineering (UBMK), 2021, pp. 98-101, doi: 10.1109/UBMK52708.2021.9559014. [Online] Available : https://ieeexplore.ieee.org/document/9559014
[47] N. Daoud, M. Eltahan and A. Elhennawi, "Aerosol Optical Depth Forecast over Global Dust Belt Based on LSTM, CNN-LSTM, CONV-LSTM and FFT Algorithms," IEEE EUROCON 2021 - 19th International Conference on Smart Technologies, 2021, pp. 186-191, doi: 10.1109/EUROCON52738.2021.9535571. [Online] Available : https://ieeexplore.ieee.org/document/9535571
[48] Patterson, J., amp; Gibson, A. (2017). Deep learning: A practitioner's approach (1st ed.). O'Reilly.
[49] Alaparthi, S., Mishra, M. (2021, February 26). Bert: A sentiment analysis odyssey - journal of marketing analytics. SpringerLink. Retrieved November 23, 2022, from https://link.springer.com/article/10.1057/s41270-021-00109-8
[50] Kula, S., Kozik, R., Choraś, M. (2021, July 22). Implementation of the Bert-derived architectures to tackle disinformation challenges - neural computing and applications. SpringerLink. Retrieved November 23, 2022, from https://link.springer.com/article/10.1007/s00521-021-06276-0
[51] Jiang, L.-G., Chan, Y.-H., Fan, Y.-C., Leu, F.-Y. (1970, January 1). You draft we complete: A Bert-based article generation by Keyword Supplement. SpringerLink. Retrieved November 23, 2022, from https://link.springer.com/chapter/10.1007/978-3-030-79728-$7_38)VanHoudt, G., Mosquera, C., Nápoles, G.(2020, May13).Areviewonthelongs$ $termmemorymodel-artificialintelligencereview.SpringerLink.RetrievedNove$ $//link.springer.com/article/10.1007/s10462-020-09838-1$
[52] Soutner, D., Müller, L. (1970, January 1). Application of LSTM neural networks in language modelling. SpringerLink. Retrieved November 23, 2022, from https://link.springer.com/chapter/10.1007/978-3-642-40585-$3_14Lee, M.-C., Chang, J.-W., Yeh, S.-C., Chia, T.-L., Liao, J.-S., Chen, X.-M.(2022, January28).Applyingattention-basedBILSTMandtechnicalindicatorsinthedesignandperformanceanalysisofS$ $neuralcomputingandapplications.SpringerLink.RetrievedNovember23, 2022, fr$ $//link.springer.com/article/10.1007/s00521-021-06828-4$
[53] Raza, S., Ding, C. (2022, January 30). Fake news detection based on news content and social contexts: A transformer-based approach - International Journal of Data Science and Analytics. SpringerLink. Retrieved November 23, 2022, from https://link.springer.com/article/10.1007/s41060-021-00302-z
[54] Reddy, S. (2021, May 7). Fake-news-detection-using-machine-learning/fakenewsdetection.ipynb at master · sushwanthreddy/fake-news-detection-using-machine-learning. Fake-News-Detection-using-Machine-Learning. Retrieved November 25, 2022, from https://github.com/SushwanthReddy/Fake-News-Detection-using-Machine-Learning/blob/master/FakeNewsDetection.ipynb
[55] Uzairrj. (2022, June 24). Uzairrj/fake-news-classifier. GitHub-uzairrj/Fake-News-Classifier. Retrieved November 25, 2022, from https://github.com/uzairrj/Fake-News-Classifier
[56]

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove the template text from your paper may result in your paper not being published.