JOIV : Int. J. Inform. Visualization, 8(2) - May 2024 643-651

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

A Multi-tier Model and Filtering Approach to Detect Fake News Using Machine Learning Algorithms

Chiung Chang Yu [a,*], Isredza Rahmi A Hamid [a], Zubaile Abdullah [a], Kuryati Kipli [b], Hidra Amnur [c]

[a] *Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, Batu Pahat, Malaysia*
[b] *Department of Electrical and Electronic Engineering, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia*
[c] *Department of Information Technology, Politeknik Negeri Padang, Padang, Indonesia*
*Corresponding author: *chiungyu24@gmail.com*

***Abstract*—Fake news trends have overgrown in our societies over the years through social media platforms. The goal of spreading fake news can easily mislead and manipulate the public's opinion. Many previous researchers have proposed this domain using classification algorithms or deep learning techniques. However, machine learning algorithms still suffer from high margin error, which makes them unreliable as every algorithm uses a different way of prediction. Deep learning requires high computation power and a large dataset to operate the classification model. A filtering model with a consensus layer in a multi-tier model is introduced in this research paper. The multi-tier model filters the news label correctly predicted by the first two-tier layer. The consensus layer acts as the final decision when collision results occur in the first two-tier layer. The proposed model is applied to the WEKA software tool to test and evaluate the model from both datasets. Two sequences of classification models are used in this research paper: LR_DT_RF and LR_NB_AdaBoost. The best performance of sequence for both datasets is LR_DT_RF which yields 0.9892 F1-Score, 0.9895 Accuracy, and 0.9790 Matthews Correlation Coefficient (MCC) for ISOT Fake News Dataset, and 0.9913 F1-Score, 0.9853 Accuracy, and 0.9455 MCC for CHECKED Dataset. This research could give researchers an approach for fake news detection on different social platforms and feature-based.**

***Keywords*— Consensus layer; fake news; machine learning; multi-tier model.**

Manuscript received 14 Nov. 2023; revised 29 Jan. 2024; accepted 20 Mar. 2024. Date of publication 31 May. 2024.

I. Introduction

Fake news contains misleading information and fabricated content, deliberately created to deceive readers. With the proliferation of social media platforms where individuals share daily updates, the spread of fake news has become alarmingly fast, resulting in significant consequences within public discourse [1]. Determining the impact of fake news is complex [2]. Most people unknowingly become intermediaries, inadvertently disseminating misinformation to their communities [3], [4], [5]. The negative impact of fake news includes distorting truth and facts, damaging reputations, and fueling hate speech and conflicts on social media [6], [7]. Preserving the integrity of information is vital to combat the issue of fake news to protect an individual's credibility and reputation from intentional attackers [8], [9], [10].

Various researchers used machine learning techniques to tackle the challenge of detecting fake news [11], [12]. Work by [11] summarized machine learning algorithms such as XGBoost, Random Forests (RF), Naïve Bayes (NB), K-Nearest Neighbours (KNN), Decision Tree (DT), and Support Vector Machine (SVM) to classify fake news, achieved over 75 percent accuracy across all employed algorithms. However, work by [12] found that when applied to small datasets, the Logistic Regression (LR) model achieved only a success rate of 63 percent to 70 percent in clustering and predictive modeling. Consequently, machine learning algorithm approaches often suffer from a high margin error, rendering these models unreliable. As the results differ, overfitting or underfitting a dataset also happens in machine learning algorithms. In contrast, a proposed deep learning model called FakeBERT [13] demonstrated a 98.90 percent accuracy for fake news detection based on a large dataset. However, deep learning models require substantial datasets and significant computational power, which are highly dependent on model complexity. Additionally, the black box problem associated with deep learning models hampers the interpretability of their outputs.

To address the challenges posed by machine learning models' error rates and the black box problem, we propose a multi-tier fake news detection model based on machine learning algorithms. The objectives of this research are:

- To identify the content-based features of fake news based on news content and user profiles using machine learning algorithms.
- To develop a multi-tier model and filtering approach to detect and classify fake news.
- To evaluate the performance of the proposed multi-tier model and filtering approach in terms of F1-Score, Accuracy, and Matthews Correlation Coefficient (MCC).

This research considers the ISOT Fake News Dataset [14] and the CHECKED Dataset [15]. Features are selected separately for each dataset, with the ISOT Fake News Dataset focusing on news content and the CHECKED Dataset relying on news content and user profiles. Consensus filtering predicts data not filtered between the first and second tiers. This consensus layer operates independently and is not influenced by decisions made in the preceding tiers. The proposed multi-tier fake news detection model is evaluated using F1-Score, Accuracy, and MCC.

II. Material and Method

Various research work on fake news detection, such as [11], [16], [17], and [18], has been done in recent years. Machine learning assists researchers in determining fake news using features defined by researchers. The computer observes the data patterns from the features we extracted in the training set and predicts better. Machine learning can also let the computer automate-learn the dataset without human control.

The study by [11] used a supervised machine learning algorithm to detect fake news, focusing on the feature extraction process. They utilized the Python sci-kit-learn library for tokenization and feature extraction due to its beneficial methods, such as Count Vectorizer and Tiff Vectorizer, used in the data pre-processing step. The chosen dataset for the study was the LIAR dataset, which consists of political news articles collected from fact-checking websites. The study employed approaches based on unigram and bigram features to extract lexical features, utilizing TD-IDF and n-gram features, which could subsequently be applied to the algorithms. All the algorithms employed achieved an accuracy rate of over 75 percent.

Work by [16] proposed the trust network construction recommendation step to detect fake news and the multiclassification approach using unlabeled data. The datasets were from the News Dataset from GitHub [19] and the Getting Real about Fake News Dataset from Kaggle [20]. Then, they were divided into four categories for multiclass classification: satire, propaganda, manufacturing, and manipulation. The news content-based feature was applied for the classification algorithms inside the trust network construction to predict the dataset. It consists of two prediction steps before going to the proposed recommendation step. After the final prediction, the final label of the unlabeled news is re-calculated according to the newly constructed trust network. The algorithm produces a list of news types to recommend to users and predicts the rates of unlabeled news. The evaluation step is repeated three times to determine the best algorithm. The logistic Regression (LR) algorithm with the BERT contextual approach is the best model, which achieved a 96 percent accuracy rate compared to other algorithms using the proposed recommendation phase approach. However, this proposed system is a sequential pipeline where the dataset is entirely dependent on each stage one by one.

Work by [17] introduced a general framework to identify the polarized news content on social media and predict future fake news topics. The Italian Facebook dataset is used in this study, which contains approximately 385,000 post samples and 104,173 entities after conducting topic extraction and sentiment analysis such as satire, propaganda, manufacturing, manipulation, and bias form. Several classification algorithms were applied, including Linear Regression (LiR), Logistic Regression (LR), SVM with a linear kernel, K-Nearest Neighbors (KNN), Neural Network Models using the Multi-layer Perceptron-L-BFGS algorithm, and Decision Trees with Gini Index. Based on the accuracy rate obtained from the comparative results, the LR algorithm demonstrated the best performance achieving a 91 percent accuracy rate.

The study by [18] proposed a content-based approach for fake news detection, using features from both the news content and social context. They proposed harmonic Boolean label crowdsourcing (HC) on social signals (HC-CB) and LR on social signals (LR-CB) as a comparison, defining a threshold value to classify posts using the content-based approach. Three datasets were used in this research to compare their proposed method to the previous method: a Facebook dataset and two datasets collected from FakeNewsNet [8]. The researchers also analyzed sensitivity to investigate the relationship between the threshold value and the classifier's accuracy. Then, they implemented the proposed detection model to chatbot with a completely independent real-world dataset and observed the results. The performance of the proposed fake news detection model, which is HC-CB, was evaluated using the accuracy performance metric, resulting in an accuracy rate of 81.7 percent when the threshold value is 4 using the real-world dataset. However, this proposed model is only trained in the Italian language, and the dataset size is small, consisting of 230 total Facebook posts.

TABLE I
EXISTING WORKS ON FAKE NEWS DETECTION APPROACH

Author	Dataset	Algorithm	Feature	Accuracy Result
Khanam et al. [4]	LIAR	LR, RF, XGBoost, NB, KNN, DT, SVM	Lexical features	All algorithms with more than 75%
Stitini et al. [8]	GitHub and Kaggle	LR, NB, DT, Linear SVM	Content-based	LR: 91%
Vedova et al. [9]	Facebook Dataset, FakeNewsNet	LR-CB and HC-CB	Content-based	HC-CB: 81.7%
Vicario et al. [10]	Italian Facebook	LiR, LR, SVM, KNN, NN, DT	Topic extraction and sentiment analysis	LR: 91%

Table I shows the comparative analysis based on existing research on fake news detection. All studies present innovative approaches to tackle the problem of fake news detection. However, further exploration and clarification regarding the methodologies and factors influencing the results would enhance the understanding and applicability of the fake news detection technique. Our work differs from previous research in such a way that we proposed the fake news detection approach by using the content-based features for the ISOT Fake News Dataset and a user profile-based features for the CHECKED Dataset. These datasets can be obtained from GitHub and Kaggle [14], [15], [21] providing a comprehensive and diverse range of data for our analysis. Moreover, we implemented a consensus filtering mechanism within our multi-tier fake News Detection Model. This approach ensures a higher level of accuracy by incorporating multiple layers of analysis and filtering, thereby reducing the likelihood of false positives or false negatives. Additionally, we conducted extensive experiments to evaluate the performance of our proposed multi-tier fake News Detection Model using existing machine learning algorithms: Random Forest, Naïve Bayes, Logistic Regression, AdaBoost, and DT algorithm. By combining innovative feature extraction techniques, consensus filtering, and rigorous algorithmic testing, our research contributes to advancing fake news detection methodologies. Our work provides valuable insights and practical solutions for addressing the challenges posed by the proliferation of fake news in today's information landscape.

III. Results and Discussion

This section discusses the methodology of a multi-tier fake news detection model with a filtering approach using machine learning algorithms.

A. Multi-tier Fake News Detection Model

The experiment involved seven distinct phases: raw data, data pre-processing, feature extraction, and feature selection, multi-tier filtering model, training and test the data using 10-fold cross-validation, parameter tuning the classification algorithms, and the final determination of whether a news article is true or fake, as depicted in Fig. 1. We used Waikato Environment for Knowledge Analysis (WEKA) tool to evaluate the performance of multi-tier filtering fake news detection model. The evaluation metrics utilized in this assessment were F1-Score, Accuracy, and Matthews Correlation Coefficient (MCC).

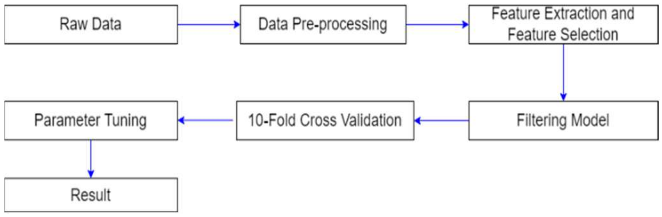

Fig. 1 Fake News Detection Approach

1) Raw Data: Raw data refers to unprocessed data that still needs to be in a format readable by computers but can be understood by humans. This data may include user inputs or computer-generated values such as timestamps. We considered two datasets, the ISOT Fake News Dataset [14] and [21] and the CHECKED Dataset [15]. The ISOT Fake News Dataset has been previously utilized by [22], [23], [24], [25], [26] to evaluate the performance of machine learning algorithms. It was also employed by [25] on fake news using capsule neural networks. An example of raw data from the ISOT Fake News Dataset is shown in Fig. 2. On the other hand, the CHECKED Dataset has been employed by [27] in their analysis of extended text feature extraction networks with data augmentation (LTFE) and other research papers [28], [29], [30]. ISOT Fake News Dataset has 21417 real news and 23481 fake news, with a total of 44898 news. Next, the CHECKED Dataset has 1760 true microblogs and 344 fake microblogs with 2104. The dataset summary includes real and fake news, and Table II calculates and presents the total number of each. Examples of raw data from the CHECKED Dataset are shown in Fig. 3.

	A	B	C	D
1	As U.S. budget fight looms, Republicans flip their fiscal script	WASHINGTON (Reuters) - The head of a conservative Republican faction in the U.S. Congress, wh	politicsNews	31-Dec-17
2	U.S. military to accept transgender recruits on Monday: Pentagon	WASHINGTON (Reuters) - Transgender people will be allowed for the first time to enlist in the U.	politicsNews	29-Dec-17
3	Senior U.S. Republican senator: 'Let Mr. Mueller do his job'	WASHINGTON (Reuters) - The special counsel investigation of links between Russia and Presiden	politicsNews	31-Dec-17
4	FBI Russia probe helped by Australian diplomat tip-off: NYT	WASHINGTON (Reuters) - Trump campaign adviser George Papadopoulos told an Australian diplo	politicsNews	30-Dec-17
5	Trump wants Postal Service to charge 'much more' for Amazon shipments	SEATTLE/WASHINGTON (Reuters) - President Donald Trump called on the U.S. Postal Service on Fi	politicsNews	29-Dec-17
6	White House, Congress prepare for talks on spending, immigration	WEST PALM BEACH, Fla./WASHINGTON (Reuters) - The White House said on Friday it was set to ki	politicsNews	29-Dec-17
7	Trump says Russia probe will be fair, but timeline unclear: NYT	WEST PALM BEACH, Fla (Reuters) - President Donald Trump said on Thursday he believes he will b	politicsNews	29-Dec-17
8	Factbox: Trump on Twitter (Dec 29) - Approval rating, Amazon	The following statementsÂ were posted to the verified Twitter accounts of U.S. President Donalc	politicsNews	29-Dec-17
9	Trump on Twitter (Dec 28) - Global Warming	The following statementsÂ were posted to the verified Twitter accounts of U.S. President Donalc	politicsNews	29-Dec-17
10	Alabama official to certify Senator-elect Jones today despite challenge: CNN	WASHINGTON (Reuters) - Alabama Secretary of State John Merrill said he will certify Democratic	politicsNews	28-Dec-17

Fig. 2 Raw data of ISOT Fake News Dataset

pic_url	video_url	comment_num	repost_num	like_num
[]	http://f.video.weibocdn.com/000lGgVrlx07AJGBulao0104120(	4	2	30
[]	http://f.video.weibocdn.com/003exorvgx07ExKn4n0I0104120	0	0	0
[]		3	2	4
[]		0	0	0
[]		16	2	9
[]	http://f.video.weibocdn.com/xuZz09X7lx07E9SjHQfC0104120(	7	126	5070
[]	http://f.video.weibocdn.com/003WEnd9gx07AFfapAdi010412	17	0	36
['https://wx3.sinaimg.cn/orj360/53ae0b70ly1gbix8lcdkwj20ku112wer.jpg',		6	1	6
['https://wx3.sinaimg.cn/orj360/005Dmppzgy1gdkdfqullmj30u0140n2u.jpg'		66	11	123
[]		185	103	758
[]		16	3	3
[]		55	2	46
[]	http://f.video.weibocdn.com/Rurkudd7lx07EaKzEQBW010412	4	0	3
[]	http://f.video.weibocdn.com/uszghlQ0lx07BRzPWhSE0104120	145	10	384
[]	http://f.video.weibocdn.com/000LDQK8gx07E9rZXfz90104120	141	1487	10733

Fig. 3 Raw data of CHECKED Dataset

TABLE II

DATASET SUMMARY

Dataset	Real News	Fake News	Total
ISOT Fake News Dataset	21417	23481	44898
CHECKED Dataset	1760	344	2104

2) Data Pre-Processing: Data Pre-Processing is the phase where the raw data is converted into a computer-readable format using tokenization or vectorization. The classification can easily observe the features of the data. Microsoft Excel is used to export and clean the data before importing it to the WEKA tool. A few actions must be carried out, such as inserting null values into the missing part of data, removing repetitive stop words, lowering the case letter, and creating a bag of words by dissembling the phrases.

3) Feature Extraction and Feature Selection: Feature extraction and feature selection are conducted to optimize the model performance. In our study, content-based features will be used in ISOT Fake News Dataset, and user profile-based features will be used in CHECKED Dataset. For ISOT Fake News Dataset, we applied the same lexical based features from the work done by [2], [23] who proposed the features like average sentence length and word counts. We proposed new features that are top 20 features on the repetitive words, average length of a post and count of @.

Next, we applied the profile-based features utilized by Yang's [15] which are number of likes, number of time comments, existing of pic_url and existing of video_url for CHECKED Dataset. We used the existing features to detect the fake news. Table III shows all the features newly proposed in this work. Table IV and Table V show features proposed by previous work [7], [17].

TABLE III

CONTENT-BASED FEATURES FOR ISOT FAKE NEWS DATASET

Features	Descriptions	Data Type
Repetitive_word_1	Top 1 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_2	Top 2 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_3	Top 3 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_4	Top 4 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_5	Top 5 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_6	Top 6 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_7	Top 7 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_8	Top 8 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_9	Top 9 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_10	Top 10 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_11	Top 11 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_12	Top 12 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_13	Top 13 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_14	Top 14 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_15	Top 15 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_16	Top 16 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_17	Top 17 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_18	Top 18 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_19	Top 19 of the repetitive word counted in an article among all the dataset	Numerical
Repetitive_word_20	Top 20 of the repetitive word counted in an article among all the dataset	Numerical
Average length of a post	Average length of an article	Numerical
Count of @	Number count of "@" in an article	Numerical

TABLE IV

CONTENT-BASED FEATURE FOR ISOT FAKE NEWS DATASET BY [23]

Features	Description	Data Type
Word_count	Number of characters in an article	Numerical
Average_sentence_length	Average length of a sentence in an article	Numerical

TABLE V

PROFILE- BASED FEATURES FOR CHECKED DATASET BY [15]

Features	Description	Data Type
Number of post shared	The number of contents shared by the user	Numerical
Number of likes	The number of likes in a microblog post	Numerical
Number of times of comments	Number of times comments in a microblog post	Numerical
Existing of pic_url	Existence of picture URL in microblog or comment	Binary
Existing or video_url	Existence of video URL in microblog	Binary

4) Filtering Model: The filtering model in Fig. 4 is proposed where the first two tier machine learning algorithms classified the fake news in sequential way. If the fake news label misclassified by any of first two tier classification, then third classification known as consensus layer will be requested to classify the misclassified fake news and sent them to the corresponding category based on the label. If there is no conflict label between the first two tier classification, then the third-tier classification will be ignored, and the consensus layer will be directly sent the results to the corresponding category: True News or Fake News.

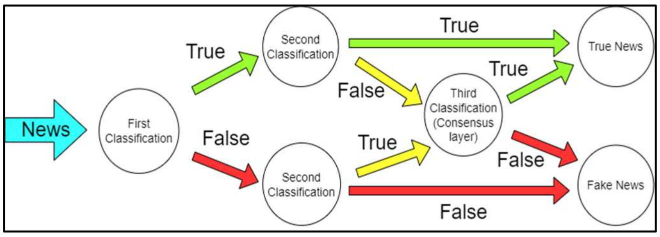

Fig. 4 Filtering model with consensus layer

Further comparison between consensus layer and first two-tier classification is not required. The existing classification algorithms made by other researchers or developers are used in our filtering model. Any classification algorithms are suitable to be applied in the filtering model. In this research paper, classification algorithms which are RF, NB, LR, AdaBoost and DT used in this fake news detection model. The pseudo code for the fake news detection model is illustrated in Fig. 5.

```
Pseudocode: Fake News Detection Model
Input: News with Label
Output: Predicted Label
Start
Split data into 10 sets using 10-fold cross
validation
Train all algorithms layer
For all news, n in dataset
    Test data with first algorithm layer
    Archive the result predicted by first
    algorithm layer, R1
    Test data with second algorithm layer
    Archive the result predicted by second
    algorithm layer, R2
    Test data with third algorithm layer
    Archive the result predicted by third
    algorithm layer, R3
    Compare R1 and R2
    If R1 = True, R2 = True Then
        n stores into True News;
    Compare with R3 if R1 and R2 collides
    ElseIf R1 = True, R3 = True Then
        n stores into True News;
    ElseIf R2 = True, R3 = True Then
        n stores into True News;
    Else
        n stores into Fake News;
    EndIf
EndFor
End
```

Fig. 5 Pseudo code for fake new detection model

5) N-Fold Cross Validation: N-fold cross validation is a train and test phase for the fake news detection model. The 10-fold cross validation is chosen to determine the best parameters and the most suitable arrangement for the algorithms used in the filtering model. First, the dataset will be shuffled randomly and split into approximately equal size of small parts dataset. For 10-fold cross validation, the dataset will be separated into 10 small parts of dataset and 9 of small dataset used as training dataset while 1 dataset used for testing. Then, the iteration of training and testing will be carried out 10 times and each iteration will be evaluated with scores. This validation is an unbiased metric hence it can avoid the algorithms overfitting or underfitting the data. This phase will be carried out by using WEKA tool where it can assist to choose the most suitable features.

6) Hyperparameter Tuning: Hyperparameter tuning is an optimization process to enhance the performance of algorithms. It can be done in manually tuning or automated tuning with the tool we used. In the multi-tier classification model, five algorithms have their own hyperparameter tuning method to implement. Since our model can run the algorithms independently, there are no consequences on other classifications performance while performing hyperparameter tuning in a classification. All the hyperparameter methods can be found and performed at each algorithm in the WEKA tool. Table VI shows the hyperparameter method used for optimizing five algorithms.

TABLE VI
HYPERPARAMETER TUNING METHOD FOR MACHINE LEARNING ALGORITHMS

Machine learning algorithm	Hyperparameter tuning method
RF	max_depth, num_features, bag_size, num_iterations
NB	Kernel Estimator
LR	L1 regularisation
AdaBoost	Classifier, num_iterations, weight_threshold
DT	min_instances_per_leaves, confidence_factor

B. Performance Metric

To evaluate the performance of the multi-tier fake news detection model, these performance metrics are considered:

1) F1-Score: The measurement metric of harmonic mean of precision and recall rate as in Equation 1. Precision indicates the rate of false positives predicted while recall indicates the rate of false negatives. High F1-Score shows the detection model performed well to predict the news correct with its label and low F1-Score shows the detection model predict the news with poor performance.

$$\mathrm{F1-Score} = \frac{2\times\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \quad (1)$$

2) Accuracy: The correct predictions about the fake news detection by the proposed model as shown in Equation 2. High accuracy value shows the detection model predicts well with the correct label. In contrast, low accuracy value shows the detection model struggles on prediction and high rate of incorrect labels are outcome.

$$\mathrm{Accuracy} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \quad (2)$$

3) Matthews Correlation Coefficient (MCC): MCC as shown in Equation 3 evaluates the model performance of four entities from 2x2 confusion matrix. The numerical value range is between -1 to 1 where 1 is the ideal binary classifier, where the detection model perfectly predicted according to the actuals; 0 indicates random binary classifier, where the detection model randomly predicted according to the actuals; and -1 indicates worst binary classifier, where the detection model not predicted according to the actuals at all.

$$MCC = \frac{TN \times TP - FN \times FP}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (3)$$

C. Experimental Setup

The experiment setup started by downloading both raw datasets from GitHub [14] and Kaggle [15] website. Next, the data cleaning phase follows the data pre-processing phase. In this phase, any invalid data is eliminated from the datasets. Blank data values are replaced with null values to ensure data consistency. Stop words and punctuation are removed to extract more valuable information and clean the data for subsequent data extraction phases. Additionally, all words are converted to lowercase. In the data pre-processing phase, the Natural Language Toolkit (NLTK) library is utilized for tokenizing the dataset. This process involves splitting the entire news content into phrases, which are then further divided into individual words.

D. Feature Extraction and Selection

In this section, each news content-based features for ISOT Fake News Dataset while news-content based and user profile-based features for CHECKED Dataset are explained. These features are used in the classification algorithms mentioned. We proposed 20 new features by calculating the repetitive word in ISOT Fake News Dataset. The formula for calculating the "Repetitive_word" can be represented as in Equation 4:

$$Repetitive\ word = \frac{Count(word)}{Total\ words} \quad (4)$$

where, count(word) represents the frequency of the most frequently occurring word in the news content from the dataset. While Total_words denote the total number of words in the news content. Table VII shows the top 20 features on the repetitive words and newly proposed features: average_length_of_a_post and count_of_@. Table VIII shows the features that have been proposed by [23]. For CHECKED dataset, we employed features that has been proposed by [15] as shown in Table IX.

TABLE VII
FEATURE MATRIX OF ISOT FAKE NEWS DATASET

Features	Description	Value
F_1	Repetitive_word_1_said	R1 = {0-28}
F_2	Repetitive_word_2_trump	R2 = {0-83}
F_3	Repetitive_word_3_would	R3 = {0-48}
F_4	Repetitive_word_4_president	R4 = {0-75}
F_5	Repetitive_word_5_people	R5 = {0-44}
F_6	Repetitive_word_6_one	R6 = {0-47}
F_7	Repetitive_word_7_state	R7 = {0-92}
F_8	Repetitive_word_8_also	R8 = {0-20}
F_9	Repetitive_word_9_new	R9 = {0-53}
F_{10}	Repetitive_word_10_reuters	R10 = {0-8}
F_{11}	Repetitive_word_11_donald	R11 = {0-74}
F_{12}	Repetitive_word_12_clinton	R12 = {0-68}
F_{13}	Repetitive_word_13_obama	R13 = {0-56}
F_{14}	Repetitive_word_14_house	R14 = {0-23}
F_{15}	Repetitive_word_15_government	R15 = {0-30}
F_{16}	Repetitive_word_16_states	R16 = {0-25}
F_{17}	Repetitive_word_17_republican	R17 = {0-25}
F_{18}	Repetitive_word_18_could	R18 = {0-12}
F_{19}	Repetitive_word_19_united	R19 = {0-42}
F_{20}	Repetitive_word_20_told	R20 = {0-13}
F_{21}	Average_length_of_a_post	R21 = {0-1}
F_{22}	Count_of_@	R22 = {0-88}

TABLE VIII
FEATURE MATRIX OF ISOT FAKE NEWS DATASET BY [23]

Features	Description	Value
F_{23}	Word_count	R23 = {1-5534}
F_{24}	Average_sentence_length	R24 = {0-252.71}

TABLE IX
FEATURE MATRIX OF CHECKED DATASET BY [15]

Features	Description	Value
F_1	Count_of_microblog_reposted	R1 = {0-1886915}
F_2	Count_of_microblog_commented	R2 = {0-73717}
F_3	Count_of_microblog_liked	R3 = {0-1179103}
F_4	Existing_of_picture_link	R4 = {0,1}
F_5	Existing_of_video_link	R5 = {0,1}

E. Constructing Feature Matrix

In this section, we construct the feature matrix of 24 features F_i ,i =1,.,24, i for ISOT dataset and the feature matrix of 5 features F_i ,i =1,….,5, i for CHECKED dataset. Features are in binary and decimal value. The Ri value for each feature is summarized in Table VII, Table VIII, and Table IX.

Let E = {e_1, e_2, ...,$e_{|E|}$} and F = {f_1, f_2, .., $f_{|6|}$ } denotes all the fake news and feature vector space respectively. So, | E | is a total fake news and | F | refer to size of feature vector. Let a_{ik} a be the value of *k*th feature of *i*th fake news. Therefore, the presentation of each fake news is $A_i = \{a_{i1}, a_{i2}, \ldots.\, a_{i[e]}\}$, and each fake news is $A = \{a_{ik}\}$ where i = 1,2,…,|F| and k =1,2,....,|E|. Where each fake news consists of A={Repetitive_word_1_said, Repetitive_word_2_trump, Repetitive_word_3_would, Repetitive_word_4_president, Repetitive_word_5_people, Repetitive_word_6_one, Repetitive_word_7_state, Repetitive_word_8_also, Repetitive_word_9_new, Repetitive_word_10_reuters, Repetitive_word_11_donald, Repetitive_word_12_clinton, Repetitive_word_13_obama, Repetitive_word_14_house, Repetitive_word_15_government, Repetitive_word_16_states, Repetitive_word_17_republican, Repetitive_word_18_could, Repetitive_word_19_united, Repetitive_word_20_told, Word_count, Average_sentence_length, Average length of a post, Count of @, Count_of_microblog_reposted, Count_of_microblog_commented, Count_of_microblog_liked, Existing_of_picture_link, Existing_of_video_link }. Then, all datasets are converted to arff format to run in the WEKA and tested using NB, RF, LR, AdaBoost, and DT algorithm which is already prepared inside the WEKA.

F. Rule of Fake News Detection Model

In this section, the fake news detection model rule is set and executed in our proposed model. The purpose of setting the rule is to filtrate any data that has the same predicted result between the first classification layer and second classification layer. This process will scope down the dataset where it only consists of collision result between the first and second classification layer. After that, the consensus layer is performed, and the performance result is analyzed. The consensus layer acts as final decision layer hence the previous two classification layers do not affect the decision predicts by consensus layer. Hence, the independent decision on each algorithm will reduce the bias value of the overall model performance. This process will go through each condition sequentially and output the corresponding result. The summary of the rule is shown in Table X.

TABLE X
RULE SUMMARY

1st classification layer	2nd classification layer	3rd classification layer	Consensus layer
True	True	True	True
True	True	False	True
True	False	True	True
False	True	True	True
True	False	False	False
False	True	False	False
False	False	True	False
False	False	False	False

Moreover, five classification algorithms can be arranged in various sequences, without repetition. Two sequences for the multi-tier filtering model are as follows:

- LR_DT_RF: In this sequence, Logistic Regression (LR) is used in the first classification layer, Decision Tree (DT) in the second classification layer, and Random Forest in filtering classification layer.
- LR_NB_AdaBoost: In this sequence, Logistic Regression (LR) is employed in the first classification layer, Naive Bayes (NB) in the second classification layer, and AdaBoost in the consensus classification layer.

These sequences represent different combinations of classification algorithms used in the multi-tier filtering model, contributing to its overall effectiveness in fake news detection.

G. F1-Score Result

The F1-Score results for the ISOT and CHECKED datasets, utilizing two different sequences of multi-tier classification and five classification algorithms (NB, LR, RF, AdaBoost, and DT), are presented in Fig. 6. For the ISOT Fake News Dataset, the LR_DT_RF model achieved F1-score of 0.9892, while the LR_NB_AdaBoost model scored 0.9350. For the CHECKED Dataset, the LR_DT_RF model scored 0.9913, and the LR_NB_AdaBoost algorithm scored a slightly lower but still commendable F1-Score of 0.9896.

Comparatively, the LR_DT_RF model outperformed the LR_NB_AdaBoost model in both datasets, achieving higher F1-Scores. The F1-Score represents an ideal balance between precision and recall, with a perfect score of 1 indicating flawless classifier performance. The high F1-Scores obtained by the LR_DT_RF model in both 49 datasets indicate its superior ability to effectively identify and classify fake news instances.

Overall, the results demonstrate the efficacy of the proposed multi-tier classification model, with the LR_DT_RF sequence showcasing particularly strong performance in distinguishing between fake and real news. These findings provide valuable insights for enhancing and optimizing the fake news detection model and contribute to the broader field of research on combating misinformation.

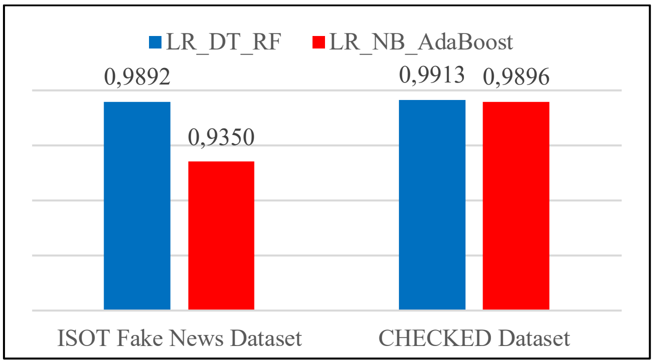

Fig. 6 F1-Score results for ISOT Fake News Dataset and CHECKED Dataset

H. Accuracy Result

The accuracy results for the ISOT and CHECKED datasets using the multi-tier fake news detection model, tested with two different sequences of multi-tier classification and five classification algorithms (NB, LR, RF, AdaBoost, and DT), are depicted in Fig. 7. For the ISOT Fake News Dataset, the LR_DT_RF model achieved a high accuracy scored of 0.9895, while the LR_NB_AdaBoost model scored 0.9329. In the CHECKED Dataset, the LR_DT_RF model obtained an accuracy scored 0.9853, and the LR_NB_AdaBoost model scored 0.9824.

The LR_DT_RF model in both datasets has outperformed the LR_NB_AdaBoost model. These algorithms exhibited higher accuracy scores, indicating a greater number of correct predictions made by the models. Overall, the accuracy results claimed the research objectives, and all the sequences in both datasets exceed 0.98 accuracy except for LR_NB_AdaBoost in ISOT Fake News Dataset. The high accuracy scores achieved by the LR_DT_RF model and other sequences indicate the effectiveness of the proposed multi-tier fake news detection model.

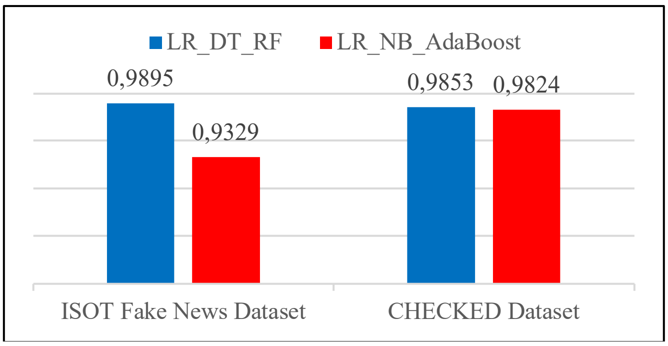

Fig. 7 Accuracy results for ISOT Fake News Dataset and CHECKED Dataset

I. Matthews Correlation Coefficient (MCC) Result

Fig. 8 shows the MCC results for ISOT and CHECKED datasets for multi-tier fake new detection model with different sequence of classification algorithm. For the ISOT Fake News Dataset, the LR_DT_RF model achieved a high MCC score of 0.9790, while the LR_NB_AdaBoost model scored 0.8736. In the CHECKED Dataset, the LR_DT_RF model obtained an MCC score of 0.9455, and the LR_NB_AdaBoost model scored 0.9347. The higher MCC scores achieved by the LR_DT_RF model in both datasets reinforce its efficacy in fake news detection and its potential for reliable and robust classification.

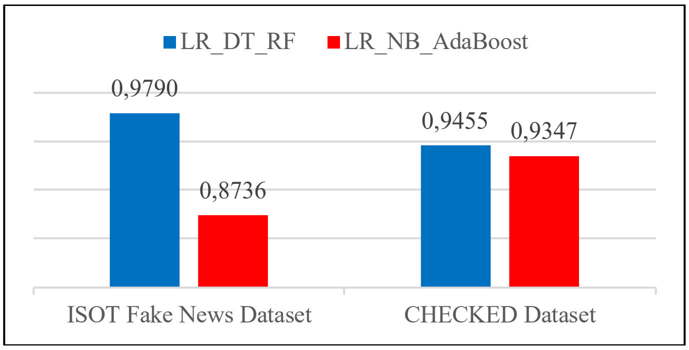

Fig. 8 MCC results for ISOT Fake News Dataset and CHECKED Dataset

IV. Conclusion

Our work proposed fake news detection model with consensus layer on news content-based features to test on ISOT Fake News Dataset and user profile-based features to test on CHECKED Dataset. We defined 24 features on news content-based for ISOT Fake News Dataset and five features on user profile-based for CHECKED Dataset. Then, all five classification layers are trained and tested with 10-fold cross validation with the features mentioned. The result is filtered based on the rule set to extract collision result that will be used in the consensus layer. Next, the evaluation process is performed on a consensus layer with the performance metric of F1-Score, Accuracy, and MCC. The proposed fake news detection model achieved the highest result on the LR_DT_RF sequence of model in overall performance in ISOT Fake News Dataset and CHECKED Dataset where the F1-Score resulted 0.9892 and 0.9913, the accuracy scored 0.9895 and 0.9853, and last the MCC obtained 0.9790 and 0.9455. In the future works, several areas can be explored to further enhance the fake news detection approach. we would investigate and evaluate additional features relevant to the dataset to enhance the model's ability to identify fake news. Exploring various textual, linguistic, and context-based features could provide valuable insights for improving classification accuracy. We would consider employing different classification algorithms such as Supper Vector Machine (SVM), K-Nearest Neighbour (KNN), and XGBoost for the classification model. The proposed model does not restrict the use of any classification model created by other researchers. This exploration can offer a deeper understanding of the best suited algorithms for fake news detection. By addressing these future research directions, this study can serve as a valuable reference and provide a comprehensive approach for researchers working on fake news detection across diverse platforms and using various feature-based models. The proposed improvements can enhance the model's accuracy, efficiency, and adaptability in tackling the pressing challenge of combating fake news. Other than fake news topic, the proposed model can also be implemented in different domains such as malware analysis, phishing analysis, and image classification. The proposed model is only executed in data analytic phase. Therefore, the extracted feature of different domains still can be used on the proposed model to outcome the results.

Acknowledgment

This research was supported by Universiti Tun Hussein Onn Malaysia (UTHM) through TIER 1 (Vot. Q167).

References

[1] X. Zhou and R. Zafarani, “A survey of fake news: Fundamental theories, detection methods, and opportunities,” *ACM Computing Surveys (CSUR)*, vol. 53, no. 5, pp. 1–40, 2020.

[2] X. Zhang and A. A. Ghorbani, “An overview of online fake news: Characterization, detection, and discussion,” Information Processing & Management, vol. 57, no. 2, p. 102025, Mar. 2020, doi:10.1016/j.ipm.2019.03.004.

[3] N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer, “Fake news on Twitter during the 2016 U.S. presidential election,” Science, vol. 363, no. 6425, pp. 374–378, Jan. 2019, doi:10.1126/science.aau2706.

[4] S. Vosoughi, D. Roy, and S. Aral, “The spread of true and false news online,” Science, vol. 359, no. 6380, pp. 1146–1151, Mar. 2018, doi:10.1126/science.aap9559.

[5] V. Balakrishnan, K. S. Ng, and H. A. Rahim, “To share or not to share – The underlying motives of sharing fake news amidst the COVID-19 pandemic in Malaysia,” Technology in Society, vol. 66, p. 101676, Aug. 2021, doi: 10.1016/j.techsoc.2021.101676.

[6] D. M. J. Lazer et al., “The science of fake news,” Science, vol. 359, no. 6380, pp. 1094–1096, Mar. 2018, doi: 10.1126/science.aao2998.

[7] S. Hangloo and B. Arora, “Fake News Detection Tools and Methods– A Review,” *arXiv preprint arXiv:2112.11185*, 2021.

[8] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, “FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media,” Big Data, vol. 8, no. 3, pp. 171–188, Jun. 2020, doi:10.1089/big.2020.0062.

[9] A. S. U and F. M. Philip, “A Hybrid Method for Fake Profile Detection in Social Network Using Artificial Intelligence,” 2021.

[10] K. Shu, S. Wang, and H. Liu, “Beyond News Contents,” Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, Jan. 2019, doi: 10.1145/3289600.3290994.

[11] Z. Khanam, B. N. Alwasel, H. Sirafi, and M. Rashid, “Fake News Detection Using Machine Learning Approaches,” IOP Conference Series: Materials Science and Engineering, vol. 1099, no. 1, p. 012040, Mar. 2021, doi: 10.1088/1757-899x/1099/1/012040.

[12] S. B. Parikh and P. K. Atrey, “Media-Rich Fake News Detection: A Survey,” 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Apr. 2018, doi:10.1109/mipr.2018.00093.

[13] R. K. Kaliyar, A. Goswami, and P. Narang, “FakeBERT: Fake news detection in social media with a BERT-based deep learning approach,” Multimedia Tools and Applications, vol. 80, no. 8, pp. 11765–11788, Jan. 2021, doi: 10.1007/s11042-020-10183-2.

[14] H. Ahmed, I. Traore, and S. Saad, “ISOT Fake News Dataset,” pp. 1–2, 2017.

[15] C. Yang, X. Zhou, and R. Zafarani, “CHECKED: Chinese COVID-19 fake news dataset,” Social Network Analysis and Mining, vol. 11, no. 1, Jun. 2021, doi: 10.1007/s13278-021-00766-8.

[16] O. Stitini, S. Kaloun, and O. Bencharef, “Towards the Detection of Fake News on Social Networks Contributing to the Improvement of Trust and Transparency in Recommendation Systems: Trends and Challenges,” Information, vol. 13, no. 3, p. 128, Mar. 2022, doi:10.3390/info13030128.

[17] M. Del Vicario, W. Quattrociocchi, A. Scala, and F. Zollo, “Polarization and Fake News: Early Warning of Potential

Misinformation Targets," Feb. 2018, [Online]. Available: http://arxiv.org/abs/1802.01400.
[18] M. L. Della Vedova, E. Tacchini, S. Moretécole, G. Ballarin, M. Dipierro, and L. De Alfaro, "Automatic Online Fake News Detection Combining Content and Social Signals," 2018. [Online]. Available: https://developers.facebook.com/docs/graph-api.
[19] H. Rashkin, E. Choi, J. Y. Jang, S. Volkova, and Y. Choi, "Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking," Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017, doi:10.18653/v1/d17-1317.
[20] R. Megan, "Getting Real about Fake News." Accessed: Jun. 18, 2023. [Online]. Available: : https://www.kaggle.com/mrisdal/fake-news/data.
[21] H. Ahmed, I. Traore, and S. Saad, "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques," Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, pp. 127–138, 2017, doi: 10.1007/978-3-319-69155-8_9.
[22] J. Shaikh and R. Patil, "Fake News Detection using Machine Learning," 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Dec. 2020, doi:10.1109/isssc50941.2020.9358890.
[23] S. Hakak, M. Alazab, S. Khan, T. R. Gadekallu, P. K. R. Maddikunta, and W. Z. Khan, "An ensemble machine learning approach through effective feature extraction to classify fake news," Future Generation Computer Systems, vol. 117, pp. 47–58, Apr. 2021, doi: 10.1016/j.future.2020.11.022.
[24] B. T.K., C. S. R. Annavarapu, and A. Bablani, "Machine learning algorithms for social media analysis: A survey," Computer Science Review, vol. 40, p. 100395, May 2021, doi:10.1016/j.cosrev.2021.100395.
[25] M. H. Goldani, S. Momtazi, and R. Safabakhsh, "Detecting fake news with capsule neural networks," Applied Soft Computing, vol. 101, p. 106991, Mar. 2021, doi: 10.1016/j.asoc.2020.106991.
[26] T. Jiang, J. P. Li, A. U. Haq, A. Saboor, and A. Ali, "A Novel Stacking Approach for Accurate Detection of Fake News," IEEE Access, vol. 9, pp. 22626–22639, 2021, doi: 10.1109/access.2021.3056079.
[27] C. Tang, K. Ma, B. Cui, K. Ji, and A. Abraham, "Long text feature extraction network with data augmentation," Applied Intelligence, vol. 52, no. 15, pp. 17652–17667, Apr. 2022, doi: 10.1007/s10489-022-03185-0.
[28] L. Hu, S. Wei, Z. Zhao, and B. Wu, "Deep learning for fake news detection: A comprehensive survey," AI Open, vol. 3, pp. 133–155, 2022, doi: 10.1016/j.aiopen.2022.09.001.
[29] L. Zhou, J. Tao, and D. Zhang, "Does Fake News in Different Languages Tell the Same Story? An Analysis of Multi-level Thematic and Emotional Characteristics of News about COVID-19," Information Systems Frontiers, Sep. 2022, doi: 10.1007/s10796-022-10329-7.
[30] Y. Yang, S. Nazir, and W. Khalil, "A probabilistic approach toward evaluation of Internet rumor on COVID," *Soft computing*, vol. 26, no. 16, pp. 8077–8088, 2022.