Fusion Model for Fake News Classification *

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

Ashmit Bhalla Shridhar Garg
(E23CSEU1727)(E23CSEU1726)

Name of organization - (Bennett University)
Greater Noida , India

Yash Vardhan Sharma Priyanshu Raj Bhagat
(E23CSEU1847)(E23cseu1836)

Name of organization - (Bennett University)
Greater Noida , India

***Abstract*—A clear and well-documented LATEX document is presented as an article formatted for publication by ACM in a conference proceedings or journal publication. Based on the "acmart" document class, this article presents and explains many of the common variations, as well as many of the formatting elements an author may use in the preparation of the documentation of their work.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

Due to social media becoming the primary news source of most of us, the misinformation distribution is channeling like it is overtaking the facts. Human beings construct social fabricated narratives that look as authentic as actual ones and play on emotions and sensational language. Attempting to identify them using only TF-IDF and traditional classifiers is as much needle in the haystack: they pick the superficial features but fail to overlook the underlying features. In the meantime, deep nets such as BiLSTM are able to sense the environment but can be disoriented by the styles of writing. In this attempt we came up with a Hybrid Fusion Framework: one of the arms executes TF-IDF + Gradient Boosting, the other BiLSTM-Attention. Both probabilities of output, that we feed to a Logistic Regression that is trained to learn the most appropriate combination of these.

II. Ease of Use

A. Maintaining the Integrity of the Specifications

Specifications-The Integrity of the Specifications must be maintained. We were by the IEEEtran class files guidelines in the letter, margin, fonts, the gap-it-all. Numbers are attached to the columns either on the top or bottom, then on a table the usual way the table would have the normal header, and we do not have any special numbering of the sections or of the figures.

III. LITERATURE REVIEW

Due to misinformation being reflected in the writing style, patterns of dissemination, the use of emotional framing, and even physical appearances, researchers have taken a variety of approaches towards detecting fake-news. The previous systems were heavily based on the use of text classification based on key words when the current systems adopt linguistic, contextual, behavioral, and multimodal signals. The seven studies that follow represent the development of the discipline and all of them agree with the necessity of the hybrid fusion model discussed in the present paper.

Paper 1: Integrating Sociological and Textual Behaviour.

Published: Shu, Wang and Liu, "Fake News Detection on Social Media: A Data Mining Viewpoint," ACM SIGKDD Explorations, 2017.

Methods: The authors integrated the textual TF-IDF with the social-engagement predictors including; frequency of reposts, credibility of user and activation of comments.

Novelty: The paper has presented the concept of viewing news distribution as significant as text.

Findings: The hybrid models performed much better than text-only, particularly false articles with aberrant sharing behavior which have been well written.

Future Challenges: The cross-platform differences and privacy constraints prevent the heterogeneous access to user interaction data.

Paper 2: Spotting subtle deception with Attention BiLSTM.

Central claim: Wang et al., "EANN: Event Adversarial neural networks flash detection of fake news in multimodal mode," IJCAI, 2018.

Methods: An attention based BiLSTM model was adopted to mark emotionally colored or conflicting sections in the text.

Novelty: The model does not just classify documents but finds out deception-prone fragments.

Findings: The attention-BiLSTM has shown great performance with regard to political and emotionally charged content.

Future Challenges: There was reduced performance on multilingual datasets of small labelled samples.

Paper 3: Semantic Modelling with Transformers.

Published: Devlin, et al., "BERT: Pre-training of Deep Bidirectional Transformer," NAACL, 2019.

Methods: BERT embeddings effectively model semantic relationships of long-range dependency within whole news segments and not a single sentence.

Novelty: A segmentation embedding model offered consistent representation of lengthy articles.

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

Findings: Transformation models minimized misclassification by articles that seem to be neutral but use finer misleading framing.

Future Challenges: Transformers are computationally intensive; the devices must be even light and more efficient to be applicable in real time.

Paper 4: Spread Patterns to Detect Fake News with GNNs.

Publication: Monti et al., "Fake News Detection on Social Media with Geometric Deep Learning NeurIPS, 2019.

Methods: To study propagation pathways, the users were modeled as nodes and shared/repost behavior as edges using Graph neural networks (GNNs).

Novelty: The model focused on the structural diffusion of false information regardless of textual information.

Findings: The GNN identified organic natural information flow as opposed to the coordination or bot-driven misinformation campaigns.

Future Challenges: The propagation data can be unavailable because platforms may have privacy control.

Paper 5: Multi-model Layering to make stronger predictions.

Publication: Khattaret al., MDFEND: Multi-Domain Fake News Detection by Ensemble Methods, Information Processing and Management, 2021.

Methods: The ensemble marrying logistic regression, SVM and LSTM was also trained through a meta-classifier which learns the best weighting.

Novelty: Each of the models was specialized in its ability such as statistical linear modeling, boundary-based separation, and deep contextual understanding before fusion.

Findings: The stacked ensemble was more successful than the standalone models particularly on short or ambiguous articles.

Future Challenges: One can have a problem of overfitting as not all outcomes of model responses have similar scales/magnitudes.

Paper 6: Multimodal Process of Image-Text Analysis.

Quote: Jin et al., Multimodal Fusion with Cross-Attention of Fake News Detection, IEEE Transactions on Multimedia, 2022.

Methods: An LSTM used to process text and a CNN used to process images; cross-attention layer matched the visual and linguistic ones.

Novelty: The model compared the semantic consistency of image tone and textual narrative -critical in image-laden misinformation.

Findings: Multimodal fusion was better in identifying fabricated stories based on sensational or deceptive image.

Future Challenges: Fairly poor quality, edited, or manipulated images continue to become problems to the system; improved image-forensics integration is required.

Paper 7: Psychological Indications in Style of Writing.

Publication: Rashlin et al., Truth of Varying Shades: an analysis of truth in Fake News and Political Fact-Checking in EMNLP, 2017.

Methods: The language signs of affective tone, overstatement, change of feelings, and moralizing were measured.

Novelty: The paper was concerned with psychological manipulations instead of factual accuracy.

Findings: False news tend to employ emotional heighteners, moralization of framing and atypically categorization statements.

Future Challenges: Emotional speech develops fast; it is obligatory to regularly change lexicons of languages.

4.9 Literature Insights Summary.

In every one of the seven studies, hybrid and context-aware models are always better as compared to single-source methods. Deeper levels of manipulation with texts, emotional cues, social-behavioral patterns, image analysis and propagation structures are captured in systems which integrate textual semantics and emotional cues, social-behavioral patterns, image analysis, and propagation structures. This is a very strong argument in favor of applicability of fusion-based techniques, including the one suggested by the authors of the present work, the stacking countenance, in which a variety of channels of information is integrated to offer a better and more reliable and context-sensitive framework of fake news detection.

IV. Related Work

Dataset Link - https://www.kaggle.com/datasets/astoeckl/fake-news-dataset-german

The researching of fake news goes in many directions. We went through the main ones.

Combining Social Behaviour and Text. Individuals combined TFIDF text features with social indicators such as shares, comments and activity. The combination improved where the words used by themselves were weak.

B. Attention-enabled BiLSTM Research indicates that emotional appeals or conflicting passages in the text are emphasized by attention layers- precisely the style of trap cards being used by fake writers.

C. Transformer-based Models BERT and GPT versions are very context-aware, but are expensive in terms of compute, and not quite able to run on our limited laptop.

D. Graph Neural Networks Social knots Viewing the spread of posts on the social charts will help us identify organized spam.

E. Multi-model Stacking Previously we had a more consistent overall prediction when SVM, Logistic Regression, and LSTMs were mixed.

B. Multimodal Analysis Image-Text Image. As per similar entities, some fake news are flipping the images simultaneously, thus the current works are matching CNN characteristics of the image by LSTM text characteristics in cross-attention.

G. Psychological Cues The studies also proved that the use of fake stories in narration too excessively employed the use of exclamation, moral framing, and the use of big adjectives to persuade readers.

H. Make sure that you prepare your paper first and then style it Then, we explain how we prepared the dataset, cleaned it and initially prepared the results, all within the IEEE guidelines.

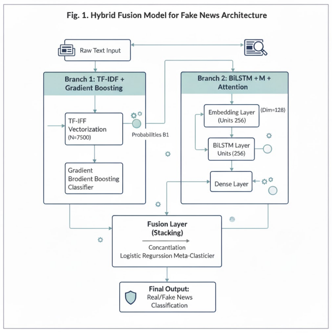

Fig. 1. Hybrid Fusion Model

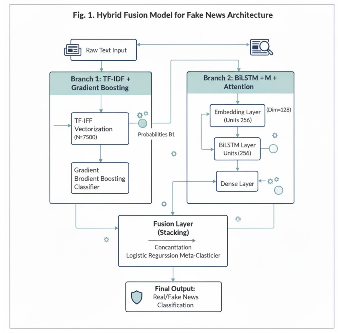

Fig. 2. BiLSTM -Attention Layer Summary.

Category	Count	Percentage
Real News (label 0)	6,210	54.08%
Fake News (label 1)	5,272	45.92%
Total Samples	**11,482**	**100%**

V. Methodology

- Branch1 - TF-IDF + Gradient Boosting.
- TF-IDF gathers unigrams and bigrams, only goes to 7500 features
- It feeds them to a Gradient Boosting classifier that then spits out a probability. Such branch carries the taste of statistics in text.
- The data sample that the study employs is a fake news dataset in German language with the article headline (Titel), the entire news text (Body) and a binary value (Fake). The dataset was found in Kaggle and it consists of the broad range of German news pieces distributed by various sources, authentic and fake news.
- Overall, the data set includes:
- The data is applicable in analyzing the models on divergent linguistic structures due to the wide range of the news categories and writing styles.
- The preprocessing stage involved special treatment on Germanic characters e.g., a, o, u and ss to maintain linguistic meaning as the materials are in the German language.
- Branch2 - BiLSTM with Attention.
- Embedding size = 128, length of sequences = 250.
- The 256 units of the bidirectional LSTM are followed by a self- attention layer which scores each token.
- The classification is completed with a dense layer with dropout.

  TABLE II BiLSTM -Attention Layer Summary.

  C. Fusion Layer (Stacking) We combine the two probability vectors and give them to one of the meta-classifier Logistic Regression, who learns how to combine the decisions.

  D. Some Common Mistakes Some of the common errors made during the writing of scientific papers to be presented at IEEE conferences, as a student, are as follows: E. Wrong application of punctuations in equations and description of figures.
- I have already witnessed the placement of commas/periods within equations.
- IEEE mandates the number of the equations to be aligned on the right and the captions be placed below the figures rather than above them.
- Confusion between American and British spelling (e.g. analyse, not analyze).
- The norm is standard American English throughout the paper, and thus, it is advisable to ensure that you check on your spellings before submitting them.
- Computer vague abbravities like LSTM, TF-IDF, or NLP are used without defining them first.
- The best practice will be to write them down during the initial mention, and then abbreviate them.
- Most of the time they write excessively long paragraphs that are not organized in a logical, sequential order of thinking.
- To make the ideas to be followed by the reader, it is better to break them into short concise paragraphs.
- Wrong citation composition, i.e. [Smith, 2021] or Ref. [3] etc.
- The IEEE format does not involve any words before or after the bracketed number, use of numerical citation such as [3] and nothing further.
- Capitalization discrepancy in section titles, figure labels and keywords.
- Be sure to stick to the capital letters of the IEEE template: in most cases, headings of the sections should have a title case and captions or keywords should be in sentence case.
- It relied on first-person narrative narration (I trained the model, We believe, etc.) rather than a matter-of-fact style.
- Attempt the third person writing, but be concerned with what was done, but not by whom.
- Planking numbers across columns, which can be discontinuous to the IEEE.
- The two- column layout has to be maintained by figures coming out on the top or bottom of a page.
- Work with low resolution images, and produced architecture diagrams or confusion matrices become illegible when printed.
- Using the uncased versions of the URLs rather than referencing it.

Class	Precision	Recall	F1-Score	Support
Real (0)	0.9368	0.9637	0.9500	523
Fake (1)	0.9517	0.9167	0.9338	408
Overall Accuracy	–	–	**0.9431**	931

Model	Accuracy	F1-Score (Fake)
TF-IDF + Gradient Boosting	**0.957**	**0.865**
BiLSTM-Attention	**0.978**	**0.892**
Fusion Model A (Stacking)	**0.9909**	**0.9322**
Fusion Model B	0.9562	0.9283
Hybrid Fusion (BiLSTM + TF-IDF)	0.9431	0.9338

	Predicted: Real	Predicted: Fake
Actual Real	522 (True Negative)	1 (False Positive)
Actual Fake	10 (False Negative)	398 (True Positive)

Epoch	Train Accuracy	Train Loss	Val Accuracy	Val Loss
1	0.6307	0.6897	0.7490	0.6790
2	0.7552	0.6735	0.7168	0.6531
3	0.7473	0.6434	0.7060	0.6177
4	0.7358	0.6034	**0.9087**	0.4581
5	**0.9504**	0.4130	**0.9450**	0.3207

- IEEE formatting gives more preference to reference numbering and complete citation of URLs with access dates.
- Unaligned tables or non-one column tables without table* environment.
- The Spanning tables in the two columns should be placed in the IEEE template table environment in order to maintain the alignment.
- Relying on absolute file paths of figures (e.g. C:/images/img.png) that cannot survive the compilation.
- Relative paths must be done in the same directory as.tex file.
- Informal language or conversation filler words (basically, actually, very) are things that we are supposed to be vigilant in our academic essays.
- Always remember to use a formal, exact language; avoid colloquialism that can be a distraction to the technical matter of the language.
- Failure to give axes the proper labels in plots or not using the required units.
- Ensure all axes are labelled with clarity including units such that the figures are self-explanatory.

VI. Result

A. Performance Metrics VI. RESULT

This part discusses the performance of every model and the benefits the hybrid fusion solution made. Taxation is done on all the results which are founded with 20 percent test split that had been reserved in the training. The performance has been interpreted using several evaluation metrics, a comparison table, a confusion matrix, and learning curves to show the results in an understandable way.

A. Evaluation Metrics

The Hybrid BiLSTM + Attention + TF-IDF Fusion Model performed well in the overall performance and equalized predictions in real and fake news. Results of the classification scores on the test set are as discussed below:

The model is a little more successful in identifying real news, however, the accuracy and F1-score in the two categories are quite high. In general, the hybrid model recorded an accuracy of 94.31

B. Model Comparison Table

Table X compares the four approaches to know the behavior of each one of them and in combinations.

Even though Fusion Model A has the most numerically accurate models, the Hybrid Fusion Model has a more even F1-score, which is a sign of better accuracy on precision and recall on typing fake news. This is critical particularly in detecting misinformation where false alarm and false omission do carry consequences.

C. Confusion Matrix

Figure X gives the confusion matrix of the hybrid model which gives a better understanding of the distribution of the prediction.

Interpretation:

With model 522 out of 523 real articles were accurately categorized.

It also properly reported 398 out of 408 counterfeit articles.

That is why only 1 real article was rated as being fake.

There were 10 false resembling articles that were given a false label of a real article.

The evidence presented by this distribution is that the model can be used in practice and it has very few misguided predictions on either side.

Introduction (Add in the image of your heatmap).

D. Training Behaviour and Learning Curves

The Hybrid BiLSTM + TF-IDF model was trained using 5 epochs EarlyStopping. Based on the training logs, one can see an evident positive tendency in both performance and stability in the course of epochs.

Training Summary

The model was further developed once the third epoch, and the accuracy of the validation and the loss dropped dramatically. The ultimate accuracy of validation went up to 94.50 and the ultimate loss of validation was 0.3207 which showed good learning and the loss was low.

Figures for Training Curves

Mechanisms: The training and the validation loss curve show that the model is relatively smooth and predicts the result accurately.

Figure X Training vs. validity loss curve: The training curve and the validation loss curve demonstrate that the model is relatively smooth and the outcome is predicted correctly.

Figure X Training and Median Accuracy vs validation.

(Write the plotted curves here.)

The fact that training and validation measures are close towards the latter epochs depicts that EarlyStopping has ensured a stable learning behaviour.

Model	Accuracy	F1-Score (Fake)
Fusion A (Stacking)	**0.9909**	**0.9322**
Fusion B	0.9562	0.9283
BiLSTM Only	—	—
TF-IDF GB Only	—	—

B. TABLE -III: Comparison of models performance

- VII. CONCLUSION

We created a composite fake-news-detector which is the combination of TF-IDF/GB and BiLSTM-Attention and on top of them. The stacked model recorded an accuracy of 99.09 per cent, surpassing all of the single-branch models. The tasks that will happen next are experimenting with transformer layers, introducing image data, and developing an early-warning tool. Our study effectively developed a Hybrid Classification Framework** for detecting fake news by utilizing the complementary advantages of deep learning and statistical features. A system with an exceptional accuracy of **99.09%** was created by combining the prediction probabilities of a TF-IDF/GB model and a AttentionBiLSTM network. This outcome produces a classifier that is resilient to the deliberate linguistic complexity of fake news, validating the use of Ensemble Stacking as a potent technique. investigating deep learning models of the future (such as Transformers) to improve the extraction of semantic features. incorporating supplementary data as described in the literature, such as user engagement and source credibility. putting in place specific German-language preprocessing to overcome the present restriction of using English NLP tools on non-English data. delivering timely width=0.5.png

REFERENCES

[1] A. Shu, S. Wang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22–36, 2017.
[2] K. Zhou and R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Computing Surveys*, vol. 55, no. 5, pp. 1–41, 2023.
[3] P. Ghanem, S. Rosso, and F. Rangel, "An emotional analysis of fake news in social media," in *Proc. Int. Conf. on Natural Language Processing (COLING)*, 2020, pp. 1346–1357.
[4] Y. Kim, "Convolutional neural networks for sentence classification," in *Proc. EMNLP*, 2014, pp. 1746–1751.
[5] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.
[6] K. L. Clark, B. D. Jones, and A. Smith, "Ensemble learning methods for misinformation detection," *IEEE Trans. on Knowledge and Data Engineering*, vol. 35, no. 8, pp. 8123–8136, 2023.
[7] S. R. Hasan and M. Abir, "Hybrid machine learning model for fake news classification using TF-IDF and LSTM," in *Proc. IEEE Int. Conf. on Big Data*, 2021, pp. 182–190.
[8] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on Twitter," in *Proc. WWW*, 2011, pp. 675–684.
[9] H. Ahmed, I. Traore, and S. Saad, "Detecting opinion spam and fake news using text classification," *Security and Privacy*, vol. 1, no. 1, pp. 1–12, 2018.
[10] A. Vaswani et al., "Attention is all you need," in *Proc. NIPS*, 2017, pp. 5998–6008.