Domain Adaptive Fake News Detection via Reinforcement Learning

Ahmadreza Mosallanezhad
amosalla@asu.edu
Arizona State University
Tempe, AZ, USA

Mansooreh Karami
mkarami@asu.edu
Arizona State University
Tempe, AZ, USA

Kai Shu
kshu@iit.edu
Illinois Institute of Technology
Chicago, IL, USA

Michelle V. Mancenido
mmanceni@asu.edu
Arizona State University
Glendale, AZ, USA

Huan Liu
huanliu@asu.edu
Arizona State University
Tempe, AZ, USA

ABSTRACT

With social media being a major force in information consumption, accelerated propagation of fake news has presented new challenges for platforms to distinguish between legitimate and fake news. Effective fake news detection is a non-trivial task due to the diverse nature of news domains and expensive annotation costs. In this work, we address the limitations of existing automated fake news detection models by incorporating auxiliary information (e.g., user comments and user-news interactions) into a novel reinforcement learning-based model called **RE**inforced **A**daptive **L**earning **F**ake **N**ews **D**etection (REAL-FND). REAL-FND exploits cross-domain and within-domain knowledge that makes it robust in a target domain, despite being trained in a different source domain. Extensive experiments on real-world datasets illustrate the effectiveness of the proposed model, especially when limited labeled data is available in the target domain.

CCS CONCEPTS

• **Applied computing** → *Document analysis*; • **Information systems** → **Social networks**; • **Computing methodologies** → **Machine learning approaches**.

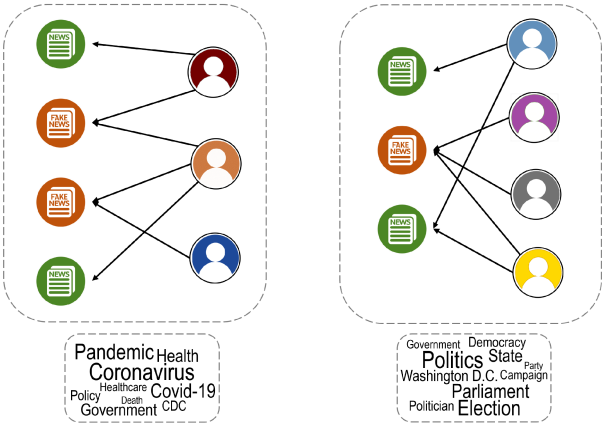
KEYWORDS

neural networks, reinforcement learning, domain adaptation, disinformation

ACM Reference Format:
Ahmadreza Mosallanezhad, Mansooreh Karami, Kai Shu, Michelle V. Mancenido, and Huan Liu. 2022. Domain Adaptive Fake News Detection via Reinforcement Learning. In *Proceedings of the ACM Web Conference 2022 (WWW '22), April 25–29, 2022, Virtual Event, Lyon, France.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3485447.3512258

WWW '22, April 25–29, 2022, Virtual Event, Lyon, France

ACM ISBN 978-1-4503-9096-5/22/04…$15.00
https://doi.org/10.1145/3485447.3512258



(a) Source Domain - Healthcare **(b) Target Domain - Politics**

Figure 1: Comparing two domains, Healthcare and Politics, in fake news detection: (1) Different word usage, (2) Event-related words/features, (3) Different user-news interaction.

1 INTRODUCTION

With people spending more time on social media platforms[1], it is not surprising that social media has evolved into the primary source of news among subscribers, in lieu of the more traditional news delivery systems such as early morning shows, newspapers, and websites affiliated with media companies. For example, it was reported that 1 in 5 U.S. adults used social media as their primary source of political news during the 2020 U.S. presidential elections [23]. The ease and speed of disseminating new information via social media, however, have created networks of disinformation that propagate as fast as any social media post. For example from the COVID pandemic era, around 800 fatalities, 5000 hospitalizations, and 60 permanent injuries were recorded as a result of false claims that household bleach was an effective panacea for the SARS-CoV-2 virus [5, 12]. Unlike traditional news delivery where trained reporters and editors fact-check information, curation of news on social media has largely been crowd-sourced, i.e., social media users themselves are the producers and consumers of information.

[1]In 2020, internet users spent an average of 145 minutes per day on social media [2].

In general, humans have been found to fare worse than machines at prediction tasks [31, 32] such as distinguishing between fake or legitimate news. Machine learning models for automated fake news detection have even been shown to perform better than the most seasoned linguists [25]. To this end, many automated fake news detection algorithms have largely focused on improving predictive performance for a specific news domain (e.g., political news). The primary issue with these existing state-of-the-art detection algorithms is that while they perform well for the domain they were trained on (e.g., politics), they perform poorly in other domains (e.g., healthcare). The limited cross-domain effectiveness of algorithms to detect fake news are mostly due to (1) the reliance of content-based approaches on word usage that are specific to a domain, (2) the model's bias towards event-specific features, and (3) domain-specific user-news interaction patterns (Figure 1). As one of the contributions of this work, we will empirically demonstrate that the advertised performance of SOTA methods is not robust across domains.

Additionally, due to the high cost and specialized expertise required for data annotation, limited training data is available for effectively training an automated model across domains. This calls for using auxiliary information such as users' comments and motivational factors [15, 37] as value-adding pieces for fake news detection.

To address these challenges, we propose a domain-adaptive model called **RE**inforced **A**daptive **L**earning **F**ake **N**ews **D**etection (REAL-FND), which uses generalized and domain-independent features to distinguish between fake and legitimate news. The proposed model is based on previous evidence that illustrates how domain-invariant features could be used to improve the robustness and generality of fake news detection methods. As an example of a domain-invariant feature, it has been shown that fake news publishers use click-bait writing styles to attract specific audiences [46]. On the other hand, patterns extracted from the social context provide rich information for fake news classification within a domain. For example, a user's comment providing evidence in refuting a piece of news is a valuable source of auxiliary information [34]. Or, if a specific user is a tagged fake news propagator, the related user-news interaction could be leveraged as an additional source of information [15].

In REAL-FND, instead of applying the commonly-used method of adversarial learning in training the cross-domain model, we transform the learned representation from the source to the target domain by deploying a reinforcement learning (RL) component. Other RL-based methods employ the agent to modify the parameters of the model. However, we use the RL agent to modify the learned representations to ensure that domain-specific features are obscured while domain-invariant components are maintained. An RL agent provides more flexibility over adversarial training because any classifier's confidence values could be directly optimized (i.e., the objective function does not need to be differentiable).

We address the challenges in domain adaptive fake news detection algorithms by making the following contributions:

- We design a framework that encodes news content, user comments, and user-news interactions as representation vectors and fuses these representations for fake news detection;
- We utilize a reinforcement learning setting to adjust the representations to account for domain-independent features;
- We conduct extensive experiments on real-world datasets to show the effectiveness of the proposed model, especially when limited labeled data is available in the target domain.

2 RELATED WORK

The proposed methodology spans the subject domains of fake news detection, cross domain modeling, and the use of reinforcement learning in the Natural Language Processing (NLP). Current state-of-the-art in these areas are discussed in this section.

2.1 Fake News Detection

Early research in fake news detection focused on extracting features from news content to either capture the textual information (such as lexical and syntactic features) [33, 36] or to detect the differences in the writing style of real and fake news such as deception [7] and non-objectivity [26]. Qian et al. propose to use convolutional neural network to detect fake news based on news content only [27]. Guo et al. propose to capture the sensational emotions to create an emotion-enhanced news representation to detect fake news [9]. Recent advancements utilize deep neural networks [16] or tensor factorization [11] to model latent news representations.

While news content-based approaches result in acceptable performance, incorporating auxiliary information improves the performance and reliability of fake news detection models [34]. For example, Tacchini et al. [40] and Guo et al. [10] aggregated users' social responses (topics, stances, and credibility), Ruchansky et al. proposed a model called CSI which uses a hybrid model on news content and users network to detect fake news [29], and Shu et al. used hierarchical attention networks to create more explainable fake news detection models [34]. In this work, we incorporate various auxiliary information (i.e., users' comments and user-news interactions) along with the news content to create a well-defined news article representation.

2.2 Cross Domain Modeling

Cross-domain modeling refers to a model capable of learning information from data in the source domain and being able to transfer it to a target domain [1]. In general, cross-domain models are categorized into *sample-level* and *feature-level* groups [47]. Sample-level domain adaptation methods focus on finding domain-independent samples by assigning weights to these instances [39, 47]. On the other hand, feature-level domain adaptation methods focus on weighting or extracting domain-independent features [18]. In addition to the aforementioned domain adaptation methods, Gong et al. combined both sample-level and feature-level domain adaptation [8] in BERT to create a domain-independent sentiment analysis model. Similarly, Vlad et al. used transfer learning on an enhanced BERT architecture to detect propaganda across domains [43]. Moreover, Zhuang et al. propose to use auto-encoders for learning unsupervised feature representations for domain adaptation [47]. The goal of this model is to leverage a small portion of the target domain data to train an auto-encoder for learning domain-independent feature representations. In this paper, we focus on feature-level

domain adaptation by leveraging reinforcement learning to learn a domain-independent textual representation.

2.3 Reinforcement Learning

In past years, reinforcement learning has shown to be useful in improving the effectiveness of NLP models. As such, Li et al. propose a method for generating paraphrases using inverse reinforcement learning [21]. Another work by Fedus et al. uses reinforcement learning to tune parameters of an LSTM-based generative adversarial network for text generation [6]. A recent work by Cheng et al. proposes to use reinforcement learning for tuning a classifier's parameters for removing various types of biases [3]. Inspired by these methods, we study the problem of domain-adaptive fake news detection using reinforcement learning.

In previous research, although reinforcement learning has been used to modify a model's parameters, little research has been done on applying it to modify the learned representations. In this work, we focus on designing a domain adaptation model with an RL agent that modifies the article's representation based on the feedback received from both the adversary domain classifier and the fake news detection component.

3 PROBLEM STATEMENT

Let $\mathcal{D}_s = \{ (\mathbf{x}_1^s, y_1^s), (\mathbf{x}_2^s, y_2^s), ..., (\mathbf{x}_N^s, y_N^s)\}$ and $\mathcal{D}_t = \{ (\mathbf{x}_1^t, y_1^t), (\mathbf{x}_2^t, y_2^t), ..., (\mathbf{x}_M^t, y_M^t)\}$ denote a set of N and M news article with labels from source and target domains, respectively. Each news article $\mathbf{x}_i$ includes a content which is a sequence of K words $\{w_1, w_2, ..., w_K\}$, a set of comments $\mathbf{c}_j \in C$, and user-news interactions $\mathbf{u}_j \in \mathcal{U}$. User-news interactions $\mathbf{u}_i$ is a binary vector indicating users who posted, re-posted, or liked a tweet about news $\mathbf{x}_i$. The goal of the reinforced domain adaptive agent is to learn a function that converts the news representation $\mathbf{x}_i \in \mathcal{D}_t$ from target domain to source domain. The problem is formally defined as follows:

Definition (Domain Adaptive Fake News Detection). Given news articles from two separate domains $\mathcal{D}_s$ and $\mathcal{D}_t$, corresponding users' comments C_s and C_t, and user-news interactions $\mathcal{U}_s$ and $\mathcal{U}_t$ from the source S and target T domains, respectively, learn a domain-independent news article representation using $\mathcal{D}_s$ and a small portion of $\mathcal{D}_t$ that can be classified correctly by the fake news classifier F.

4 PROPOSED MODEL

In this section, we describe our proposed model, **RE**inforced domain **A**daptation **L**earning for **F**ake **N**ews **D**etection (REAL-FND). The input of this model is the news articles from source domain $\mathcal{D}_s$ and a portion (γ) of the target data set $\mathcal{D}_t$. As shown in Figure 2, the REAL-FND model has two main components: (1) the news article encoder, and (2) the reinforcement learning agent. In the following subsections, we explain these two components in detail.

4.1 News Article Encoder

The problem of detecting fake news requires learning a comprehensive representation that includes information about news content and its related auxiliary information. In this paper, we consider news comments and user-news interactions as auxiliary information for fake news detection. Recurrent neural networks (RNN) such as Long Short Term Memory (LSTM) has been proven to be effective in modeling sequential data [36]. However, Transformers, due to their attentive nature, create a better text representation that includes vital information about the input in an efficient manner [42]. Thus, in this work, we use Bidirectional Encoder Representations from Transformers (BERT) in creating the representation vector for the news content.

BERT is a pre-trained language model that uses transformer encoders to perform various NLP-based tasks such as text generation, sentiment analysis, and natural language understanding. Previous research have used BERT to achieve state-of-the-art performance on different applications of sentiment analysis tasks [20, 42, 44]. Although BERT has been pre-trained on a large corpus of textual data, it should be fine-tuned to perform well on a specific task [22, 44]. In this paper, to create a well-defined news content representation, we fine-tune BERT using news articles $\mathbf{x} = \{w_1, w_2, ..., w_K\}$ from source domain dataset $\mathcal{D}_s$ and a portion γ of the target dataset $\mathcal{D}_t$.

In addition to the news content $\mathbf{x}_i$, we also consider the article's comments and user-news interactions. In the experiments, we show that using auxiliary information leads to better detection performance as the model accounts for user's reliability and feedback on the news article. Nonetheless, previous studies also have shown that comments on news articles can improve fake news detection [34] by extracting semantic clues confirming or disapproving the authenticity of the content. Due to the fact that not all comments are useful for fake news classification, we use Hierarchical Attention Network (HAN) to encode the comments of a news article. The hierarchical structure of HAN facilitates the importance of every comment, as well as the salient word features. To create a representation for the news article comments, we pre-train HAN by stacking it with a feed-forward neural network classifier. After pre-training HAN, we remove the feed-forward classifier and only use the HAN to encode the news article comments. Note that in case a news article does not have any comments, we use vectors with zero values for comments representation.

Moreover, in addition to the news article comments, we also consider the user-news interactions to improve fake news detection. For a news article $\mathbf{x}$, the user-news interactions $\mathbf{u}$ is a binary vector where $\mathbf{u}_i$ indicates user i has tweeted, re-tweeted, or commented on a tweet about that news. Thus, the user-news interactions vector is a representation of the user behaviour toward a news article. To encode this information, we use a feed-forward neural network that takes the binary vector of user-news interaction as input and returns a representation containing important information about that interactions.

After constructing the representation networks (i.e., HAN, BERT, and the feed-forward neural network) and pre-training BERT and HAN, we concatenate the output of these three components into one vector $E = (h'_{comments}||h'_{content}||h'_{interactions})$ (|| indicates concatenation) and pass it to another feed-forward network to combine these information into a single vector $\mathbf{E}'$ [14]. Once we stacked the representation network with a feed-forward neural network classifier, we train a fake news classifier using the source domain data $\mathcal{D}_s$ and a portion γ of the target domain data $\mathcal{D}_t$. After training the fake news classifier F, we freeze the weights of the

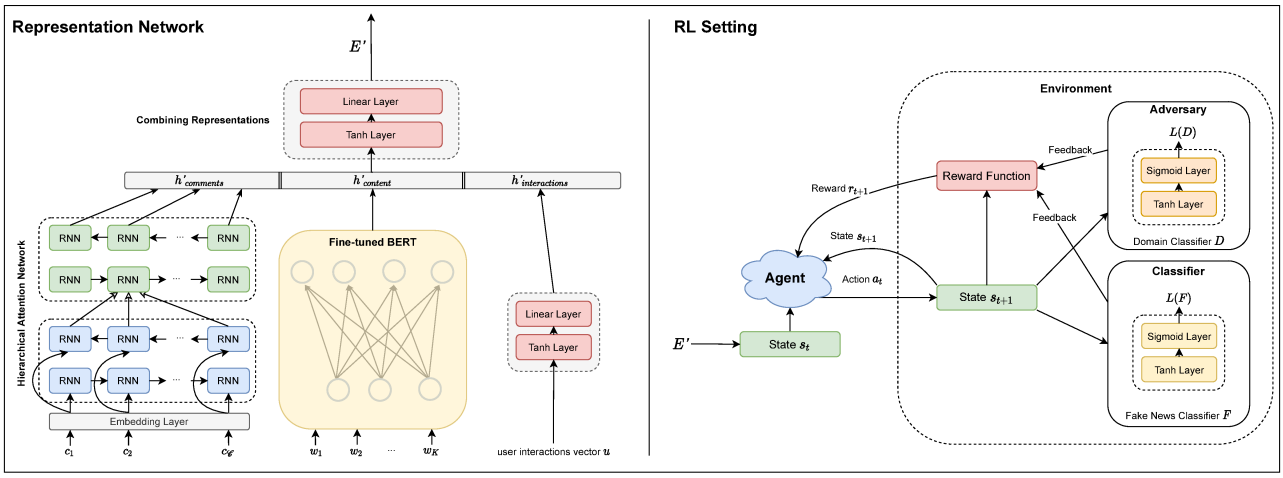


Figure 2: The proposed REAL-FND architecture. REAL-FND has two main components, the representation network, and the reinforcement learning agent. The representation network learns a representation including information about news content, comments, and user-news interactions. The RL agent learns a method to transfer the learned representation from source domain to the target domain.

representation network and train a domain classifier D using the representation $\mathbf{E}'$ of the news articles. By the end of this process, we have created three sub-components: (1) a representation network that encodes news content, comments, and user-news interactions, (2) a single-domain fake news classifier, and (3) a domain classifier. In the following section, we discuss the second component of the model (i.e., Figure 2-RL Setting) that uses reinforcement learning to create a domain adaptive representation using $\mathbf{E}'$.

4.2 Reinforced Domain Adaptation

Inspired by the success of RL [13, 24, 41], we model this problem using RL to automatically learn how to convert textual representations from both source domain data $\mathcal{D}_s$ and target domain data $\mathcal{D}_t$. Instead of applying a commonly-used approach of adversarial learning to train both F and D classifiers, we utilized an RL-based technique. RL would transform the representation to a new one such that it works well on the fake news classifier F, but not on the D classifier (i.e., the adversary). In this approach, the agent interacts with the environment by choosing an action a_t in response to a given state s_t. The performed action results in state s_{t+1} with reward r_{t+1}. The tuple $(s_t, a_t, s_{t+1}, r_{t+1})$ is called an experience which will be used to update the parameters of the RL agent.

In our RL model, an agent is trained to change the news article representation $\mathbf{E}'$ into a new representation that deceives the domain classifier, but preserves the accuracy for the fake news classifier. The RL agent learns this transition by changing the values in the input vector $\mathbf{E}'$ and receiving feedback from both fake news classifier F and domain classifier D. To create an RL setting, we define four main parts of RL in our problem, i.e., environment, state, action, and reward.

- **Environment**: the environment in our model includes the RL agent, pre-trained fake news classifier F, and the pre-trained domain classifier D. In each turn, the RL agent performs an action on the news article representation E' (known as state s_t) by changing one of its values (performing action a_t). The modified representation is passed through classifiers F and D to get their confidence scores to calculate the reward value. Finally, the reward value and the modified representation is passed to the agent as reward r_{t+1} and state s_{t+1}, respectively.
- **State**: the state is the current news article representation $\mathbf{E}'$ that describes the current situation for the RL agent.
- **Actions**: actions are defined as selecting one of the values in news article representation $\mathbf{E}'$ and changing it by adding or subtracting a small value σ. The total number of actions are $|E'| \times 2$.
- **Reward**: since the aim is to nudge the agent into creating a domain adaptive news article representation, the reward function looks into how much the agent was successful in removing the domain-specific features from the input representation. Specifically, we define the reward function at state s_{t+1} according to the confidence of both the domain classifier and the fake news classifier. Considering the news article embedding $\mathbf{E'}_{t+1}$ with its label i (being fake or real) and domain label j at state s_{t+1}, we formally define the reward function as follows:

$$r_{t+1}(s_{t+1}) = \alpha Pr_F(l = i|\mathbf{E'}_{t+1}) - \beta Pr_D(l = j|\mathbf{E'}_{t+1}) \quad (1)$$

where l indicates the label.

4.3 Optimization Algorithm

Given the RL setting, the aim is to learn the optimal action selection strategy $\pi(s_t, a_t)$. Algorithm 1[2] shows this optimization process. At each timestep t, the RL agent changes one of the values in the news article representation, $s_t = \mathbf{E'}_t$, and gets the reward value, r_{t+1}, based on the modified representation $s_{t+1} = \mathbf{E'}_{t+1}$ and the

[2]The source code will become publicly available upon acceptance.

Algorithm 1 The Learning Process of REAL-FND

Require: News article representations $\mathbf{E}' \in \mathcal{D}$ - fake news classifier F - domain classifier D - parameters α, β, and λ - terminal time T.

Initialize state s_t and memory M.
while training is not terminal **do**
  $s_t \leftarrow \mathbf{E}'$
  for $t \in \{0, 1, ..., T\}$ **do**
    Choose action a_t according to current distribution $\pi(s_t)$
    Perform a_t on s_t and get (s_{t+1}, r_{t+1})
    $M \leftarrow M + (s_t, a_t, r_{t+1}, s_{t+1})$
    $s_t \leftarrow s_{t+1}$
    for each timestep t, reward r in M_t **do**
      $G_t \leftarrow \sum_{i=1}^{t} \lambda^i r_{i+1}$
    end for
    Calculate policy loss according to Equation 2
    Update the agent's policy according to Equation 3
  end for
end while

selected action a_t. The goal of the agent is to maximize its reward according to Equation 1.

To train the agent, we use the REINFORCE algorithm which uses policy gradient to update the agent [45]. Considering the agent's policy according to parameters θ as $\pi_\theta(s_t, a_t)$, the REINFORCE algorithm uses the following loss function to evaluate the agent:

$$\mathcal{L}(\theta) = \log(\pi_\theta(s_t, a_t) \cdot G_t) \tag{2}$$

where $G_t = \sum_{i=1}^{t} \lambda^i r_{i+1}$ is the cumulative sum of discounted reward, and λ indicates the discount rate. The gradient of the loss function is used to update the agent:

$$\nabla\theta = \mathrm{lr}\nabla_\theta \mathcal{L}(\theta) \tag{3}$$

where lr indicates the learning rate.

5 EXPERIMENTS

In the designed experiments, we try to understand how the differences between domains affect the performance of fake news detection models. Moreover, we investigate how the proposed RL-based approach will overcome performance degradation due to the differences in news domains. Our main evaluation questions are as follows:

- **Q1.** How well does the current fake news detection methods perform on social media data?
- **Q2.** How well does the proposed model detect fake news on a *target domain* $\mathcal{D}_t$ after training the model on a *source domain* $\mathcal{D}_s$?
- **Q3.** How do auxiliary information contribute to the improvement of the fake news detection performance?

Q1 evaluates the quality of fake news detection models. We answer this question by training and testing fake news detection models on the same domain and comparing their performance. **Q2** studies the effect of domain differences on the performance of fake news detection models. We use a similar approach to **Q1** to answer this question by training on one domain, but testing on another. In **Q3** we perform ablation studies to analyze the impact of auxiliary information, the reward function parameters α and β, and the portion of target domain data γ needed to achieve an acceptable detection performance.

	Politifact	GossipCop
# True News	2,645	3,586
# Fake News	2,770	2,230
# News	5,415	5,819
# News with Comments	415	5,819
# Users	60,053	43,918
# Unique Users	100,520	

Table 1: The statistics of *Politifact* and *GossipCop* datasets. The *Politifact* dataset contains news articles from both FakeNewsNet and the dataset provided by Rashkin et al. [28], while *GossipCop* contains news articles from FakeNewsNet only.

5.1 Datasets

We use the well-known fake news data repository FakeNewsNet [35] which contains news articles along with their auxiliary information such as users' metadata and news comments. These news articles have been fact-checked with two popular fact-checking platforms - *Politifact* and *GossipCop*. Politifact fact-checks news related to the U.S. political system, while GossipCop fact-checks news from the entertainment industry. In addition to the existing Politifact news articles from FakeNewsNet, we enrich the dataset by adding $5,000$ annotated Politifact news from the dataset introduced by Rashkin et al. [28]. Table 1 shows the statistics of the final dataset. The Politifact news articles from [28] includes truth rating from 0 to 5, in which we only consider news with label $\in \{0, 4, 5\}$.

5.2 Data Pre-processing

Each news articles from the final dataset contains news content, users' comments, and their meta-data. We pre-process the textual data (i.e., news content and users' comments) by removing the punctuation, mentions, and out-of-vocabulary words. Further, as BERT has a limitation of getting 512 words as input, we truncate the news content and every comment to include its first 512 words. Finally, we utilize the users' meta-data to create user-news interaction matrix by tracking every user's interactions with news articles.

5.3 Implementation Details

The training process has been conducted in three stages: (1) pre-training the representation networks, (2) training the fake news and domain classifiers, and (3) training the RL agent. In what follows, we will expand the implementation details of each stage.

Pre-training: In this stage, we fine-tune BERT and HAN networks to generate a reasonable text representation from news content and users' comments, respectively. Motivated by the low memory consumption of the distilled version of Bert [30], we used the base model of Distilled BERT for creating the textual representation of the news contents. The model was fine-tuned for 3 iterations using a classifier on top of it. Moreover, to create the representations

related to the user's comments, we pre-trained HAN by stacking it with a simple fake news classifier and fine-tuning it for 5 iterations. After fine-tuning both models, the classifier module in both distilled Bert and HAN were removed. These pre-trained networks were placed in the final architecture of our model.

Training classifiers F and D: With passing the news content through the representation network in Figure 2, we trained a fake news classifier and a domain classifier using the news content representations. During the training, we did not update the weights of BERT and HAN networks. In this stage, a dropout with $p = 0.2$ was used for both fake news and domain classifiers. Both classifiers use a similar feed-forward neural network with a single hidden layer of 256 neurons. The networks are trained using Adam optimizer with a learning rate of $1e-5$ and a Cross Entropy loss function:

$$\mathcal{L}_{CE} = -\frac{1}{M}\sum_{i=1}^{M}(y_i \log(p_i) + (1-y_i)\log(1-p_i)) \tag{4}$$

Training the RL agent: Once finished with the first two stages, the fake news classifier and the domain classifier are placed as the reward function of the RL setting to train the agent. We train the agent using Algorithm 1 for $2,000$ episodes performing only $T = 20$ actions. As the long term reward function is important, we used a discount rate of $\lambda = 0.99$. In updating the agent network, we applied Adam optimizer and set the parameters with $\alpha = \beta = 0.5$, $\sigma = 0.01$, and learning rate $lr = 1e-4$. The RL agent uses a feed-forward neural network with 2 hidden layers of 512 and 256 neurons.

5.4 Baselines

For evaluating the effectiveness of REAL-FND on fake news detection task, we compare it to the baseline models described below. For a comprehensive comparison we consider state-of-the-art baselines that (1) only use news content (BERT-BiLSTM and BERT-UDA), (2) use both news content and users' comments (TCNN-URG and dEFEND), (3) consider users' interactions (CSI), and (4) considers propagation network (RoBERT-EMB). In addition to these baselines, we consider two variants of REAL-FND, Simple-REAL-FND and Adv-REAL-FND that use bi-directional GRU instead of BERT/HAN and adversarial training, respectively. These two baselines help us to study the effectiveness of using a complex architecture such as BERT and RL for domain adaptive fake news detection.

- **TCNN-URG (URG)** [27]: Based on TCNN [17], this model uses convolutional neural networks to capture various granularity of news content as well as including users' comments.
- **CSI** [29]: CSI is a hybrid model that utilizes news content, users' comments, and the news source. The news representation is modeled using LSTM neural network using Doc2Vec [19] that outputs an embedding for both news content and users' comments. For fair comparison, we disregarded the news source feature.
- **dEFEND (DEF)** [34]: This model uses a co-attention network to model both news content and users' comments. dEFEND captures explainable content information for fake news detection.
- **BERT-UDA (UDA)** [8]: This model uses feature-based and instance-based domain adaptation on BERT to create domain-independent news content representation.
- **BERT-BiLSTM (BBL)** [43]: This model uses a complex neural network model including BERT [42], bi-directional LSTM layer, and a capsule layer to classify news content. This model uses transfer learning to work across different domains.
- **RoBERTa-EMB (EMB)** [39]: This model adopt instance-based domain adaptation. It uses RoBERTa-base [22] to create news content representation, while uses an unsupervised neural network to create the propagation network's representation [38]. To evaluate this method on single-domain, we disable the domain adaptation process.
- **Simple-REAL-FND (SRE)**: To study the effect of using complex networks such as BERT and HAN for creating the representation of the news content and users comments, we replace BERT and HAN with a bi-directional Gated Recurrent Unit (BiGRU) stacked with a location-based attention layer [4]. Specifically, we use two separate BiGRU neural networks to create $h'_{content}$ and $h'_{comments}$. To create $h'_{comments}$, we concatenate all of the comments and select the first 256 words.

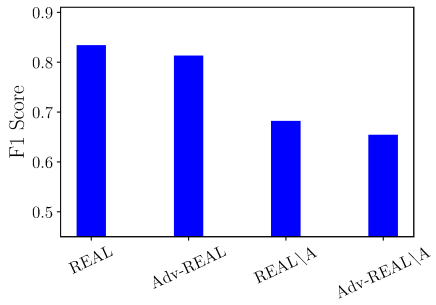
5.5 Experimental Results

To answer questions **Q1** and **Q2**, we trained REAL-FND and the baselines on a source domain S and tested them on both source domain S and target domain T. We used k-fold validation and calculated the average AUC and F1 scores. For single-domain and cross-domain, we set the number for folds as $k = 9$ and $k = 10$, respectively. In single-domain training, we use the source domain data $\mathcal{D}_s$, while for training REAL-FND and the baselines in a cross-domain setting, we use the source domain data $\mathcal{D}_s$ combined with portion γ of the target domain data $\mathcal{D}_t$. In the subsequent subsection it will be shown in Figure 4c that performance improvements taper off after using 30%($\gamma = 0.3$) of the target domain data. Finally, to answer **Q3**, we perform an ablation study to measure the impact of using auxiliary information and reinforcement learning.

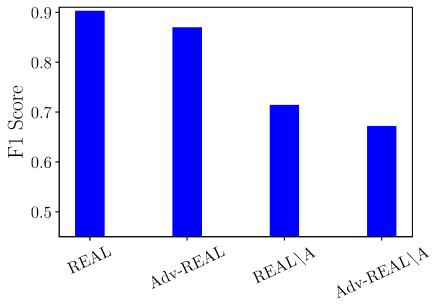
Fake News Detection Results (Q1). Table 2 shows a comparison of the baselines with REAL-FND. Training is applied on a single domain dataset $\mathcal{D}_s$ and tested on both source and target domains. For this experiment, we removed the domain classifier's feedback from the RL agent by setting the parameter $\beta = 0$. From the results we conclude that (1) all baseline models and REAL-FND perform reliably well when both training and testing news come from a single domain, and (2) due to the considerable decrease in performance when testing with news from another domain, it appears that the *Politifact* and *GossipCop* news domains have different properties. These results imply that the evaluated models are not agnostic to domain differences. REAL-FND performed better than baselines for the majority of scenarios. The only case where REAL-FND is under-performed is when the model is trained and tested on the *GossipCop* dataset. The better performance of REAL-FND on the *Politifact* domain suggests that BBL may have overfitting issues and REAL-FND benefits the use of auxiliary information for creating a more general fake news classifier that performs well on both domains. It is worth mentioning that EMB and BBL models are similar to each other in terms of model architecture except that EMB utilizes the user's propagation network as well. Comparing the

Model	GossipCop → Politifact		GossipCop → GossipCop		Politifact → Politifact		Politifact → GossipCop	
	F1	AUC	F1	AUC	F1	AUC	F1	AUC
CSI	0.532±0.11	0.563±0.14	0.811±0.05	0.832±0.08	0.756±0.02	0.785±0.10	0.589±0.10	0.610±0.13
URG	0.448±0.11	0.450±0.11	0.770±0.04	0.792±0.10	0.526±0.01	0.741±0.08	0.460±0.09	0.532±0.15
DEF	0.621±0.10	0.687±0.11	0.857±0.03	0.893±0.07	0.768±0.01	0.793±0.06	0.423±0.06	0.561±0.11
BBL	0.695±0.07	0.715±0.09	**0.918**±0.02	0.947±0.08	0.824±0.02	**0.868**±0.04	0.558±0.09	0.597±0.05
UDA	0.673±0.08	0.695±0.09	0.901±0.02	0.923±0.06	0.727±0.02	0.762±0.02	0.596±0.09	0.602±0.07
EMB	0.704±0.07	0.719±0.08	0.914±0.04	**0.952**±0.05	0.835±0.04	0.852±0.03	0.586±0.06	0.601±0.05
SRE	0.656±0.09	0.714±0.11	0.867±0.05	0.931±0.07	0.720±0.01	0.751±0.05	0.429±0.08	0.521±0.09
REAL-FND	**0.726**±0.11	**0.728**±0.10	0.905±0.10	0.933±0.08	**0.844**±0.05	0.810±0.08	**0.601**±0.09	**0.612**±0.11

Table 2: *Single-domain:* Performance of REAL-FND and baselines, trained *only on one domain* and tested on both *Politifact* and *GossipCop* domains (Train → Test). The values indicate *F1* and *AUC* measures on the test set.



(a) GossipCop as source, Politifact as target domain



(b) Politifact as source, GossipCop as target domain

Figure 3: The impact of auxiliary information and reinforcement learning for cross-domain fake news detection. Different versions of REAL-FND are evaluated on the target domain.

results between these two models also reveals that using auxiliary information can be helpful in detecting fake news.

Cross-Domain Results (Q2). Table 3 shows the performance of models on the target domain. In this experiment, we used $\gamma = 0.3$ of nine-folds of the target domain dataset $\mathcal{D}_t$ in addition to all the source domain dataset $\mathcal{D}_s$. Performance measures are calculated based on the tenth-fold of the target domain dataset. The results indicate that REAL-FND detects fake news in the target domain most efficiently, in comparison to the baselines. For example, compared with the best baselines, both F1 and AUC scores have improved. This indicates that most baselines suffer from overfitting on the source domain. Moreover, the results show that using a small portion of the target domain dataset can lead to a large increase in the cross-domain classification, indicating that the RL agent learns more domain-independent features by the feedback received from the domain classifier. It is also notable that Simple-REAL-FND (SRE) perform surprisingly well despite using weaker network architecture than the most baselines.

Impact of RL and Auxiliary Information (Q3). To show the relative impact using reinforcement learning and auxiliary information, we created the following variants of REAL-FND:

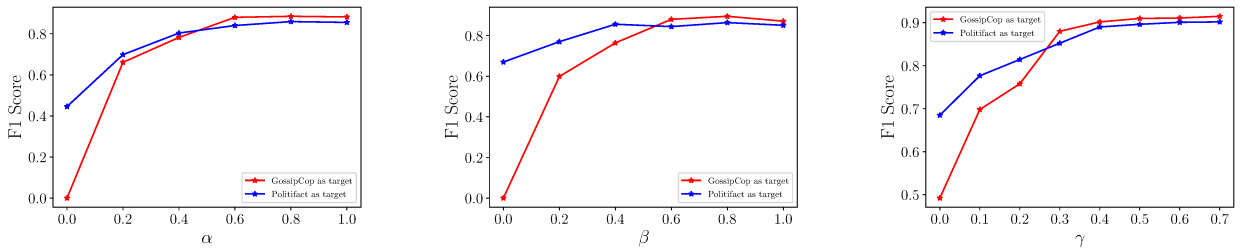
- **REAL-FND \A**: To study the effects of using auxiliary information in fake news detection, we created a variant of REAL-FND which does not use users' comments and user-news interactions. Thus, this model only uses BERT to create news article representations.

Model	GossipCop → Politifact		Politifact → GossipCop	
	AUC	F1	AUC	F1
CSI	0.581±0.03	0.547±0.02	0.612±0.03	0.598±0.02
URG	0.503±0.04	0.486±0.02	0.545±0.04	0.471±0.02
DEF	0.712±0.02	0.634±0.01	0.583±0.02	0.583±0.01
BBL	0.748±0.01	0.711±0.01	0.634±0.01	0.634±0.01
UDA	0.812±0.02	0.778±0.01	0.702±0.01	0.702±0.01
EMB	0.870±0.01	0.876±0.02	0.846±0.03	0.795±0.03
SRE	0.885±0.03	0.881±0.04	0.838±0.03	0.791±0.05
REAL-FND	**0.901**±0.02	**0.892**±0.03	**0.862**±0.05	**0.815**±0.04

Table 3: *Cross-domain:* Cross-domain fake news detection results on the target domain. source → target indicates the models have been trained on the source domain and tested on the target domain. In this case, we use the source domain data in addition to a small portion of the target domain data.

- **Adv-REAL-FND (ARE)**: To study the effect of the RL agent in domain adaptation, we use adversarial training to create a domain adaptive news article representation. In this model, we use the representation network, fake news classifier F, ARE domain classifier D in an adversarial setting using the following loss function to train a fake news classifier and representation network that is capable of creating domain independent features.

$$\min_F \max_D \mathcal{L}_{CE}(F) - \mathcal{L}_{CE}(D), \tag{5}$$



(a) Effect of the α parameter (b) Effect of the β parameter (c) Effect of the target domain portion γ

Figure 4: The impact of reward function parameters and the portion of the target domain data. Blue line shows the result of REAL-FND on *GossipCop*, while the red line shows the F1 for REAL-FND trained on the *Politifact* domain.

where L is the binary cross entropy loss similar to Equation 4. In this model we do not need to pre-train the fake news classifier F and the domain classifier D. Instead, we train the representation network, fake news classifier F, and domain classifier D as a whole using Equation 5.

- **Adv-REAL-FND \A**: In this model, we removed both RL component and the auxiliary information.

According to Figure 3, it is notable that removing the users' comments, and the user-news interactions severely impacts the performance of cross-domain fake news detection. Although using adversarial training for cross-domain fake news detection performs well, reinforcement learning allows us to adopt the news article representation using any pre-trained domain and fake news classifier in the reward function without considering its differentiability.

5.6 Parameter Analysis

Figure 4 shows the results of varying parameters in the reward function (i.e., Equation 1) and the portion of target domain data. We perform sensitivity analysis on the reward function parameters α, β by fine-tuning the parameters across the range $[0.0, 1.0]$. We also examine the effect of the γ parameter by changing it across the range $[0.0, 0.7]$. We evaluate the effect of varying these parameters by evaluating the F1 score of cross-domain fake news detection.

Figure 4a shows the F1 score for different values of α when $\beta = 0.5$. The α parameter controls the amount of the contribution of the fake news classifier in the reward function. This figure indicates that the performance does not increase when $\alpha \geq 0.6$ and even lead to a decrease in the performance if we use a higher value than β.

Figure 4b illustrates the results of varying β when $\alpha = 0.5$. By using $\beta = 0.0$, REAL-FND does not consider the feedback of the domain classifier, thus, the performance on the target domain will be dismal. By increasing β, RL penalizes the agent more when the agent cannot decrease the confidence of the domain classifier. Similar to the α parameter, the performance does not increase by a large margin by using $\beta \geq 0.6$.

To analyze the impact of γ, we use various portions of the target domain dataset to train REAL-FND. Intuitively, using all the data from target domain will improve the performance; however, in real-world scenarios, we may not have access to a comprehensive dataset from the target domain and using such dataset can lead to an expensive computation time. In real-world scenarios, domain-specific data is usually sparse and training on 100% of the target domain's data is impractical and unrealistic. Based on our analysis, using 30% of target domain dataset seems to be sufficient for training a reliable domain adaptive fake news detection model.

6 CONCLUSION AND FUTURE WORK

Collecting and integrating news articles from different domains and providing human annotations by fact-checking the contents for the purpose of aggregating a dataset are resource-intensive activities that hinder the effective training of automated fake news detection models. Although many deep learning models have been proposed for fake news detection and some have exhibited good results on the domain they were trained on, we show in this work that they have limited effectiveness in other domains. To overcome these challenges, we propose the **RE**inforced domain **A**daptation **L**earning for **F**ake **N**ews **D**etection (REAL-FND) task which could effectively classify fake news on two separate domains using only a small portion of the target domain data. Further, REAL-FND also leverages auxiliary information to enhance fake news detection performance. Experiments on real-world datasets show that in comparison to the current SOTA, REAL-FND adopts better to a new domain by using auxiliary information and reinforcement learning and achieves high performance in a single-domain setting.

Enhancements on REAL-FND could focus on cross-domain fake news detection under limited supervision to address the effects of limited or imprecise data annotations by applying weak supervision learning in a domain adaptation setting. Additionally, REAL-FND could benefit from methods that automatically defend against adversarial attacks. For example, malicious comments that promote a fake news article could degrade the performance of REAL-FND because of REAL-FND's contingency on users' comments as auxiliary information. Finally, we also propose addressing the inconsistencies in multi-modal information due to the breadth of the types and sources of information that are required for a cross-domain fake news classifier such as REAL-FND.

ACKNOWLEDGMENTS

This material is based upon work supported by ONR (N00014-21-1-4002), and the U.S. Department of Homeland Security under Grant Award Number 17STQAC00001-05-00[3]. Kai Shu is supported by the NSF award #2109316.

[3]Disclaimer: "The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security."

REFERENCES

[1] Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. 2010. A theory of learning from different domains. *Machine learning* 79, 1-2 (2010), 151–175.

[2] Published by Statista Research Department and Sep 7. 2021. Daily social media usage worldwide. https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/

[3] Lu Cheng, Ahmadreza Mosallanezhad, Yasin Silva, Deborah Hall, and Huan Liu. 2021. Mitigating bias in session-based cyberbullying detection: A non-compromising approach. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 2158–2168.

[4] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555* (2014).

[5] Alistair Coleman. 2020. 'hundreds dead' because of covid-19 misinformation. https://www.bbc.com/news/world-53755067

[6] William Fedus, Ian Goodfellow, and Andrew M Dai. 2018. MaskGAN: better text generation via filling in the_. *arXiv preprint arXiv:1801.07736* (2018).

[7] Song Feng, Ritwik Banerjee, and Yejin Choi. 2012. Syntactic stylometry for deception detection. In *ACL*.

[8] Chenggong Gong, Jianfei Yu, and Rui Xia. 2020. Unified Feature and Instance Based Domain Adaptation for End-to-End Aspect-based Sentiment Analysis. In *EMNLP*.

[9] Chuan Guo, Juan Cao, Xueyao Zhang, Kai Shu, and Huan Liu. 2019. DEAN: Learning Dual Emotion for Fake News Detection on Social Media. *arXiv preprint arXiv:1903.01728* (2019).

[10] Han Guo, Juan Cao, Yazi Zhang, Junbo Guo, and Jintao Li. 2018. Rumor detection with hierarchical social attention network. In *CIKM*.

[11] Seyedmehdi Hosseinimotlagh and Evangelos E Papalexakis. 2018. Unsupervised content-based identification of fake news articles with tensor decomposition ensembles. In *Proceedings of the Workshop on Misinformation and Misbehavior Mining on the Web (MIS2)*.

[12] Md Saiful Islam, Tonmoy Sarkar, Sazzad Hossain Khan, Abu-Hena Mostofa Kamal, SM Murshid Hasan, Alamgir Kabir, Dalia Yeasmin, Mohammad Ariful Islam, Kamal Ibne Amin Chowdhury, Kazi Selim Anwar, et al. 2020. COVID-19–related infodemic and its impact on public health: A global social media analysis. *The American journal of tropical medicine and hygiene* 103, 4 (2020), 1621.

[13] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. 1996. Reinforcement learning: A survey. *Journal of artificial intelligence research* 4 (1996), 237–285.

[14] Mansooreh Karami, Ahmadreza Mosallanezhad, Michelle V Mancenido, and Huan Liu. 2020. " Let's Eat Grandma": When Punctuation Matters in Sentence Representation for Sentiment Analysis. *arXiv preprint arXiv:2101.03029* (2020).

[15] Mansooreh Karami, Tahora H Nazer, and Huan Liu. 2021. Profiling Fake News Spreaders on Social Media through Psychological and Motivational Factors. In *Proceedings of the 32nd ACM Conference on Hypertext and Social Media*. 225–230.

[16] Hamid Karimi and Jiliang Tang. 2019. Learning hierarchical discourse-level structure for fake news detection. *arXiv preprint arXiv:1903.07389* (2019).

[17] Yoon Kim. 2014. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882* (2014).

[18] Wouter M Kouw, Laurens JP Van Der Maaten, Jesse H Krijthe, and Marco Loog. 2016. Feature-level domain adaptation. *The Journal of Machine Learning Research* (2016).

[19] Quoc Le and Tomas Mikolov. 2014. Distributed representations of sentences and documents. In *ICML*.

[20] Xinlong Li, Xingyu Fu, Guangluan Xu, Yang Yang, Jiuniu Wang, Li Jin, Qing Liu, and Tianyuan Xiang. 2020. Enhancing BERT representation with context-aware embedding for aspect-based sentiment analysis. *IEEE Access* 8 (2020), 46868–46876.

[21] Zichao Li, Xin Jiang, Lifeng Shang, and Hang Li. 2017. Paraphrase generation with deep reinforcement learning. *arXiv preprint arXiv:1711.00279* (2017).

[22] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692* (2019).

[23] Amy Mitchell and Mark Jurkowitz. 2020. Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable. https://www.journalism.org/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/

[24] Ahmadreza Mosallanezhad, Ghazaleh Beigi, and Huan Liu. 2019. Deep reinforcement learning-based text anonymization against private-attribute inference. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. 2360–2369.

[25] Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, and Rada Mihalcea. 2018. Automatic Detection of Fake News. In *Proceedings of the 27th International Conference on Computational Linguistics*. Association for Computational Linguistics, Santa Fe, New Mexico, USA, 3391–3401. https://aclanthology.org/C18-1287

[26] Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein. 2017. A stylometric inquiry into hyperpartisan and fake news. *arXiv preprint arXiv:1702.05638* (2017).

[27] Feng Qian, Chengyue Gong, Karishma Sharma, and Yan Liu. 2018. Neural User Response Generator: Fake News Detection with Collective User Intelligence.. In *IJCAI*.

[28] Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. 2017. Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 conference on empirical methods in natural language processing*. 2931–2937.

[29] Natali Ruchansky, Sungyong Seo, and Yan Liu. 2017. Csi: A hybrid deep model for fake news detection. In *CIKM*.

[30] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108* (2019).

[31] Tal Schuster, Roei Schuster, Darsh J. Shah, and Regina Barzilay. 2020. The Limitations of Stylometry for Detecting Machine-Generated Fake News. *Computational Linguistics* 46, 2 (06 2020), 499–510. https://doi.org/10.1162/coli_a_00380 arXiv:https://direct.mit.edu/coli/article-pdf/46/2/499/1847559/coli_a_00380.pdf

[32] Shaban Shabani and Maria Sokhn. 2018. Hybrid machine-crowd approach for fake news detection. In *2018 IEEE 4th International Conference on Collaboration and Internet Computing (CIC)*. IEEE, 299–306.

[33] Kai Shu, Amrita Bhattacharjee, Faisal Alatawi, Tahora H Nazer, Kaize Ding, Mansooreh Karami, and Huan Liu. 2020. Combating disinformation in a social media age. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10, 6 (2020), e1385.

[34] Kai Shu, Limeng Cui, Suhang Wang, Dongwon Lee, and Huan Liu. 2019. dEFEND: Explainable Fake News Detection. In *KDD*.

[35] Kai Shu, Deepak Mahudeswaran, Suhang Wang, Dongwon Lee, and Huan Liu. 2020. Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big data* 8, 3 (2020), 171–188.

[36] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter* 19, 1 (2017), 22–36.

[37] Kai Shu, Suhang Wang, and Huan Liu. 2019. Beyond news contents: The role of social context for fake news detection. In *WSDM*. ACM.

[38] Amila Silva, Yi Han, Ling Luo, Shanika Karunasekera, and Christopher Leckie. 2020. Embedding Partial Propagation Network for Fake News Early Detection.. In *CIKM (Workshops)*.

[39] Amila Silva, Ling Luo, Shanika Karunasekera, and Christopher Leckie. 2021. Embracing Domain Differences in Fake News: Cross-domain Fake News Detection using Multi-modal Data. *arXiv preprint arXiv:2102.06314* (2021).

[40] Eugenio Tacchini, Gabriele Ballarin, Marco L Della Vedova, Stefano Moret, and Luca de Alfaro. 2017. Some like it hoax: Automated fake news detection in social networks. *arXiv preprint arXiv:1704.07506* (2017).

[41] Hado Van Hasselt, Arthur Guez, and David Silver. 2016. Deep reinforcement learning with double q-learning. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 30.

[42] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *NeurIPS*.

[43] George-Alexandru Vlad, Mircea-Adrian Tanase, Cristian Onose, and Dumitru-Clementin Cercel. 2019. Sentence-Level Propaganda Detection in News Articles with Transfer Learning and BERT-BiLSTM-Capsule Model. In *Proceedings of the Second Workshop on NLP for Internet Freedom*. 148–154.

[44] Hu Xu, Bing Liu, Lei Shu, and Philip S Yu. 2019. BERT post-training for review reading comprehension and aspect-based sentiment analysis. *arXiv preprint arXiv:1904.02232* (2019).

[45] Junzi Zhang, Jongho Kim, Brendan O'Donoghue, and Stephen Boyd. 2020. Sample efficient reinforcement learning with REINFORCE. *arXiv preprint arXiv:2010.11364* (2020).

[46] Xinyi Zhou, Atishay Jain, Vir V Phoha, and Reza Zafarani. 2020. Fake news early detection: A theory-driven model. *Digital Threats: Research and Practice* 1, 2 (2020), 1–25.

[47] Fuzhen Zhuang, Xiaohu Cheng, Ping Luo, Sinno Jialin Pan, and Qing He. 2015. Supervised representation learning: Transfer learning with deep autoencoders. In *AAAI*.