FakEDAMR: Fake News Detection using Abstract Meaning Representation

Shubham Gupta, Narendra Yadav, Sainathreddy Sankepally, Suman Kundu, *Member, IEEE*

***Abstract*—Given the rising prevalence of disinformation and fake news online, the detection of fake news in social media posts has become an essential task in the field of natural language processing (NLP). In this paper, we propose a fake detection model named, FakEDAMR that encodes textual content using the Abstract Meaning Representation (AMR) graph, a semantic representation of natural language that captures the underlying meaning of a sentence. The graphical representation of textual content holds longer relation dependency in very few distances. A new fake news dataset, FauxNSA, has been created using tweets from the Twitter platform related to 'Nupur Sharma' and 'Agniveer' political controversy. We represent each sentence of the tweet using AMR graph and then use this in combination with textual features to classify fake news. Experimental results on two different sets of features show that adding AMR graph features improves F1-score and accuracy. In the experiments, Random Forest with AMR-encoded features outperforms other models in Feature-set 1, achieving 88.90%, 89.48%, 87.09% accuracy and 85.92%, 88.69%, 86.70% F1-score on the FauxNSA, Covid19-FND, and KFN datasets, respectively. However, when Feature-set 2 is used, BiLSTM with AMR-encoded features emerges as the top-performing model. It achieves highest accuracy and F1-score of 93.96% and 91.96% on the dataset. It also maintains high performance on Covid19-FND and KFN datasets, with accuracy and F1-scores of 93.26% and 93.20% on Covid19-FND, and 93.52% and 93.52% on KFN, respectively.**

***Index Terms*—Fake News Detection, Natural Language Processing, AMR Graph, India Nupur Sharma protest, India Agniveer protest.**

I. Introduction

Social media became essential for communication and information sharing. However, news shared over social media platforms lacks cross-referencing, allowing the spread of misinformation. Interestingly, it appears that the rate at which fake news is shared on Twitter exceeds that of genuine news [19]. Figure 1 presents some examples of fake news which spread through various media platforms, including Twitter. Many ML/DL methods were proposed to identify fake news from social media [8]. These existing methods focused on syntactic features and did not investigate how semantic features of news content affect ML models. However, complex semantic features are seen to improve the performance of different NLP tasks such as event detection [7], abstractive summarization [15], and question answering [14] in machine learning. Considering this one may ask "Does incorporating complex semantic features of sentences enhance the performance of fake news detection models too?"

Fig. 1. Examples of false information related to the 'Nupur Sharma' and 'Agniveer' controversy showcased in the images (Courtesy: Boomlive). The images depict various misleading claims, including: a) Russia, Netherlands, France and 34 other countries are supporting India and Nupur Sharma. b) Nupur Sharma is arrested and in the jail. c) Oppressors are damaging the railway line in the protest of Aginveer scheme.

The present study proposed a fake news detection model, FakEDAMR, that classify tweets as genuine and fake information, by introducing graph-based semantic features with syntactic and lexical features of the sentences. The main contribution of our work is to use deep semantic representation from the features of the Abstract Meaning Representation (AMR) graph. AMR helps to better extract the relationships between entities far apart in the text with minimum cost. This approach reduces the emphasis on syntactic features and collapses certain elements of word category, such as verbs and nouns, as well as word order and morphological variations. To the best of our knowledge, this is the first study to examine the semantics features of AMR graphs for detecting fake news. We curated a fake news dataset, namely, FauxNSA, related to the well-known controversies *Nupur Sharma* and *Agniveer* in India. Tweets with a list of curated hastags (Table I) on said topics are collected from the Twitter platform, in two different languages - Hindi and English. We extracted AMR graphs from each text document by using STOG model [25]. We encoded AMR graphs using graph embedding and combined them with the syntactic features of the text used in state-of-the-art model [23]. Finally, the resulting embedding vector, which included both semantic and syntactic features, is fed into a deep-learning model to predict the probability of fake and real. We have experimented our model on two publicly available datasets (Covid19-FND[20], KFN[13]) and FauxNSA. Our experiments demonstrated an improvement in accuracy of 2-3% over all the datasets when the AMR graph features were included with exiting textual features in the model. In particular, we sum up contributions of this study as follows:

- A new fake news dataset, FauxNSA has been created, focusing on a well-known controversy in India involving

S. Gupta, N. Yadav and S. Kundu are with Department of Computer Science and Engineering, Indian Institute of Technology Jodhpur, India-342030. S. Sankepally is with International Institute of Information Technology Naya-Raipur, India-492109.
E-mail: gupta.37@iitj.ac.in, yadav.42@iitj.ac.in, sankepally-sainathreddy@gmail.com, suman@iitj.ac.in

individuals named ‘Nupur Sharma’ and ‘Agniveer’. This dataset comprises tweets pertaining to religion, politics, and terrorism. To ensure the dataset’s credibility, a meticulous methodology was employed to gather both fake and real information from reliable sources.

- In this study, a new model called FakEDAMR is introduced for detecting fake news. The model incorporates Abstract Meaning Representation (AMR) to capture intricate semantic details, resulting in a significant improvement of 2-3% in accuracy on two publicly available datasets (Covid19-FND [20] and KFN [13]), as well as FauxNSA dataset. This novel approach effectively enhances the ability to identify and distinguish fake news articles.
- Further, this study aims to determine the relative importance of AMR graph features compared to text-based features within a model and evaluate the model performance in handling unknown samples.

The rest of the paper is organized as follows: Section II reports the related work. Section III and IV describe the working methodology and experimental setup. Section V reports the results with comparative analysis. Ablation study is presented in Section VI. Finally, Section VII concludes the research outcome.

II. Related Work

Fake news detection has been extensively studied recently using Natural Language Processing. Oshikawa et al. [17] clarify the distinction between detecting fake news and related concepts, including rumor detection, and provide an overview of current data sets, features, and models. As mentioned in the introduction, Castillo et al. [3] created a set of 68 features in the identification of false information. They used propagation tree over the feature set to identify whether the news is false or not. An extension to the lexical-based analysis model is used in [16] by incorporating speaker profile details into an attention-based long short-term memory (LSTM) model. Zervopoulos et al. [24] created a set of 37 handcrafted features that includes morphological (e.g., part of speech), vocabulary (e.g., type-to-token ratio), semantic (e.g., text and emoji sentiment), and lexical features (e.g., number of pronouns) to predict the false news using traditional ML algorithms. Further, in 2022 [23], they have extended the research to run different feature set with complex deep learning models.

AMR is a graph-based representation of natural language that accurately captures the complex semantics of a sentence in a way that is both language-independent and computationally tractable. A growing number of researchers are investigating how to use the information stored in the AMR graphs and its representations to assist in the resolution of other NLP problems. AMR has been successfully applied to more advanced semantic tasks such as entity linking [18], abstractive summarization [15], question answering [14], and machine translation [11]. Garg et al. [6] were the first to employ AMR representation for extracting interactions from biomedical text, utilizing graph kernel methods to determine if a given AMR subgraph expresses an interaction. Aguilar et al. [1] and Huang et al. [9] had conducted research and indicated that the semantic structures of sentences, such as AMR introduced in [2], encompass extensive and varied semantic and structural information concerning events. AMR graph has never been explored in the Fake News detection, however, considering its capability to determine the trigger words by extracting complex semantic information, AMR graphs have potential to improve efficiency of existing fake news detection methodologies. Recently, graph based fake news detection models [5, 10, 22] are proposed. Xu et al. [22] used evidence based encoded features with graph neural network to achieve better performance. The model effectively captures and incorporates long-distance semantic dependencies among scattered relevant snippets through neighborhood propagation. KAN [5] and FinerFact [10] models are proposed which takes social user information and Wikipedia information as an evidence to enhance the performance of the fake news detection model.

III. Methodology

The methodology in this study is divided into two parts: 1) curation of the proposed data set FauxNSA, and 2) fake news detection model FakEDAMR. Figure 2 shows the methodology, and the description of each step is provided in the following sections.

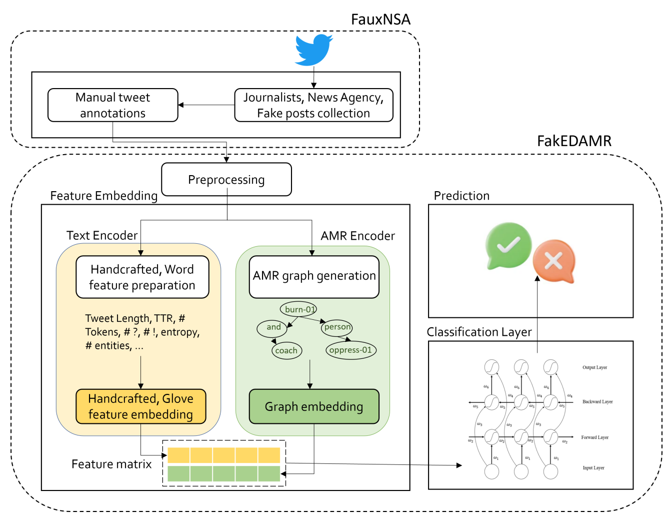

Fig. 2. Structural outline of proposed fake news detection methodology.

1) FauxNSA: Fake News Dataset on Nupur Sharma and Agniveer controversy:

a) Fake news dataset: The data set was gathered from the Twitter platform between May and September 2022 using the Twitter Academic API’s full-archive search over the political controversy ‘Nupur Sharama’ and ‘Agniveer’. This controversy holds the data related to religion, political, and terrorist issues. The methodology to collect the tweets can be broken down as follows. First, a list of curated hashtags mentioned in Table I related to ‘Nupur Sharama’ and ‘Agniveer’ controversy is manually constructed. Tweets were captured through Twitter

TABLE I
LIST OF CURATED HASHTAGS USED TO SCRAP TWEETS FROM TWITTER PLATFORM.

Category	Hashtags
Nupur Sharma Controversy	#NupurSharmaControversy, #Jamamasjid, #Nupur_Sharma, #NupurSharma, #HinduRashtra, #HindusUnderAttack, #SarTanSeJuda, #KanahiyaLal, #NupurSharmaBJP, #IsupportNupurSharma
Agniveer Controversy	#AgnipathRecruitmentScheme, #Agnipath, #Agniveer, #AgnipathProtests

API consisting at least one hashtag from the list. Total 31,889 tweets in Hindi and English languages including 31 features such as account information (display name, # of followers), tweet information (text, hashtag, URLs), and network information (quote, like, reply) are collected. We used popular fact checking websites such as BoomLive[1], NewsChecker[2], AltNews[3], etc. to annotate the data for fake news which were then manually searched over social media platforms and carefully annotated by two human annotators. We have also collected tweets from the 42 verified fact-checker Twitter accounts (PIBFackCheck, ABCFactCheck, etc.) from Twitter platform. After annotating and rigorously manually checking the Hindi tweets, they were translated into English using the Google Translate API available on the Google Cloud platform to fed into the model. This was done to ensure consistency in the dataset. After performing all the filtering process, we have 4632 tweets that are fake news.

b) Real news dataset: In the collection of real news, we have used the same approach discussed in [23]. That is, we have considered journalists and news agency as trustworthy source and collected tweets related to *Nupur Sharama* and *Agniveer* controversy. Overall, the account of 34 news agencies and 82 account of journalists with a global outreach are identified and gathered. Total 4657 tweets are collected and verified with the human annotators to make data set for the real news.

Figure 3 shows the word cloud of the collected true and fake data for Hindi and English languages. We can observe from that all the keywords are related to 'Nupur Sharama' and 'Agniveer' controversy only. After gathering tweets from news agencies, journalists and fake news sources, a comparison of their characteristics was made to determine their similarities. Specifically, the average number of hashtags per tweet was found to be 2.95 in the fake news data set, 3.12 for tweets posted by journalists, and 2.7 for those posted by news agencies. Additionally, the mean number of URLs per tweet was found to be 0.42 in the fake news data set, 0.55 for tweets posted by journalists, and 1.18 for those posted by news agencies. The statistics shows that collected data from all the sources is showing almost similar properties. Finally, the data set consists of 9289 tweets with 4632 fake and 4657 real tweets.

[1]https://www.boomlive.in/
[2]https://newschecker.in/
[3]https://www.altnews.in/

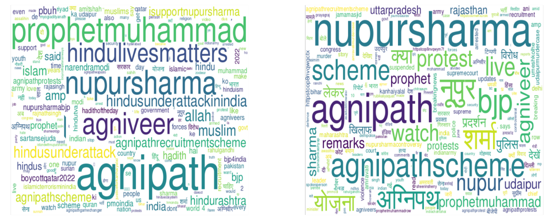

Fig. 3. Frequency word clouds of a) fake and b) true tweets collected from Twitter over 'Nupur Sharama' and 'Agniveer' controversy.

A. FakEDAMR: Fake nEws Detection using Abstract Meaning Representation

Proposed model, FakEDAMR, takes preprocessed text as input and predicts whether text document is real or fake. FakEDAMR comprises three primary components:

a) Text encoder: Research in the field of Natural Language Processing (NLP) has long focused on effectively representing sequential data. In line with previous studies, we have employed two different approaches to encode the sequence of tokens. The first approach involves the use of handcrafted features, consisting of 37 specific features outlined in [23]. The second approach utilizes GloVe embedding [21], which are pre-trained using a Twitter-based corpus comprising 27 billion tokens. This embedding maps each word to a d-dimensional vector. Although each word is represented by a fixed-size vector, the length of tweets can vary. To tackle this issue, we employ post-padding, which involves adding padding tokens at the end of a tweet to match the length of the longest tweet (100 tokens). Mathematically, the cost function of the GloVe embedding for a word in a word sequence (represented as $w =< w_1, ..., w_k >$) can be expressed as follows:

$$J = \sum_{p,q=1}^{V} f\left(N_{pq}\right)\left(w_p^T \tilde{w}_q + b_p + \tilde{b}_q - \log N_{pq}\right)^2 \quad (1)$$

Here, $f\left(N_{pq}\right)$ is weighting function, w_p^T is a context word vector and $\tilde{w}_q$ is out of context word vector and b_p, $\tilde{b}_q$ are bias terms. In Eq. 1, bias terms are also learned along with weight vector. Finally, we get the text embedding vector $t = ([t_i]_{i=1}^{i=d};\ t_i \in \mathbb{R}^{1\times m})$, where d is the fixed dimension and m is the maximum number of tokens.

b) AMR encoder: It is important to understand the complex semantics of the sentences which is not included in any previous work on fake news. we use AMR graphs for the same. AMR is a sembanking language that utilizes a rooted, directed, labeled, and mostly acyclic graph structure to capture the complete meanings of sentences. It employs multi-layer linguistic analysis techniques, including PropBank frames, non-core semantic roles, co-reference, named entity annotation, modality, and negation, to express the semantic structure of a sentence. AMR graph is composed of nodes represent semantic concepts such as entities, events, and attributes, and edges represent the relationships between those concepts, labeled with semantic roles such as agent, patient, location, time,

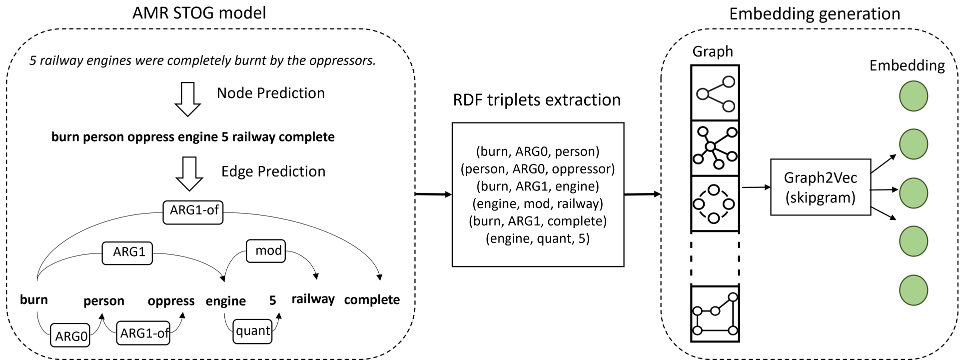

Fig. 4. Process Flow of AMR Encoder: Text-to-Graph Conversion, RDF Triplet Extraction, and Graph Embedding Generation.

and manner. Our approach involves several steps to generate an AMR graph from each text document (Figure 4) and extraction of RDF (Resource Description Framework) triplets from generated AMR graph. These triplets are represented in the form of *(subject, relation, object)*. Subsequently, we create a final graph using the extracted edges from the RDF triplets. Finally, we fed this converted graph into the Graph2Vec model to obtain the AMR graph embedding.

The AMR graph conversion process of each text in the document utilizes STOG model [25]. STOG model breaks down the sequence-to-graph task into two main components: node prediction and edge prediction.

In node prediction, the model takes an input sequence $w =< w_1, \ldots, w_k >$, where each word w_a is part of the sentence. It sequentially decodes a list of nodes $v =< v_1, \ldots, v_k >$ and assigns their indices $i =< i_1, \ldots, i_k >$ deterministically using the equation:

$$P(\boldsymbol{v}) = \prod_{a=1}^{k} P\left(v_a \mid v_{<a}, i_{<a}, w\right) \tag{2}$$

For edge prediction, given an input sequence w, a list of nodes v, and indices i, the model searches for the highest scoring parse tree y within the space $\mathcal{Y}$ of valid trees over v, while adhering to the constraint of i. A parse tree y represents a collection of directed head-modifier edges, depicted as:

$$y = \{(v_a, v_b) | 1 \leq a, b \leq k\} \tag{3}$$

To efficiently find the highest scoring parse tree (i.e., maximum spanning arborescence), the model utilizes a scoring mechanism used in [12].

$$\operatorname{parse}(\boldsymbol{v}) = \underset{y \in \mathcal{Y}(\boldsymbol{v})}{\arg\max} \sum_{(v_a, v_b) \in y} \operatorname{score}\left(v_a, v_b\right) \tag{4}$$

After obtaining the parse tree, the model proceeds with a merging operation to reconstruct the standard Abstract Meaning Representation (AMR) graph by combining nodes that share identical indices. Once we have the AMR tree denoted as y, we extract the RDF triplets from it. These triplets are represented as $t = \{(v_1, r_1, v_2), \ldots, (v_{k-1}, r_j, v_k)\}$. Each triplet consists of a subject v_a, a concept r_k, and an object v_b.

Using the extracted RDF triplets, we construct the final graph denoted as $g = (v, e, r)$. Here, v represents the set of vertices, specifically $v = \{v_1, \ldots, v_k\}$, r corresponds to the set of concepts obtained from the RDF triplets, i.e., $r = \{r_1, \ldots, r_j\}$. Lastly, e represents the set of edges in the graph, which is defined as $e = \{(v_a, r_j, v_b) | \exists\, v_a, v_b \in v \text{ and } r_j \in r\}$. In other words, the edges in e connect the vertices v_a and v_b using the relation r_j. This process of extracting RDF triplets and constructing the final graph enables the representation and analysis of the AMR graph, capturing the semantic relationships between entities and facilitating further processing and interpretation.

Afterward, a list of graph G, where each graph $g \in G$ represents one text, is passed as input to the Graph2Vec model, specifically the skip-gram model, to obtain the final embedding. The Graph2Vec model processes the AMR graph and generates embeddings by considering the graph structure and the relationships between its elements. The resulting embedding is obtained from the last hidden layer of the model, capturing the learned representation of the AMR graph in a vector form. Finally, we get the graph embedding vector $u = ([u_i]_{i=1}^{i=d};\ u_i \in \mathbb{R}^{n \times 1})$, where d is the fixed dimension and n is number of sentence in the document.

c) Classification layer: After getting the text embedding $t_{(m \times d)}$ and graph embedding $u_{(n \times d)}$, we get final embedding x by Eq. 5, where $|$ represents concatenation operation:

$$x = \left(t_{(m \times d)} \mid u_{(n \times d)}\right)_{(m+n) \times d} \tag{5}$$

Finally, BiLSTM model is used as classification layer to identify tweets in fake or real news using the prepared feature embedding x.

$$f_t = \operatorname{sigmoid}\left(W_{fx} x_t + W_{fh} h_{t-1} + b_f\right) \tag{6}$$
$$i_t = \operatorname{sigmoid}\left(W_{ix} x_t + W_{ih} h_{t-1} + b_i\right) \tag{7}$$
$$c_t = c_{t-1} \odot f_t + i_t \odot \tanh\left(W_{cx} x_t + W_{ch} h_{t-1} + b_c\right) \tag{8}$$
$$o_t = \operatorname{sigmoid}\left(W_{ox} x_t + W_{oh} h_{t-1} + b_o\right) \tag{9}$$
$$h_t = o_t \odot \tanh\left(c_t\right) \tag{10}$$

TABLE II
DISTRIBUTION OF DATA FOR DATASETS: A) COVID19-FND, B) KFN, AND C) FAUXNSA (OURS)

Dataset	Covid19-FND	KFN	FauxNSA
# Real	5100	10387	4657
# Fake	5600	10413	4632
# Total	**10700**	**20800**	**9289**

Where $x_t \in \mathbb{R}^{n \times d}$ is the input vector, $W \in \mathbb{R}^{l \times n}, b \in \mathbb{R}^{v}$ and the superscripts n and l depict the dimension of the input vector and the number of words in the dataset or vocabulary at any time t, respectively. For an input vector x_t, h_{t-1} and c_{t-1} are previously hidden and cell state, whereas the current hidden and cell state h_t and c_t. The above output represents the LSTM network. Finally, output of the BiLSTM can be summarized by concatenating the forward and backward state as $h_t = \left[\overrightarrow{h}_t, \overleftarrow{h}_t\right]$. At the output layer, it employed binary cross-entropy as the loss function to identify probability of true label $p(y_i)$ for real/fake classification.

IV. EXPERIMENTAL SETUP

A. Dataset

Other than our dataset, we have also used two publicly available dataset Covid-FND [20] and KFN [13]. Statistics of the data is given in the table II and detailed description of each is described below.

1) Kaggle Fake News (KFN) [13]: The Kaggle Fake News dataset includes 20,387 news items which spans the fields of politics, commerce, and technology, contains an evenly distributed mix of real and fake news pieces.

We ignored additional information, such as article content and author information, and concentrated only on the news titles for our analysis. Although the limited textual length available for classification makes it difficult to detect fake news with this option, it is consistent with earlier studies that show that a sizeable portion of fake news is disseminated on social media platforms like Twitter that have strict character limits.

2) Covid19-FND [20]: This dataset consists social media posts and articles related to COVID-19. These posts and articles are labeled as either real or fake. They specifically collected data from social media platforms that are actively used for peer communication and information sharing, including news, events, and social phenomena. To gather fake claims, authors referred to fact-checking websites such as Politifact, NewsChecker, Boomlive, and others. Additionally, they utilized tools like Google Fact-Check Explorer and the IFCN chatbot. Real news articles were collected from verified Twitter handles. This dataset contains 10,700 fake and real news related to COVID-19.

B. Metric

Four quality metrics, namely, Precision, Recall, F1-score, and Accuracy are considered for comparative study. Although accuracy is frequently employed as the primary metric for classification tasks, it might not be appropriate for unbalanced data sets. In these situations, we need to take into account alternative metrics that offer a more thorough review. One such metric is the F1-score, which takes into account both recall and precision and provides a balanced assessment by using their harmonic mean.

C. Implementation details

We have developed and tested our code in Keras (Python library). We partitioned each data set into training, validation, and testing sets, following a 70:20:10 split. This approach maintains the proportional representation of classes within both the train, validation, and test sets, enabling a robust evaluation of the model performance across various data sets. We have used basic preprocessing, like removing URLs, stopwords, etc., on each text document of the data set. We have incorporated AMR graph features on the feature sets proposed by [23]. They used two feature sets: Feature-set 1 adopts a feature engineering approach, where the chosen features are hand-crafted, including various categories such as morphological, vocabulary, and lexical features. Feature-set 2 employs tokenization of each tweet's text and conversion into word embedding. GloVe embedding [21], pre-trained with a Twitter-based corpus of 27 billion tokens, is used to map each word to a 100-dimensional vector. Despite each word being mapped to a fixed-size vector, tweet length still varies; to address this issue, post-padding (i.e., padding at the end of a tweet) is used to match the longest tweet (approximately 100 tokens). Therefore, a tweet in Feature-set 2 is presented as a 100x100 matrix. Although the size of Graph2Vec can vary based on the length of the AMR graph, we have fixed the dimension to 100, considering the length of the tweet is fixed in the Twitter platform. We evaluated Feature-set 1 on Naive Bayes, SVM, C4.5, random forests, and Feature-set 2 on CNN, C-LSTM, and BiLSTM. For the purpose of training each model on the data sets, we carried out three distinct trials with various seed values. The performance metric was then generated using the test data set findings, taking into account the best-performing trial. Model configuration, such as the number of hyper-parameters and number of layers used in the model, is kept the same as in the research [23].

V. RESULTS

We evaluated the performance of AMR with different feature sets on different ML/DL algorithms. Table III presents the comparison results of different models. It is evident that incorporating AMR semantic features into the feature sets significantly improves the performance of the models. Among the models evaluated using Feature-set 1, Random Forest with AMR-encoded feature sets achieves the highest accuracy of 88.90% and an F1-score of 85.92% on the FauxNSA (proposed) dataset. Furthermore, it also achieves the highest accuracy of 89.48% and 87.09%, along with F1-scores of 88.69% and 86.70%, on the publicly available datasets Covid19-FND and KFN, respectively.

BiLSTM with AMR-encoded features outperforms other models in the case of Feature-set 2. The model achieved an accuracy of 93.96% and an F1-score of 91.96% on our data set. Similar performance is observed on the other two publicly available data sets as well, where the accuracy and F1-scores of 93.26% and 93.20% on Covid19-FND, and 93.52% and 93.52% on KFN, respectively, are achieved.

VI. Ablation Study

A. Performance analysis of pretrained FakedAMR on South African Dataset

In order to assess the effectiveness of our model on an unknown set of samples, we utilized a South African fake news dataset (SA1) [4] obtained from various South African websites such as MyBroadband[4], News24[5], and MM Africa[6]. This dataset includes 807 fake articles in total, along with details like the article's title, publication date, and URL.

To evaluate the performance of our best-trained model on this South African dataset, we employed the models trained on the Covid19-FND, KFN, and FauxNSA datasets. The results revealed that the models trained on KFN and FauxNSA exhibited remarkable accuracy, achieving **97.78%** and **95.83%** respectively, on the unknown South African dataset. However, the model trained on the Covid19-FND dataset displayed a significantly lower accuracy of 39.38%. One possible explanation for this disparity is that the Covid19-FND dataset predominantly contains tokens related to the Covid-19 disease, while the KFN and FauxNSA datasets encompass a wider range of topics such as politics, commerce, and technology, which are more diverse and akin to the South African dataset.

B. Effect of AMR graph features

We conducted an investigation to understand why AMR features enhance the accuracy of the model. To provide evidence for our hypothesis, we visualized the features of the concatenation layer in the model, which merges the AMR features with textual features. In our experiment, we randomly selected three samples each from three datasets that demonstrated improved prediction accuracy when using both textual and AMR features compared to using only textual features in the model.

Figure 5 clearly illustrates the impact of AMR features on the model's performance. It shows that the AMR features create a distinct and decisive boundary that aids the model in effectively distinguishing between target classes. On the other hand, the textual features are scattered across dimensions and do not significantly contribute to the final prediction regarding the authenticity of the content (fake or real). These findings confirm that the integration of AMR features with textual features significantly enhances the model's ability to establish the correct decision boundary during the final prediction.

Through this series of experiments, we have successfully demonstrated the substantial contribution of AMR features in conjunction with textual features, enabling the model to create accurate decision boundaries and improve overall prediction accuracy.

[4]https://mybroadband.co.za/forum/threads/list-of-known-fake-news-sites-in-south-africa

[5]https://exposed.news24.com/the-website-blacklist/

[6]https://mediamonitoringafrica.org/

VII. Discussion and Conclusion

In this paper, we show that detecting fake news requires a more sophisticated understanding of the semantic relationships between trigger words and entities in the text. We demonstrated that how Abstract Meaning Representation (AMR) graph improves the fake news detection model and we concluded that semantic features are just as important as linguistic and syntactic features for identifying fake news in posts. Our experimental results indicate that for Fetaure-set 1, Random Forest with AMR-encoded features outperforms other traditional ML models with 88.90% accuracy and 85.92% F1-score. On the other hand for Feature-set 2, BiLSTM with AMR encoded features outperforms other methods, achieving 93.96% accuracy and 91.96% F1-score on our proposed dataset (FauxNSA) along with BiLSTM with AMR encoded features achieves accuracy and F1-scores of 93.26% and 93.20% on Covid19-FND, and 93.52% and 93.52% on KFN, two publicly available datasets respectively. In future, we are exploring the way to embed AMR graph with pretrained transformers based models such as Bert, XLM-Roberta, Electra, etc. Also, we are interested in exploring the more ways to encode AMR knowledge in order to increase the performance of existing fake news models.

References

[1] Jacqueline Aguilar, Charley Beller, Paul McNamee, Benjamin Van Durme, Stephanie Strassel, Zhiyi Song, and Joe Ellis. 2014. A Comparison of the Events and Relations Across ACE, ERE, TAC-KBP, and FrameNet Annotation Standards. In *Proceedings of the Second Workshop on EVENTS: Definition, Detection, Coreference, and Representation*. Association for Computational Linguistics, Baltimore, Maryland, USA, 45–53.

[2] Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2013. Abstract Meaning Representation for Sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*. Sofia, Bulgaria, 178–186.

[3] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2013. Predicting information credibility in time-sensitive social media. *Internet Research: Electronic Networking Applications and Policy* 23 (10 2013).

[4] Harm de Wet and Vukosi Marivate. 2021. Is it Fake? News Disinformation Detection on South African News Websites. In *2021 IEEE AFRICON*. 1–6.

[5] Yaqian Dun, Kefei Tu, Chen Chen, Chunyan Hou, and Xiaojie Yuan. 2021. KAN: Knowledge-aware Attention Network for Fake News Detection. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 1 (May 2021), 81–89.

TABLE III
COMPARISON OF THE PERFORMANCE OF DIFFERENT MODELS ON FEATURE-SETS (FS) 1 AND 2 FOR DATASETS COVID19-FND, KFN, AND FAUXNSA (OURS).

Model	Feature-set	Covid19-FND				KFN				FauxNSA			
		Precision	Recall	F1	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1	Accuracy
Naive Bayes	FS1	76.06	67.81	71.67	71.55	67.25	79.90	73.03	69.42	82.19	80.85	81.51	80.11
	FS1+AMR	78.83	66.93	72.39	72.58	67.46	79.59	73.10	70.13	84.41	82.06	83.21	80.81
SVM	FS1	82.05	81.47	81.76	80.71	85.91	83.80	84.84	83.71	84.58	82.87	83.71	84.35
	FS1+AMR	81.70	82.39	82.04	82.24	85.66	83.84	84.73	84.65	83.71	83.63	83.67	85.31
C4.5	FS1	81.94	79.57	80.74	79.85	80.81	79.56	80.18	80.28	86.08	83.16	84.59	83.32
	FS1+AMR	82.70	80.04	81.34	80.03	81.42	80.87	81.14	81.67	87.64	83.47	85.50	86.04
Random Forest	FS1	86.01	90.02	87.96	88.26	88.86	84.89	86.83	86.92	87.37	84.05	85.67	86.35
	FS1+AMR	**86.25**	**91.28**	**88.69**	**89.48**	**88.90**	**85.12**	**86.70**	**87.09**	**87.72**	**84.21**	**85.92**	**88.90**
CNN	FS2	91.16	91.30	91.20	91.21	91.74	91.27	91.50	91.52	89.25	84.35	86.34	89.64
	FS2+AMR	92.65	92.75	92.69	92.71	92.42	92.18	92.29	92.11	90.14	89.99	90.06	92.10
C-LSTM	FS2	91.49	91.46	91.47	91.51	91.58	91.58	91.57	91.57	91.61	86.13	88.34	91.11
	FS2+AMR	92.87	92.93	92.90	92.95	93.38	93.28	93.23	93.24	91.85	88.68	90.23	91.89
BiLSTM	FS2	91.55	91.71	91.54	91.55	92.44	92.40	92.41	92.36	90.60	86.83	88.45	91.12
	FS2+AMR	**93.43**	**93.08**	**93.20**	**93.26**	**93.55**	**93.54**	**93.52**	**93.52**	**92.25**	**91.67**	**91.96**	**93.96**

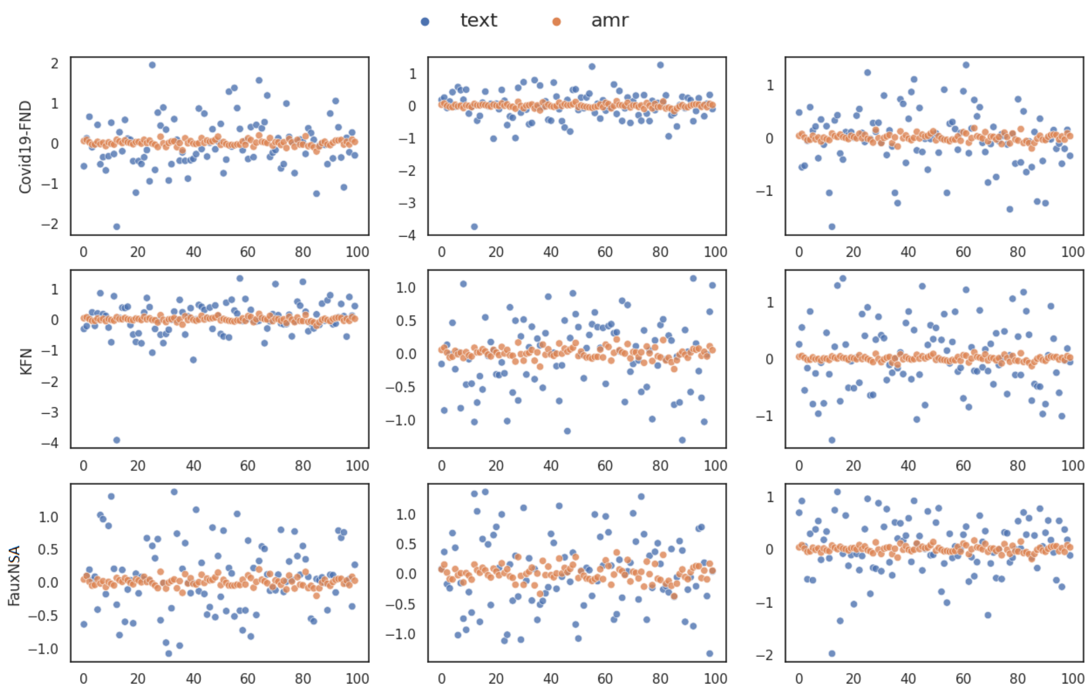

Fig. 5. Comparative analysis of AMR and text features in three datasets: Covid19-FND, KFN, and FauxNSA. Graph is plotted for three correct predicted samples by model where x-axis represents the feature index and y-axis represents its corresponding value.

[6] Sahil Garg, Aram Galstyan, Ulf Hermjakob, and Daniel Marcu. 2016. Extracting Biomolecular Interactions Using Semantic Parsing of Biomedical Text. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (Phoenix, Arizona) *(AAAI'16)*. AAAI Press, 2718–2726.

[7] Shubham Gupta and Suman Kundu. 2023. Interaction graph, topical communities, and efficient local event detection from social streams. *Expert Systems with Applications* 232 (2023), 120890.

[8] Linmei Hu, Siqi Wei, Ziwang Zhao, and Bin Wu. 2022. Deep learning for fake news detection: A comprehensive survey. *AI Open* 3 (2022), 133–155.

[9] Lifu Huang, Taylor Cassidy, Xiaocheng Feng, Heng Ji, Clare R. Voss, Jiawei Han, and Avirup Sil. 2016. Liberal Event Extraction and Event Schema Induction. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Berlin, Germany, 258–268.

[10] Yiqiao Jin, Xiting Wang, Ruichao Yang, Yizhou Sun, Wei Wang, Hao Liao, and Xing Xie. 2022. Towards fine-grained reasoning for fake news detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 5746–5754.

[11] Bevan Jones, Jacob Andreas, Daniel Bauer, Karl Moritz Hermann, and Kevin Knight. 2012. Semantics-based machine translation with hyperedge replacement grammars. In *Proceedings of COLING 2012*. 1359–1376.

[12] Eliyahu Kiperwasser and Yoav Goldberg. 2016. Simple and Accurate Dependency Parsing Using Bidirectional LSTM Feature Representations. *Transactions of the Association for Computational Linguistics* 4 (2016), 313–327.

[13] William Lifferth. 2018. Fake News. https://kaggle.com/competitions/fake-news.

[14] Jungwoo Lim, Dongsuk Oh, Yoonna Jang, Kisu Yang, and Heuiseok Lim. 2020. I Know What You Asked: Graph Path Learning using AMR for Commonsense Reasoning. In *Proceedings of the 28th International Conference on Computational Linguistics*. International Committee on Computational Linguistics, Barcelona, Spain (Online), 2459–2471.

[15] Fei Liu, Jeffrey Flanigan, Sam Thomson, Norman Sadeh, and Noah A. Smith. 2015. Toward Abstractive Summarization Using Semantic Representations. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Denver, Colorado, 1077–1086.

[16] Yunfei Long, Qin Lu, Rong Xiang, Minglei Li, and Chu-Ren Huang. 2017. Fake News Detection Through Multi-Perspective Speaker Profiles. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. Asian Federation of Natural Language Processing, Taipei, Taiwan, 252–256.

[17] Ray Oshikawa, Jing Qian, and William Yang Wang. 2018. A Survey on Natural Language Processing for Fake News Detection. *CoRR* abs/1811.00770 (2018). arXiv:1811.00770

[18] Xiaoman Pan, Taylor Cassidy, Ulf Hermjakob, Heng Ji, and Kevin Knight. 2015. Unsupervised entity linking with abstract meaning representation. In *Proceedings of the 2015 conference of the north american chapter of the association for computational linguistics: Human language technologies*. 1130–1139.

[19] John H Parmelee and Shannon L Bichard. 2011. *Politics and the Twitter revolution: How tweets influence the relationship between political leaders and the public*. Lexington books.

[20] Parth Patwa, Shivam Sharma, Srinivas Pykl, Vineeth Guptha, Gitanjali Kumari, Md Shad Akhtar, Asif Ekbal, Amitava Das, and Tanmoy Chakraborty. 2021. Fighting an Infodemic: COVID-19 Fake News Dataset. In *Combating Online Hostile Posts in Regional Languages during Emergency Situation*, Tanmoy Chakraborty, Kai Shu, H. Russell Bernard, Huan Liu, and Md Shad Akhtar (Eds.). Springer International Publishing, Cham, 21–29.

[21] Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. GloVe: Global Vectors for Word Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Doha, Qatar, 1532–1543.

[22] Weizhi Xu, Junfei Wu, Qiang Liu, Shu Wu, and Liang Wang. 2022. Evidence-Aware Fake News Detection with Graph Neural Networks. In *Proceedings of the ACM Web Conference 2022* (Virtual Event, Lyon, France) *(WWW '22)*. Association for Computing Machinery, New York, NY, USA, 2501–2510.

[23] Alexandros Zervopoulos, Aikaterini Alvanou, Konstantinos Bezas, Asterios Papamichail, Manolis Maragoudakis, and Katia Kermanidis. 2022. Deep learning for fake news detection on Twitter regarding the 2019 Hong Kong protests. *Neural Computing and Applications* 34 (01 2022), 1–14.

[24] Alexandros Zervopoulos, Aikaterini Georgia Alvanou, Konstantinos Bezas, Asterios Papamichail, Manolis Maragoudakis, and Katia Kermanidis. 2020. Hong Kong Protests: Using Natural Language Processing for Fake News Detection on Twitter. In *Artificial Intelligence Applications and Innovations*, Ilias Maglogiannis, Lazaros Iliadis, and Elias Pimenidis (Eds.). Springer International Publishing, Cham, 408–419.

[25] Sheng Zhang, Xutai Ma, Kevin Duh, and Benjamin Van Durme. 2019. AMR Parsing as Sequence-to-Graph Transduction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Florence, Italy, 80–94.