RAYW1

Social-Content-Fusion Fake News Detection with Heterogeneous Graph Neural Networks

by

Shang-Ling HSU, Tianle LI, Yanjia LI, Yushi SUN

RAYW1

Advised by

Prof. Raymond Chi-Wing WONG

Submitted in partial fulfillment

of the requirements for COMP 4981

in the

Department of Computer Science and Engineering

The Hong Kong University of Science and Technology

2020-2021

Date of submission: February 12th, 2021

Contents

Contents **3**

1 Introduction **5**

1.1 Overview . . . 5

1.2 Objectives . . . 5

1.3 Challenges . . . 6

1.4 Literature Survey . . . 7

1.4.1 Content-based Methods . . . 7

1.4.2 Social Context based Methods . . . 12

2 Methodology **14**

2.1 Design . . . 14

2.1.1 Datasets . . . 14

2.1.2 Model Framework . . . 16

2.2 Implementation . . . 20

2.2.1 Data Partition . . . 20

2.2.2 Data Normalization . . . 20

2.2.3 Weight Initialization . . . 21

2.2.4 Batch Size . . . 21

2.2.5 Co-attention and attention mechanisms . . . 21

2.3 Experiment . . . 21

2.3.1 Evaluation Metrics . . . 22

2.3.2 Experiment Setup . . . 22

2.4 Upcoming tasks . . . 24

3 Project Planning **25**

3.1 Distribution of Work . . . 25

3.2 GANTT Chart . . . 25

4 Required Hardware & Software **25**

4.1 Hardware Requirements . . . 25

4.2 Software Requirements . . . 26

5 References **26**

6 Appendix A: Meeting Minutes **29**

6.1 Minutes of the 1^{st} Project Meeting . . . 29

6.2 Minutes of the 2^{nd} Project Meeting . . . 30

6.3 Minutes of the 3^{rd} Project Meeting (Internal) 31
6.4 Minutes of the 4^{th} Project Meeting 33
6.5 Minutes of the 5^{th} Project Meeting (Internal) 34
6.6 Minutes of the 6^{th} Project Meeting (Internal) 35
6.7 Minutes of the 7^{th} Project Meeting (Internal) 37
6.8 Minutes of the 8^{th} Project Meeting (Internal) 38
6.9 Minutes of the 9^{th} Project Meeting 39
6.10 Minutes of the 10^{th} Project Meeting (Internal) 40
6.11 Minutes of the 11^{th} Project Meeting (Internal) 41
6.12 Minutes of the 12^{th} Project Meeting (Internal) 42
6.13 Minutes of the 13^{th} Project Meeting (Internal) 43
6.14 Minutes of the 14^{th} Project Meeting (Internal) 44
6.15 Minutes of the 15^{th} Project Meeting (Internal) 44
6.16 Minutes of the 16^{th} Project Meeting 45
6.17 Minutes of the 17^{th} Project Meeting (Internal) 46
6.18 Minutes of the 18^{th} Project Meeting (Internal) 47
6.19 Minutes of the 19^{th} Project Meeting (Internal) 48
6.20 Minutes of the 20^{th} Project Meeting (Internal) 49

1 Introduction

1.1 Overview

In this era of advanced technology and high digital literacy, the cost of spreading information has been reduced drastically. The prevalence of mobile devices and the nearly ubiquitous access to the Internet facilitate much communication among human beings, breaking the geographic limitations that have existed for thousands of years. Due to the increasingly lower cost of information dissemination, almost everyone can access social media networks, in which they may forward news with a few simple clicks or taps. Nonetheless, it is both a blessing and a curse. A new challenge, which is the rapid dissemination of fake news, has emerged.

Whether malicious or inadvertent, the propagation of fake news has become faster than any other time in history, especially with AI and swarm computing to multiply it exponentially and target individuals based on their user profiles, thus appealing specifically to each person's emotions. Basically, fake news is digital gossip, slander and libel that can harm people's reputations, cause people to make uninformed decisions and even shape the collective consciousness of the masses and control what people believes is true and real. It's the ideal way for technocrats to perform social engineering. Therefore trying to understand, identify and mitigate fake news is a worthy task.

Fake news detection has become a hot research topic recently, especially on social networks. In previous studies, researchers have usually defined *fake news* as *intentional and verifiabe false news published by a news outlet* [5]. Specifically, a manipulated image, an untruthful claim, a mismatch between a picture and a caption, etc., may all contribute to a piece of fake news. Moreover, since social media is of paramount importance in the propagation of fake news, fake news detection *on social media* is the main focus of our research.

Some previous studies addressing fake news detection, were *content-based*, classifying the news by its text and images. Others have attempted to address the problem based on the *social media context*, utilizing the propagation patterns of the news on social media networks. However, as far as we know, none of them fused the information from both the social media context and content. Fortunately, a Heterogeneous Graph Neural Network (HetGNN) [34] was recently proposed and has yielded some promising results in some downstream applications, making encoding a heterogeneous graph with nodes fusing different information possible.

To this end, we propose a novel fake news detection model. We utilize both information from social media context and the news content to classify the news, based on the concept of HetGNN.

1.2 Objectives

The main goal of this project is to develop an approach for better fake news detection on social networks. We are mainly focusing on the following objectives:

1. **Propose a fake news detection model with deeply-fused information from both the news content and the social media context of the news.** To the best of our knowledge, this has not been done before. Adopting the concept of HetGNN [34], we propose a model that may achieve the deep fusion of these two kinds of information.

2. **Utilize HetGNN in fake news detection**. HetGNN is a novel model which succeeds in various tasks, but as far as we know, it has not yet been applied in fake news detection. As we see the benefits and potential of adopting it in our work, we introduce its usage in fake news detection.

3. **Evaluate the overall performance of the proposed model** with extensive experiments on large-scale public datasets. We are using the FakeNewsNet (Twitter) dataset [23] and the Weibo dataset [9], which are the most common datasets used in the literature reviewed.

4. **Compare the performance of the proposed model with the state-of-the-art baselines**. We selected four previous studies to be our baselines. Two of them are based on content, and the other two are based on social media context. For both content-based and social context based methods, we have chosen one influential work that has been cited and compared frequently in recent works, and one convincing, recent work that outperforms other state-of-the-art models.

5. **Conduct ablation studies on the proposed model**. To demonstrate the efficacy of each component in our design, we will examine each component thoroughly with ablation tests, showing the necessity of the component in the proposed model.

6. **Explore the potential of explainability of the model.** As the proposed model has the potential of being explainable, we are exploring and evaluating this potential qualitatively.

1.3 Challenges

We identified some challenges shared by most previous studies in fake news detection, and the lack of fusion and the explainability are the ones shared by most existing works. We hope to overcome these challenges in our model.

1. **Lack of Fusion.** The first general challenge is that existing works do not combine the information from both news posts content and the social media context information from the dissemination process of news well. We show in our Literature Survey that existing methods either focus on processing news' visual and textual content or analysing the user profile information and the dissemination patterns of news posts in a social media network. Our work aims at filling the gap in this field by proposing a novel model that utilizes both information from news content and the social media context.

2. **Explainability.** The second general challenge lies in the explainability. In pursuit of mere accuracy, most of the previous studies in fake news detection were not explainable. Nonetheless, explainability is essential in this field because the content of fake news, especially the malicious content, is often deliberately designed to deceive readers. Without explainability, the detection results seem unconvincing to readers and may turn out to be useless. To improve the practicality of our approach, we are utilizing an attention layer cross the heterogeneous network to capture the entities with more salient weights, making the proposed model more explainable.

3. **Adapted model.** In our project, we are adapting HetGNN[34] model to the fake news detection task. The challenge lies in the original visual and textual feature extractors used in HetGNN, which may not be efficient in capturing the fake news content features. We are utilizing the novel visual and textual feature extraction methods proposed by the existing models in the field of fake news detection to deal with this challenge. Besides, the definition of nodes and vertices in our fake news detection task is different from that in the original HetGNN model, which can lead to potential problems. We are overcoming this challenge by conducting experiments on different ways of defining the nodes and vertices, hoping to eventually come up with an appropriate definition of the components in the graphs for the fake news detection task.

1.4 Literature Survey

The dissemination of fake news has become a major problem in social media. The main challenge in the task of fake news detection is how to use multi-modal information properly to distinguish fake news from real news. Different kinds of fake news detection methods have been proposed by researchers nowadays. We classify them into two categories: content-based methods and social context based methods.

1.4.1 Content-based Methods

Content-based methods have been proven to be useful in fake news detection tasks by many previous studies [10][20][30]. In general, the content-based methods can be further classified into visual content-based, textual content-based and multi-modal content-based methods. The advantages of content-based methods are as follows: These methods directly process the news content and do not focus on the time-consuming analysis of social context information about the dissemination of fake news on social network. Besides, content-based methods transform the original fake news detection task into a classification task of news content. The classification models with given news content are mature and practical, is proven by previous studies. The disadvantages also exist: The content-based methods ignore the potential useful information from the social context. As pointed out by [7], social context information contains valuable information for fake news detection. Therefore, the ignorance of social context information limits the advancement of the content-based methods.

1. **Visual Content**

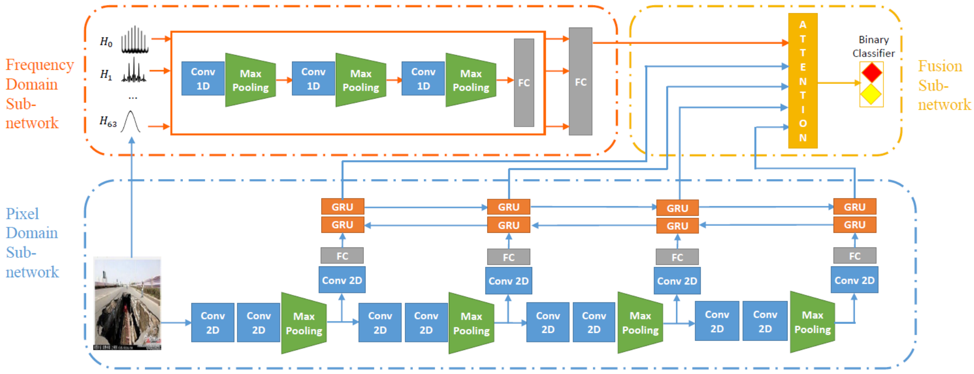

Figure 1: MVNN Architecture [20]

- Multi-domain Visual Neural Network (MVNN), proposed by [20], is a fake news detection neural network that utilizes visual content in posts to make judgments regarding truthfulness. In general, this model is easy to implement and use.

  The model takes images in a pixel domain as input, transforms the images into its frequency domain, and then takes the images in the pixel domain and frequency domain into their respective subnetworks, as shown in Figure 1. Then it fuses the outputs from the two subnetworks and uses an attention mechanism to judge on whether or not the posts are fake.

  The frequency domain subnetwork in MVNN aims at identifying heavy re-compression traces, such as block effects and the low image quality that usually occur in tampered images attached to fake news posts [20]. Thus, it makes use of CNN to collect these features in the frequency domain, and then it outputs the feature extracted to the attention layer.

  The pixel domain subnetwork is exploited to identify the visual impact [9] and emotional provocations [25][27] from the misleading images of fake news posts. A multi-branch CNN-RNN network is utilized to extract those emotional and visual features from low to high semantic levels [12].

  In the fuse subnetwork, features captured by the frequency domain subnetwork together with the features collected at different levels by the pixel domain subnetwork are fed into an attention layer [2] . The subnetwork then outputs the final decision to dynamically classify the fake news posts.

  MVNN has great contributions: it identifies the common visual features that appear in fake news posts and shows that these features are essential for the fake news detection task using visual content. Moreover, this model takes only images as input, which avoids the time-consuming processing of textual data and social context information. In addition, it is a novel model that outperforms existing state-of-the-art models at the moment.

However, its drawback is also obvious: MVNN cannot process posts without images attached. It also fails to fuse the visual and textual content of news posts, which may be particularly helpful for fake news detection proven by other models.

Thus, we decided to learn from MVNN to reinforce our own image feature extraction component, since MVNN has been proven to be useful in extracting image features from both pixel and frequency domains and we need this model to generate image embeddings for our model to replace the original CNN model used in HetGNN [34].

2. **Textual Content**

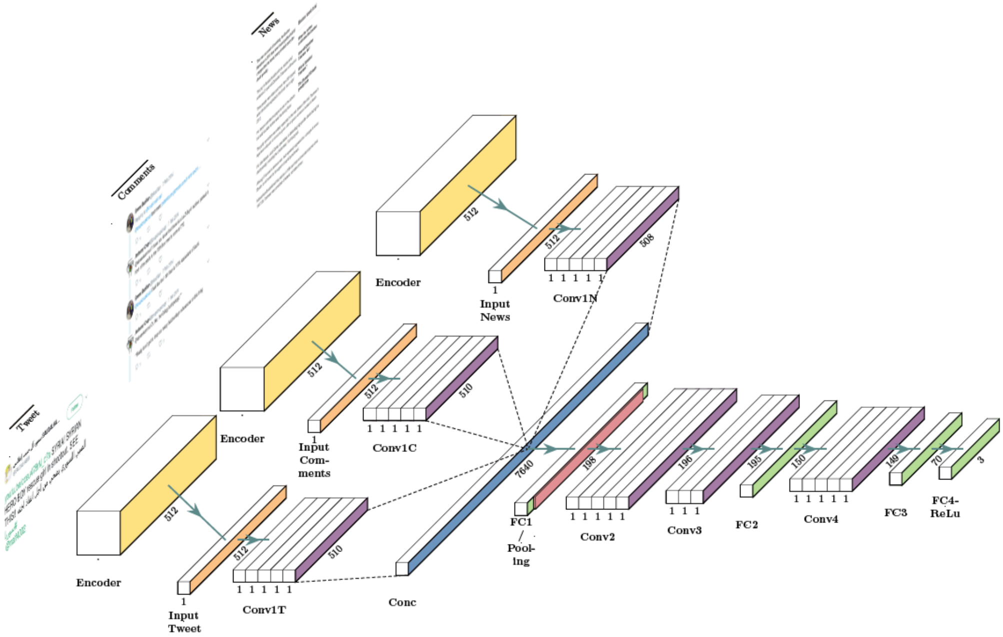

Figure 2: EMET Architecture [22]

- Embeddings from Multilingual-Encoder Transformer (EMET), proposed by [22], utilizes textual information from posts only to detect fake news. In general this model avoids the time-consuming analysis of social context and the accuracy of it is satisfactory.

  As shown in Figure 2, the network takes tweet contents, tweet comments and the fact-checked news searched by using key words from tweets as input and outputs the predicting labels about the news posts.

  The three sources of textual input in this network are encoded by encoders to form

512-dimension feature vectors respectively. Then, these three vectors are used for an 1D-convolution. The results are then concatenated to form a 7640-dimension feature vector. Then, a convolutional neural network is applied to this feature vector to achieve the final prediction result.

EMET's novel contribution is that it considers fact-checked news searched by using key words from tweets as input. This had not been previously proven to be useful by other studies. However, EMET's design also has its weakness. Since fact-checked news may not be available when fake news first emerges, this model may not be applicable at the initial stage of dissemination of fake news. Pointed out by [7], the initial stage is actually the crucial stage for the mitigation and prevention of fake news dissemination. Thus, this model may not have great practical use in reality.

Thus, we do not adopt this model due to the drawbacks mentioned previously. However, we are applying the word encoder structure in the word embedding generation component of our model.

3. **Multi-modal Content**

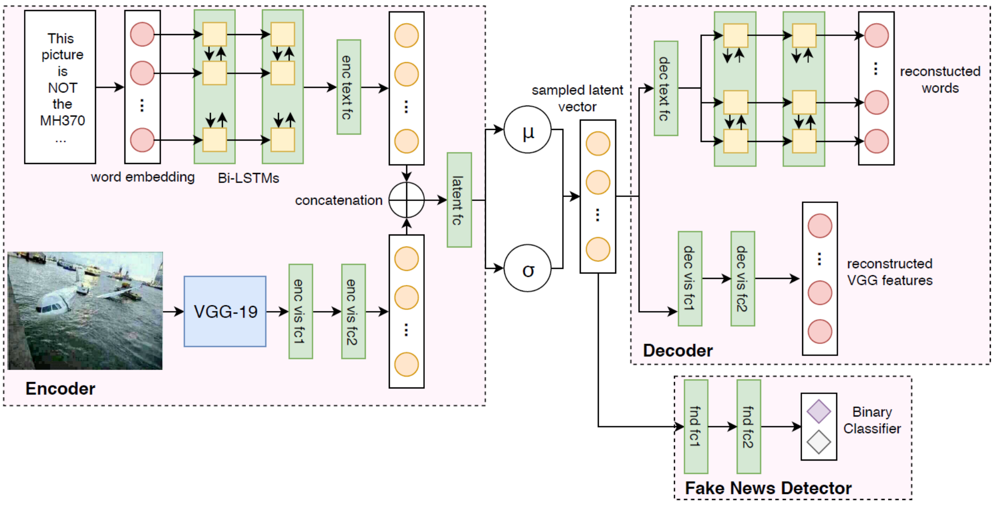

Figure 3: MVAE Architecture [10]

- Multimodal Variational AutoEncoder (MVAE), proposed by [10], takes both textual contents and the images attached to posts as input and outputs the predicting labels of fake news. In general, this model is a novel framework to combine the information from both textual and visual domains, which has been foundational to many later works.

  The textual encoder part of this model exploits a Long-Short Term Memory (LSTM) structure to encode the textual input as a feature vector, while the image encoder simply utilizes a VGG network to extract visual features. These two feature vectors are then concatenated and passed to the decoder and fake news detector. The decoder is used to decode the textual and visual contents so as to construct a reconstruction loss, which is common in the design of encoder-decoder structures for the supervision of this model. The fake news detector utilizes the fully-connected network to transform the concatenated vector into the binary representation of fake news labels.

The main contribution of MVAE is that it further improved the structural design of EANN model [30] by employing the encoder-decoder structure. The design of MVAE carefully fuses the textual and visual contents of fake news posts. It has also proven that both textual and visual contents are helpful for the task of fake news detection, which serves as theoretical support for our method. A possible improvement is to enable the usage of social context information, such as user characteristics as input.

Thus, we decided to utilize the word embedding generation model used in this work, since transforming words into vectors is its essential function, and it has been proven to be useful and efficient.

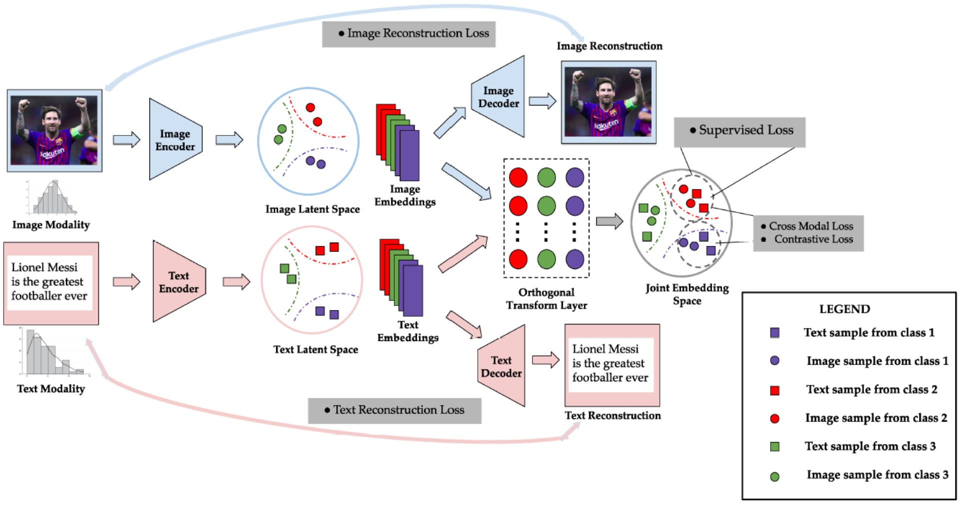

Figure 4: COBRA Architecture [28]

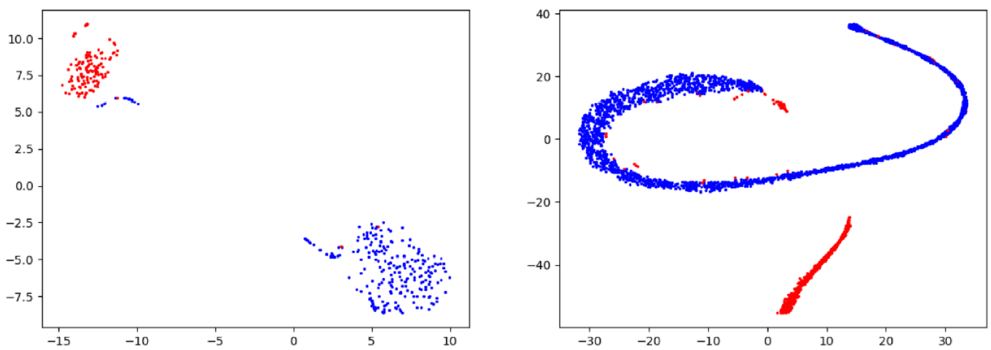

Figure 5: t-SNE transformed data points for fake news (red) and real news (blue) in COBRA [28]

- COnstractive Bi-modal Representation Algorithm (COBRA), proposed by [28], is a cross-modal embedding generation method. The authors have further extended this method to fit the fake news detection task. The goal of COBRA is to transform multi-modal information, including textual and visual contents into a common manifold, such that the representations between different classes are invariant among modalities and discriminatory.

The encoder-decoder structure is applied to both textual and visual content to calculate the reconstruction loss for different modalities. The encoded feature vectors are then passed through an orthogonal transform layer to gain the final joint embedding representations for each class. t-SNE transformation [29] as shown in Figure 5, is applied on the final joint embedding space to transform the high-dimensional embedding into 3-dimensional space, and then a binary classifier can be applied to distinguish fake news data points from the real news data points.

COBRA is a novel multi-modal framework that is used to transform textual and visual contents into a common space, while preserving modality invariance and class discrimination. It outperforms the current state-of-the-art content-based models on fake news detection task.

We are employing the idea of COBRA to reinforce our design. Since it has been proven to be particularly useful to project the input contents onto a latent space with careful design of contrastive loss and cross-modality loss to facilitate the modality invariant and class discriminatory requirements, we believe that it can also improve the original design of textual and visual content processing in HetGNN model [34].

1.4.2 Social Context based Methods

Apart from content-based methods, the importance of social context analysis has been shown by many recent works [7][15][24]. Instead of analysing the news post content to do prediction, these methods identify the fake news through their dissemination patterns. Social context information analysis can sometimes be more accurate in comparison to the content-based methods. However, the drawbacks are also obvious: the social context based methods always have complex structures, and the training process can be very time-consuming. Social context based methods in general require more detailed data about the dissemination process of news post. Thus, dataset collection can also be a potential obstacle for the development of these methods.

1. **Social Context**

   - Graph-aware Co-Attention Networks (GCAN), proposed by [15], focuses on tweet contents, users' profiles as well as the sequence of retweets to judge whether news posts are fake or not. It pays little attention to the tweet content.

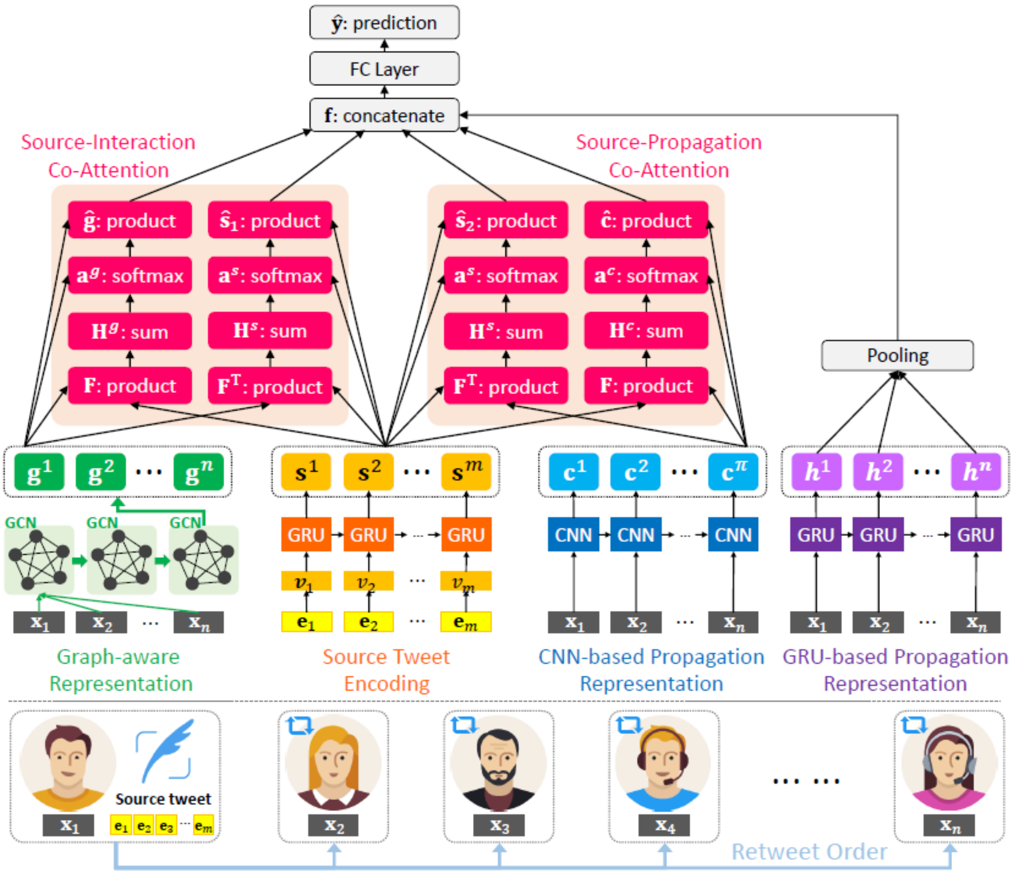

Figure 6: GCAN Architecture [15]

GCAN generally contains five components: 1) a user characteristics extractor, which extracts features on the users' participation patterns on a social network and encodes them into graphs; 2) a source tweet encoder, which encodes the news posts into vector representations; 3) user propagation representation networks, which utilizes the retweet sequences and the extracted users' characteristics to model and represent the propagation pattern of news posts. Both CNN-based and GRU-based propagation representations are employed; 4) a co-attention mechanism, which utilizes the attention mechanism to capture the correlation between user representation, encoded news contents and propagation representations; 5) a prediction layer, which concatenates the outputs from the other four components and outputs the prediction result.

As shown in Figure 6, users' profile information is encoded by the user characteristics extractor, while the news content is encoded using a source tweet encoder, and the retweeting sequences are processed with propagation representation networks (which are CNN-based and GRU-based). These encoded features except the GRU-based propagation representation is passed to the co-attention layer. Then the co-attention layer output together with the GRU-based propagation representation are passed to the final prediction layer.

The key benefit of GCAN is that it does not require users' comments, which are heterogeneous to some extent. Besides, this work outperforms current state-of-the-art social context based models by a large extent. Moreover, this work has great explainability thanks to the co-attention mechanism employed. Future improvement for this work can potentially be in the news content encoding part, which could allow multi-modal content, as used in MVAE [10] and COBRA [28], to improve the source tweet encoding quality.

We are using the user propagation representation mechanism employed in this model in our own design, since the original HetGNN [34] design does not pay great attention to the time sequence between the edges and this work actually proved this to be helpful for fake news detection task.

2 Methodology

2.1 Design

2.1.1 Datasets

To provide a fair evaluation on our proposed model, we have three main criteria for dataset selection:

- If the news in the dataset include multimedia information and their propagation paths as required by the heterogeneous graph neural network[34].
- If the news in the dataset are all verified by official and reputable platforms or committees. The verification given by the authorities promises the objective ground-truth labeling of the news.
- If the dataset contains news covering a various range of topics and if the news are collected from different backgrounds(*e.g.*, communities, languages). The diversity of the news avoids the model from over-fitting specific events as well as tests the applicability of the model.

Based on the above standards, we found two qualified publicly-available datasets. They are *FakeNewsNet*[23] and *Weibo dataset*[16], which were released by [23] and [16], respectively. Besides the text contents as most existing datasets have, both of them also contains visual contents and social context information. Visual contents are related to attached images or user avatars, and social context includes the source of the contents and the propagation patterns (e.g., sequence of retweet users). Below are the detailed introduction of the two datasets.

- ***FakeNewsNet***

  FakeNewsNet[23] is a comprehensive collection of human-labeled short statements crawled from two well-known fact-checking platforms: GossipCop and PolitiFact. GossipCop investigates celebrity-reporting while PolitiFact mainly focuses on political news in the U.S. Both of them are non-profit voluntary project launched by the journalists and editors working in some credible news agencies. The news included in FakeNewsNet are those widely spread in English-speaking communities, especially in the U.S. society, and are mainly about entertainment stories and politics-involved statements. The detailed statistics of the dataset are shown in Table 1.

Table 1: Statistics of FakeNewsNet

Platform	**PolitiFact**	**GossipCop**
# fake news	432	5,323
# true news	624	16,817
# fake news with propagation network	351	3,684
# true news with propagation network	277	6,945
# tweets	275,058	1,058,330
# retweets	293,438	530,833
# replies	125,654	530,833
# users	384,813	739,166

- ***Weibo Dataset*** This dataset is also used in [16] to detect the fake news on Weibo, one of the largest social platform in China. The dataset contains posts covering a wide range of topics, from education to public hygiene, which were published on Weibo by official/personal accounts, in the period from 2012 to 2016. The fake news in the dataset are all verified by the official rumor debunking system on Weibo and are manually checked by a committee consisting of users with high reputation. The detailed statistics about this dataset are shown in Table 2.

Table 2: Statistics of Weibo Dataset

# fake news	2,313
# true news	2,351
# users	3,719,681
# retweets	3,757,123
# images	3,851

2.1.2 Model Framework

Inspired by multiple related literature [10][30][31] based on multi-modal information, we devised a framework to exploit heterogeneous neural network to deeply fuse images, texts and social network components. We considered both global interpretation and local comparison as it is still unresolved how disparate features extracted from diverse modals can be semantically combined and utilized in a more intuitive and cunning way.

1. Input Components

   - Tweet text
   - Tweet images
   - Tweet status & user profile
   - Social network relationship

   These four major components are the inputs of our model. The first three are based on the content of posts in social media like Twitter and Weibo, while the last one indicates the relationship between users within both forwarding and followers.

2. Feature Encoding

   - **Tweet Text**: We are utilizing Par2Vec [19] to process the text contents. In the meantime, a powerful BERT pre-trained method, RoBERTa [13] is implemented as a strengthened and cogent representation of texts which has been verified its superiority among copious down stream tasks in NLP field. To boost the interpretability of our model, we also encode the text word by word to make preparation for the implementation of co-attention [14] with visual features in the future.

   - **Tweet Images**: We are adopting ResNet [6] to encode images rather than computation-intensive VGG16 [26] used in [34] expecting a more efficient implementation as well as a better performance as displayed in [1]. We transform the encoded visual features into 1D tensors to feed into the HetGNN [34]. Meanwhile, we also transform them into 2D tensors to interact with co-attention [14] mechanism for the explanability of our model in the future.

- **Tweets status & User profile**: There are a variety of miscellaneous information we could gain from the datasets [23] [16] about both the users themselves and the posts status. We are utilizing the same encoding method as [15] to represent the discrete and numerical features such as whether the users have been verified, the number of their followers and likes received by their tweets. In the meantime, profile images and descriptions as visual and textual features are encoded with cutting-edge methods as mentioned above.

- **Social Network relationship**: We construct a social network graph with heterogeneous node types as users and posts. If a user forwards a post, the user node will be linked to the original post node. And all the users, who have forwarded the same post, are fully connected with each other. Specifically, there are two types of edges in our heterogeneous graph, which are user-to-user and user-to-post. The information within the social network can be propagated through these links.

3. HetGNN

With diverse kinds of attributes within the same entity and different types of entities, HetGNN [34], a graph neural network resolving both graph configuration heterogeneity and node content heterogeneity, used in the first place to play a primary role of aggregating various features with supplementary information extracted from feature encoding as well as the corresponding knowledge graph [8] compared to the original one. In addition, we attempt to embed the social context information into our graph neural network to deeply fuse user relationship with content-based features as graph could be treated as a social network simultaneously. The overview of our model framework is displayed in Figure 7.

- **Heterogeneous Graph**

  As shown in Figure 8, we extract two elementary types of nodes from the datasets [23]

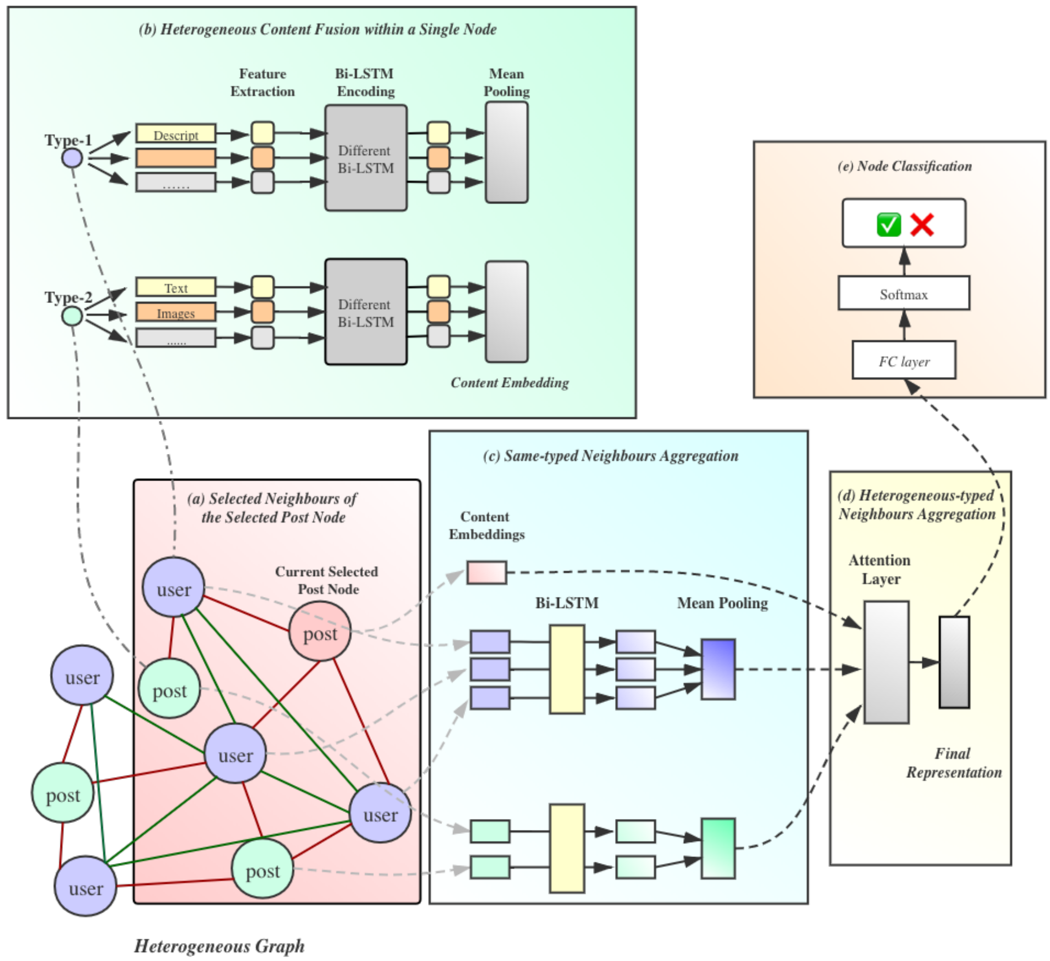

Figure 7: **(a) Sample neighbors of a specific node; (b) Aggregate the heterogeneous contents within s single node and figure out the similarity score between same-type nodes which would be utilized in (a) for sampling; (c) Aggregate same-type neighbors of the targeted node; (d) Aggregate heterogeneous-type neighbor representations obtained in (c) for the targeted node; (e) Connect the final representation to a fully connected layer and a Softmax function to get the prediction result.**

[16]: user and tweet. We expect to see a rationale fusion and aggregation within both a single node having miscellaneous features like images, texts and numerical counts and a neighbor group with distinctive type of nodes upon building subtle relationship or edges between either heterogeneous or homogeneous type of nodes. For instance, for the homogeneous type of nodes, users can be connected as long as they forward the same news. And heterogeneous type of nodes can be directed spontaneously with the relationship that a user has released a tweet.

- **Sample Heterogeneous Neighbors**

  To obtain neighbor candidates of post nodes for a fixed size, the strategy of random

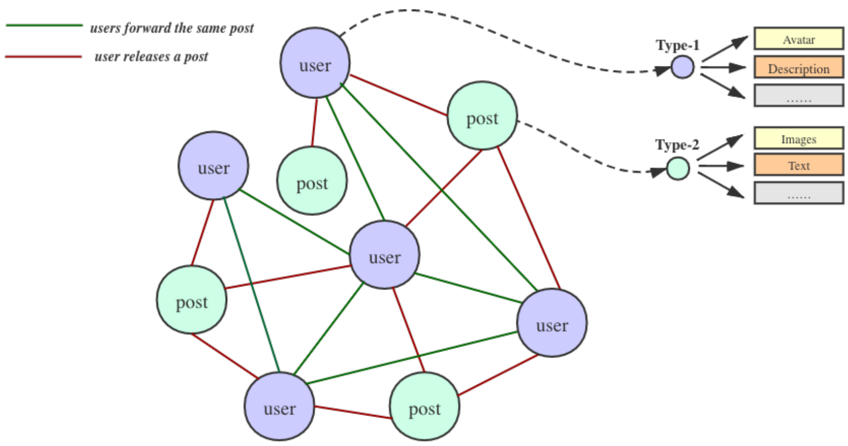

Figure 8: Heterogeneous Graph

walk with stationary length and homogeneous neighbor grouping would be exploited as mentioned in [34]. We set the size of post neighbours and user neighbours to be five and twenty respectively.

- **Encode Heterogeneous Contents within a Node**
  - Transform heterogeneous content features by fully connected layer.
  - Capture distinctive features within a post node or user node by Bi-LSTM (HetGNN model) or attention mechanism (HetGNN-att model).
  - Obtain a general input embedding of a node by mean pooling.
- **Aggregate Heterogeneous Neighbors**
  - Employ Bi-LSTM or attention mechanism to aggregate all the nodes with same type within the sampled neighbor group
  - Exploit the attention mechanism to aggregate heterogeneous types of nodes in the neighbor group with dexterously attuned weights.

Upon the detailed implementation of the redesigned HetGNN, a node classification problem can be formulated to judge whether a tweet-type node is fake or not after training so as to achieve our goal.

4. **Explainability**

   As employed in GCAN [15], co-attention mechanism [14] can attend to heterogeneous features and indicate the outstanding parts of features that play a more paramount role in making the final predictions. Therefore, to boost the explanability of our model, we utilize the co-attention mechanism to supervise the contributing weights within the encoded images-to-text features, text-to-user-neighbour features and text-to-post-neighbour features. We expect it can identify the specific words in the text, the vital objects in the images and the pivotal user or post neighbours that exert most crucial impact on the predictions.

2.2 Implementation

Following the structure described in 7, we implemented the structure of HetGNN. This section will introduce the special settings we used in our implementation.

2.2.1 Data Partition

Following the setting in [16], each dataset has been randomly split into the training set, the validation set and the testing set approximately with a ratio of 8:1:1. The proportion of the fake news in each sub-set is approximately the same as that in the whole dataset. Furthermore, posts about similar events have been carefully grouped into three sets in order to maintain a balanced distribution of contents between the training set, validation set and the testing set. Otherwise, the model may over-fit certain topics frequently occur in the training set.

2.2.2 Data Normalization

The progress was not as smooth as we expected at the beginning: Although we implemented the complete structure correctly as described in the Design part and carried out the experiments on Weibo dataset, the loss was stuck at around 0.693 (ln2) and the accuracy is fluctuate at around 50%. This implied that the model was randomly guessing the answer. A thorough examination was carried out and we found that the reason was that the values of the features varied a lot between textual and visual inputs. Aware of this, we normalized the textual and visual features collected before putting them into the network and the problem was solved.

2.2.3 Weight Initialization

The training process was occasionally influenced by the weight initialization. Originally, we did not consider the weight initialization and just use the default random values to construct the model, which led to large variance between different training trials, causing great challenges for parameters finetuning. In order to achieve a more stable performance of our model, we decided to adopt Xavier weight initialization [4] method in our model. The result turned out to be satisfactory, the model became more stable and the performance was more consistent between different training trials.

2.2.4 Batch Size

Our initial design was an online setting, which means the batch size was set to 1. However, during experiments, we found that the model would easily overfit after around 5 epochs even if dropout layers were added. We analyzed the reasons and suspected that this kind of overfitting might be the result of our initial online setting. Since if the batch size is set to 1, then the weights might be influenced a lot by the inherent noise in the dataset. Thus, we refined our code to enable batch size larger than 1 and this actually resulted in a more stable training process.

2.2.5 Co-attention and attention mechanisms

Inspired by GCAN [15], we decided to modify the original coattention mechanism used in the heterogeneous neighbor aggregation part to improve the explanability of our model. The implementation is completed and the on-going experiments are carrying out currently.

Apart from this, we also replaced Bi-RNN network used to aggregate neighbours with attention layers to compare the model performance. The experiment results will be discussed in 2.3.2. We named the model with HetGNN structure and attention mechanism as HetGNN-att model.

2.3 Experiment

Extensive experiments are conducted with view to giving answers to following questions:

Q1: How does our proposed model perform *vs.* state-of-the-art baselines for fake news detection, such as MVAE [10], SVM-RBF [32], DTC [3], RFC [11], SVM-TS [18], GRU-2 [17], CAMI [33] and CSI [21]?

Q2: How does our model perform with use of Bi-LSTM (i.e. HetGNN model), compared with the use of attention mechanism (i.e. HetGNN-att model)?

2.3.1 Evaluation Metrics

We use two metrics, Accuracy score and F1 score, to evaluate the performance of our HetGNN model and other baseline models. The accuracy score is the ratio between number of correctly predicted events and total number of events in the test set. The F1 score measures the balance between the precision and the recall.

2.3.2 Experiment Setup

We trained our models (using either par2vec or RoBERTa word embedding) on Weibo dataset, which was partitioned into train, validation and test sets with ration 8:1:1. Each model was trained for 40 epochs with a patience of 5. We selected the best models based on the validation loss and obtained the metrics scores on their performance on test set.

- **Model comparison**

  To answer **[Q1]** and **Q2**, we plan to run our proposed model as well as other baseline models on the *Weibo dataset*[16] and *FakeNewsNet*[23] and compare their performances using accuracy score and F1 score. The experiment results are shown in the table below:

Dataset	Method	Accuracy	Fake News			Real News		
			Precision	Recall	F1	Precision	Recall	F1
Weibo	MVAE	0.769	0.784	0.749	0.766	0.755	0.790	0.772
	SVM-RBF	0.818	0.822	0.812	0.817	0.815	0.824	0.819
	DTC	0.831	0.847	0.815	0.831	0.815	0.847	0.830
	RFC	0.849	0.786	0.959	0.864	0.947	0.739	0.830
	SVM-TS	0.857	0.878	0.830	0.857	0.947	0.739	0.830
	GRU-2	0.910	0.876	0.956	0.914	0.952	0.864	0.906
	CAMI	0.933	0.921	0.945	0.933	0.945	0.921	0.932
	CSI	0.932	0.938	0.924	0.931	0.926	0.940	0.933
	HetGNN-att	0.929	0.906	0.962	0.933	0.957	0.894	0.925
	HetGNN(RoBERTa)	0.944	0.928	0.967	0.947	0.963	0.920	0.941
	HetGNN(par2vec)	**0.951**	0.925	0.983	0.953	0.981	0.916	0.947

Table 3: Experiment Results on Weibo

- **Result Analysis**

  We tested the HetGNN model with two different word embeddings: RoBERTa and par2vec. After tuning the parameters subtly, we achieved test accuracy as 0.951 on HetGNN with par2vec word embedding, which is higher than all the other baselines in Table 3. The trend of loss and training & validation accuracies for both HetGNN (RoBERTa) and HetGNN (par2vec) have been recorded as Figure 9, Figure 10, Figure 11 and Figure 12.

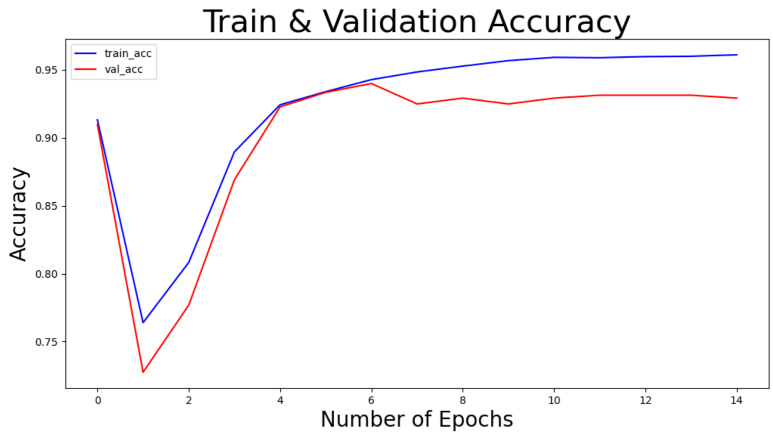

Figure 9: HetGNN (RoBERTa) Train & Validation Accuracy Trend

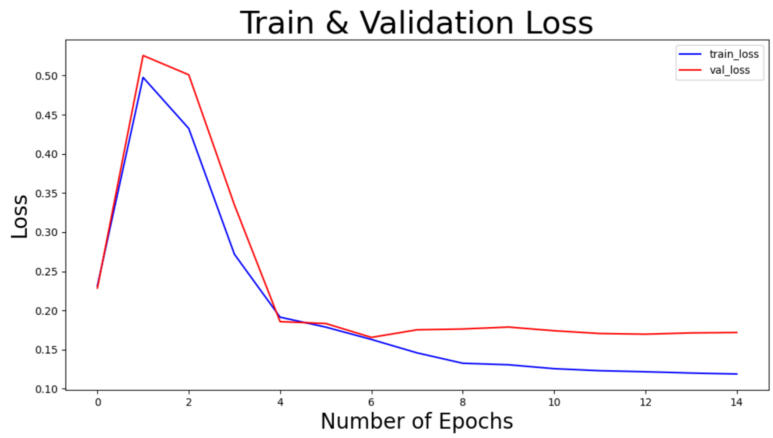

Figure 10: HetGNN (RoBERTa) Train & Validation Loss Trend

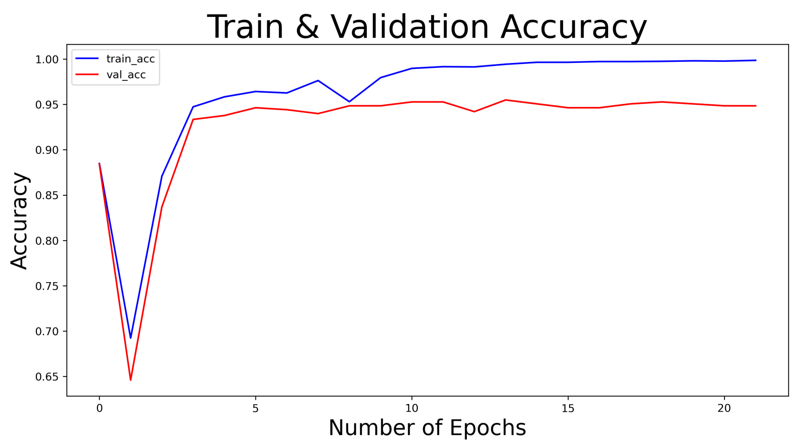

Figure 11: HetGNN (par2vec) Train & Validation Accuracy Trend

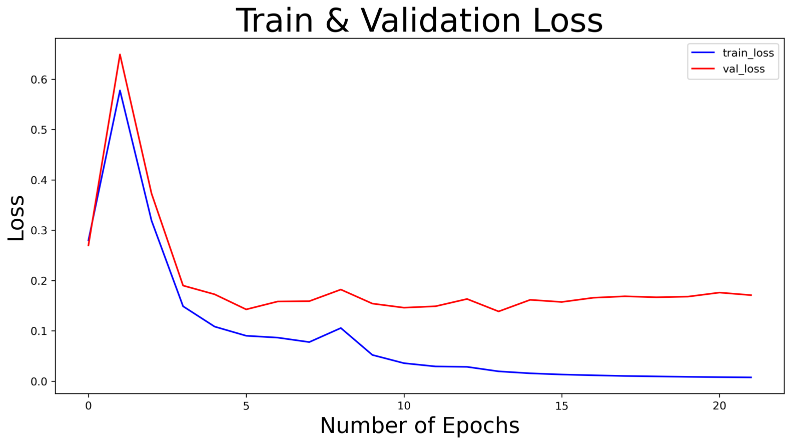

Figure 12: HetGNN (par2vec) Train & Validation Loss Trend

2.4 Upcoming tasks

1. Further conduct literature reviews about applications of object captioning and knowledge graph.
2. Utilize the newly computed post-post similarity score to refine the random walk neighbors generation process.
3. Construct the Heterogeneous Graph Network model and continue the undertake experiments on *FakeNewsNet*[23] and *Weibo dataset*[16].
4. Evaluate and improve results of the experiments.
5. Make comparison with other state-of-the-art baseline models and analyze the effectiveness of each model components through the ablation study.

3 Project Planning

3.1 Distribution of Work

Table 4: Distribution of Work

Task	**Shang-Ling**	**Tianle**	**Yanjia**	**Yushi**
Do the literature survey	○	○	○	●
Formulate objectives	●	○	○	○
Design the model	○	●	○	○
Implement the model	○	○	●	○
Evaluate and improve	○	○	○	●
Evaluate the model on large datasets	●	○	○	○
Compare our model with the baseline models	○	●	○	○
Release the model	○	○	●	○
Write the final paper	○	○	○	●
Produce the project video	●	○	○	○

3.2 GANTT Chart

Table 5: GANTT Chart

Task	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Jan**	**Feb**	**Mar**	**Apr**	**May**
Do the literature survey	■	■	■	■							
Formulate objectives		■	■								
Design the model		■	■	■	■	■	■	■			
Implement the model			■	■	■	■	■	■			
Evaluate and improve			■	■	■	■	■	■	■		
Evaluate the model on large datasets				■	■	■	■	■	■	■	
Compare our model with the baseline models					■	■	■	■	■	■	
Write the final paper									■	■	■
Release the model										■	
Produce the project video										■	■

4 Required Hardware & Software

4.1 Hardware Requirements

Table 6: Hardware Requirements

Item	**Specification**
Development PC	MacBook Pro (Retina, 15-inch, Mid 2015)
Server workstation	ASUS WS X299 SAGE CEB
Server CPU	Intel Core i9-10920X (12 cores, 24 thres, 3.5GHz)
Server GPU	ASUS RTX2080Ti Turbo 11GD6
Server RAM	32GB

4.2 Software Requirements

Table 7: Software Requirements

Item	Version	Specification
Development OS	MacOS Catalina 10.15.6	Environment for development
Server OS	CentOS linux	Environment for evaluation
Git	2.23.0 or after	Version control
Miniconda	4.8.4 or after	Package control
OpenSSL	1.0.2k-fips	Secure communications over networks
Python	3.6.10 or after	Programming language
Pytorch	1.6.0 or after	Machine learning library
tcsh	6.18.01 or after	Linux command shell

5 References

[1] V. Atliha and D. Šešok, "Comparison of vgg and resnet used as encoders for image captioning," in *2020 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream)*, 2020, pp. 1–4.

[2] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, 2014.

[3] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *Proceedings of the 20th international conference on World wide web*, 2011, pp. 675–684.

[4] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2010, pp. 249–256.

[5] Y. Han, S. Karunasekera, and C. Leckie, "Graph neural networks with continual learning for fake news detection from social media," *arXiv preprint arXiv:2007.03316*, 2020.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[7] M. R. Islam, "Detecting and mitigating rumors in social media," Ph.D. dissertation, Virginia Tech, 2020.

[8] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: Representation, acquisition and applications," *arXiv preprint arXiv:2002.00388*, 2020.

[9] Z. Jin, J. Cao, J. Luo, and Y. Zhang, "Image credibility analysis with effective domain transferred deep networks," *arXiv preprint arXiv:1611.05328*, 2016.

[10] D. Khattar, J. S. Goud, M. Gupta, and V. Varma, "Mvae: Multimodal variational autoencoder for fake news detection," in *The World Wide Web Conference*, 2019, pp. 2915–2921.

[11] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang, "Prominent features of rumor propagation in online social media," in *2013 IEEE 13th international conference on data mining.* IEEE, 2013, pp. 1103–1108.

[12] P. J. Lang, "A bio-informational theory of emotional imagery," *Psychophysiology*, vol. 16, no. 6, pp. 495–512, 1979.

[13] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "Roberta: A robustly optimized bert pretraining approach," *ArXiv*, vol. abs/1907.11692, 2019.

[14] J. Lu, J. Yang, D. Batra, and D. Parikh, "Hierarchical question-image co-attention for visual question answering," *CoRR*, vol. abs/1606.00061, 2016. [Online]. Available: http://arxiv.org/abs/1606.00061

[15] Y.-J. Lu and C.-T. Li, "Gcan: Graph-aware co-attention networks for explainable fake news detection on social media," *arXiv preprint arXiv:2004.11648*, 2020.

[16] J. MA, W. GAO, P. MITRA, S. KWON, B. J. JANSEN, K.-F. WONG, and M. CHA, "Detecting rumors from microblogs with recurrent neural networks," in *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, 2016, pp. 3818–3824.

[17] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, and M. Cha, "Detecting rumors from microblogs with recurrent neural networks," 2016.

[18] J. Ma, W. Gao, Z. Wei, Y. Lu, and K.-F. Wong, "Detect rumors using time series of social context information on microblogging websites," in *Proceedings of the 24th ACM international on conference on information and knowledge management*, 2015, pp. 1751–1754.

[19] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 3111–3119. [Online]. Available: http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf

[20] P. Qi, J. Cao, T. Yang, J. Guo, and J. Li, "Exploiting multi-domain visual information for fake news detection," in *2019 IEEE International Conference on Data Mining (ICDM).* IEEE, 2019, pp. 518–527.

[21] N. Ruchansky, S. Seo, and Y. Liu, "Csi: A hybrid deep model for fake news detection," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2017, pp. 797–806.

[22] S. Schwarz, A. Theóphilo, and A. Rocha, "Emet: Embeddings from multilingual-encoder transformer for fake news detection," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 2777–2781.

[23] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media," *arXiv preprint arXiv:1809.01286*, 2018.

[24] K. Shu, D. Mahudeswaran, S. Wang, and H. Liu, "Hierarchical propagation networks for fake news detection: Investigation and exploitation," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 14, 2020, pp. 626–637.

[25] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD explorations newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[26] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014. [Online]. Available: http://arxiv.org/abs/1409.1556

[27] C. R. Sunstein, *On rumors: How falsehoods spread, why we believe them, and what can be done*. Princeton University Press, 2014.

[28] V. Udandarao, A. Maiti, D. Srivatsav, S. R. Vyalla, Y. Yin, and R. R. Shah, "Cobra: Contrastive bi-modal representation algorithm," *arXiv preprint arXiv:2005.03687*, 2020.

[29] L. van der Maaten and G. Hinton, "Viualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, pp. 2579–2605, 11 2008.

[30] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, and J. Gao, "Eann: Event adversarial neural networks for multi-modal fake news detection," in *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining*, 2018, pp. 849–857.

[31] Y. Wang, S. Qian, J. Hu, Q. Fang, and C. Xu, "Fake news detection via knowledge-driven multimodal graph convolutional networks," in *Proceedings of the 2020 International Conference on Multimedia Retrieval*, ser. ICMR '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 540–547. [Online]. Available: https://doi.org/10.1145/3372278.3390713

[32] F. Yang, Y. Liu, X. Yu, and M. Yang, "Automatic detection of rumor on sina weibo," in *Proceedings of the ACM SIGKDD workshop on mining data semantics*, 2012, pp. 1–7.

[33] F. Yu, Q. Liu, S. Wu, L. Wang, T. Tan *et al.*, "A convolutional approach for misinformation identification." in *IJCAI*, 2017, pp. 3901–3907.

[34] C. Zhang, D. Song, C. Huang, A. Swami, and N. V. Chawla, "Heterogeneous graph neural network," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 793–803.

6 Appendix A: Meeting Minutes

6.1 Minutes of the 1^{st} Project Meeting

1. Arrangement
   - Date: April 26, 2020
   - Time: 4:00pm
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li,
     Prof. Raymond Chi-Wing Wong
   - Absent: None
   - Recorder: Shang-Ling Hsu, Yanjia Li
2. Approval of minutes
   - This was the first formal group meeting, so there were no minutes to approve.
3. Report on progress
   - All team members have read the instructions of the project listed in the CSE FYP system.
   - All team members have studied some related works regarding the topic.
4. Discussion items
   - There are many different definitions for fake news. We may define one on our own if we wish to. Some definitions include:
     - Images edited by people
     - The mismatch between the title and the content of a piece of news
     - An logically-misleading argument that describes only part of a fact
     - The credibility of news sources
   - Some possible research directions for us:
     - Propose a new method solving fake news detection from a new perspective

  - Evaluate existing works concerning fake news detection, find the most accurate one, and compare the pros and cons of the previous studies (The downside of this: no innovation is involved. It is only a comparison of the performance of the previous studies.)
  - Implement a system, detecting potential fake news and warning users
- Build knowledge graphs from text, and compare different graphs describing the same events
- Graph Neural Networks may also be used.

5. Goals for the coming week
   - Read three related works cherry-picked by Prof. Wong.
6. Meeting adjournment and next meeting
   - The meeting was on time.
   - The next meeting will be at 5pm on August 5, 2020 via Zoom.

6.2 Minutes of the 2nd Project Meeting

1. Arrangement
   - Date: August 5, 2020
   - Time: 5:00pm
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun,
     Prof. Raymond Chi-Wing Wong
   - Absent: None
   - Recorder: Shang-Ling Hsu, Yanjia Li
2. Approval of minutes
   - The minutes of last meeting were approved without amendment.
3. Report on progress
   - All team members have read the instructions of the Final Year Project online.
   - All team members have studied the 3 related works given by Prof. Wong in advance.
4. Discussion items
   - Confirm a new member: Yushi Sun.

- We may choose from the following three possible directions for this project:
  (a) Develop a fake news detection system (more application-oriented)
  (b) Do a survey for papers published in the last 5-10 years
  (c) Develop a new model for fake news detection (more research-oriented)
    - Define a new problem
    - Propose a better solution for an existing problem, addressing the limitations of the current methods
- We have to finish our proposal by mid-September, which takes 5% of the total FYP score. A sample proposal is provided in the CSE FYP system, but the requirement, which is 8-10 pages, is too short for the proposal. Thus, we may write more if we want to.
- Find a feasible dataset and describe it in the proposal. It is acceptable if we have not found all datasets exhaustively by the time we submit our proposal.
- A standard procedure of doing research:
  (a) Study the related works
  (b) Find the limitations of the related works
  (c) Proposal a new model to address the limitations

5. Goals for the coming week
   - Read the papers which cite the three papers provided by Prof. Wong and compare their performance.
   - Do a presentation summarizing the related works in the next meeting.
6. Meeting adjournment and next meeting
   - The meeting was on time.
   - The next meeting will be internal (without Prof. Wong) at 7pm on August 14, 2020 via WeChat.

6.3 Minutes of the 3rd Project Meeting (Internal)

1. Arrangement
   - Date: August 14, 2020
   - Time: 7:00pm
   - Place: Online via WeChat
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun

- Absent: Prof. Raymond Chi-Wing Wong
- Recorder: Shang-Ling Hsu

2. Approval of minutes
   - The minutes of last meeting were approved without amendment.
3. Report on progress
   - All team members have studied the papers distributed to them, citing the three provided by Prof. Wong.
   - All team members have written detailed memorandum and summary of the papers in a format approved by the team members.
4. Discussion items
   - Each of us talked about the papers we were responsible for respectively to the others.
   - The related works in fake news detection can be roughly categorized into:
     - Content-based: classify a piece of news typically by the image and text
     - Social-context based: classify a piece of news by its retweet or reply pattern
     - Fused and others: the model merging features from different sources
   - There is a variety of methods to conduct fake news detection:
     - The degree to which the title matches the contents of a piece of news
     - Do a "reverse text search" to see if the event described in the news is different from what have been fact-checked
     - Utilize the citation relationship among news to build a graph, and treat it as a node classification task
     - Cross-lingual learning: see if a model trained in one language with rich data (e.g. English) can be somehow applied to another language with fine-tuning
5. Goals for the coming week
   - Prepare a presentation summarizing our literature review for the meeting with Prof. Wong.
6. Meeting adjournment and next meeting
   - The meeting was on time.
   - The next meeting will be with Prof. Wong at 2pm on August 19, 2020 via Zoom.

6.4 Minutes of the 4^{th} Project Meeting

1. Arrangement

   - Date: August 19, 2020
   - Time: 2:00pm
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun,
     Prof. Raymond Chi-Wing Wong
   - Absent: None
   - Recorder: Shang-Ling Hsu, Yanjia Li

2. Approval of minutes

   - The minutes of last meeting were approved without amendment.

3. Report on progress

   - All team members have studied the papers citing the three provided by Prof. Wong.
   - All team members have prepared the presentation for literature review.

4. Discussion items

   - Confirm our direction: propose a new model which addresses the fake news detection problem
   - Good summarization skills for literature review:
     - Layout a framework for the previous studies
     - See where each related work is in the framework of the lite
     - See where our work will be in the framework
   - Potential features to be used: Twitter posts, comments, feedback (number of "likes"), news text, news image
   - Potential dataset: Weibo and Twitter
   - Potential adoption of heterogeneous graph neural networks (HetGNN) in this task
     - It is very new and yields high accuracy in other domains.
     - Hierarchical aggregation: aggregate heterogeneous nodes
     - Model the entities in our task as nodes, and define the initial weights and the transition (i.e. edges). Attention mechanisms may be adopted.

- Hyper-parameter tuning can be arduous. When frustrated by tuning the parameters, we may think about chemistry researchers, who stand in the lab everyday instead of asking the computer to conduct the experiments and relax like us. In addition, they need to do postdoctoral research for many years before becoming a professor, which is worse than us.
- Compared to Tensorflow 1.0, Pytorch is suggested for research due to its simplicity.

5. Goals for the coming week
   - Check whether the information we need to train our model are all included in the public Weibo and Twitter datasets, so the research is feasible.
   - Pick a few state-of-the-art related works to be our baselines.
   - Check if HetGNN has been adopted in fake news detection.
   - Discuss our proposal
6. Meeting adjournment and next meeting
   - The meeting was on time.
   - The next meeting will be internal at 7pm on September 2, 2020 via WeChat.

6.5 Minutes of the 5^{th} Project Meeting (Internal)

1. Arrangement
   - Date: September 2, 2020
   - Time: 7:00pm
   - Place: Online via WeChat
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong
   - Recorder: Shang-Ling Hsu
2. Approval of minutes
   - The minutes of last meeting were approved without amendment.
3. Report on progress
   - Each of us has found and checked the datasets we were assigned to study.
   - Each of us has read and summarized a few more related works, including two regarding HetGNN.
4. Discussion items

- Together, we picked four baselines, which are the state-of-the-art methods conducted from different perspectives.
- We distributed the jobs regarding the proposal writing. Each of us is responsible for writing one or more sections of the proposal.
- We discussed how we may apply HetGNN in fake news detection. There are many possibilities (the following items may intersect):
  - Consider both social context and content
  - Consider content only.
  - Each node represents a piece of news.
  - Each node represents a tweet.
  - Each user can also be a node.
  - Each source can also be a node.
- We may refer to COBRA's idea, projecting news into the same latent space.

5. Goals for the coming week
   - Finish our respective assigned part of the proposal.
   - Test our the baseline models and see if they perform as good on our datasets (Twitter and Weibo) as stated in their papers.
   - Finish crawling the datasets.
6. Meeting adjournment and next meeting
   - The meeting was on time.
   - The next meeting will be before or after the submission time of the proposal.

6.6 Minutes of the 6^{th} Project Meeting (Internal)

1. Arrangement
   - Date: September 23, 2020
   - Time: 8:00pm
   - Place: Online via WeChat
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong
   - Recorder: Shang-Ling Hsu
2. Approval of minutes

- The minutes of last meeting were approved without amendment.

3. Report on progress

    - Each of us has finished our part of the proposal report.
    - Each of us has revised our part of the proposal report according to the feedback from Prof. Wong and the tutor.

4. Discussion items

    - The model proposed in the proposal is still yet to be improved. We do not have to stick to it if we come up with better ideas.
    - We may improve the performance of HetGNN in this task by more meaningful encoding of the features, as our features are diverse and deeply correlated across modals and news. A better fusion method may bring much improvement on the performance.
    - Since we are not sure about whether it is feasible to adopt the model on our task, we will try adopting the model first and evaluate the performance, analyze the drawbacks, and decide how we will improve the model. Even if we do not use the model in the end, it will be an interesting baseline.
    - The progress report is due on October 23rd, so we should be aware of our progress.
    - The details of "random walk" are not described clearly either in the paper nor in the source code. We have to figure out what we can do with it. (Or email the authors for the details?)
    - The Twitter dataset being downloaded on the server is too slow. The issue may be caused by the reach of the API key limit. We will try to apply for a new one in order to download the dataset.

5. Goals for the coming week

    - Read the source code of HetGNN.
    - See how to adopt the Weibo dataset into the model.
    - Ask Prof. Wong for the server access
    - Create Twitter Developer accounts for more API keys

6. Meeting adjournment and next meeting

    - The meeting was on time.
    - The next meeting will be at 8:00pm on September 30, 2020.

6.7 Minutes of the 7^{th} Project Meeting (Internal)

1. Arrangement
   - Date: September 30, 2020
   - Time: 8:00pm
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong
   - Recorder: Shang-Ling Hsu
2. Approval of minutes
   - The minutes of last meeting were approved without amendment.
3. Report on progress
   - Each of us has read the source code of HetGNN.
   - Each of us has thought about how to adopt the Weibo dataset into the model.
   - We asked Prof. Wong for the server access.
   - We created a Twitter Developer account for one more API key.
4. Discussion items
   - The source code of HetGNN is hard-coded for author-paper-venue application, which makes it difficult to be adopted directly on our task. We may have to:
     (a) Modify tremendous amount of code from the original implementation.
     (b) Rewrite our own version of HetGNN, adopting the concept only.
   - There are several ways to adopt HetGNN on our task. As there are no relevant suggestions on how we may apply the model, we may start from the most intuitive way. In our formulation, each tweet, each retweet, each author, and each source are all separate nodes. In contrast, the replies to the tweets may be regarded as part of the tweet nodes instead of separate nodes. Alternatively, we may disregard the replies for now.
   - We may conduct some experiments on the dataset we have now to see if there is any temporal information to be captured. For example, can the content of a news be misunderstood after being forwarded too many times? This requires us to have a deeper understanding on the mechanism of retweeting.
   - As we have at least three kinds of nodes: user, source, and news, is it possible for us to do multi-tasking, i.e. extent the classification goal from just "fake news detection" to also "social bot detection" and "unreliable news source detection?" As they should be highly correlated.

- Which features in the user information will we use? We will look at the ones utilized in previous studies for comparisons.
- We confirmed the job distribution for the LANG 4030 proposal report.

5. Goals for the coming week

   - Extract and encode the text in the dataset.
   - Download and encode the images.
   - Extracts useful data (e.g. user-tweet relationships, user profiles) from the datasets

6. Meeting adjournment and next meeting

   - The meeting was on time.
   - The next meeting will be at 8:00pm on October 8, 2020.

6.8 Minutes of the 8^{th} Project Meeting (Internal)

1. Arrangement

   - Date: October 8, 2020
   - Time: 8:30pm
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong
   - Recorder: Shang-Ling Hsu

2. Approval of minutes

   - The minutes of last meeting were approved without amendment.

3. Report on progress

   - The images have been downloaded and encoded.
   - The text has been encoded.
   - Useful data (e.g. user-tweet relationships, user profiles) have been extracted from the dataset.

4. Discussion items

   - Due to the limit of Twitter API, downloading profile images takes too long. We will not use it for now.
   - Retweets in Twitter may include images, but this is not the case in Weibo.

- For comparison, we choose to use user vector that is similar to GCAN. 10 features are included, mainly images, text, and other attributes.
- New Twitter API keys will be added to download the Twitter dataset to speed up the download.
- The text embedding model is self-supervised and thus can be fine-tuned on the dataset directly, but this is not the case for images. Hence, we will use the image encoder trained on ImageNet.
- The implementation pf saving the encoding of images and text directly instead of embedding them in real-time makes it impossible to optimize the images and text encoders. It is acceptable if the whole model is so complex and difficult to train that the quality of text and image embeddings are relatively less important as we will focus on the training of the remaining part of the model.
- The dimension of the embedding in each domain should be unified. For now, text=768, image=512, and it is less for user embedding. Some projection function or its approximation may be required.
- As we are implementing our baseline, we stick to the original design for now. Any improvements will be our next steps, e.g. frequency domain, better text model, attentions for explainability.
- Heterogeneous Attentive Neural Networks (HAN) may also be useful for us. We will consider it when improving our own model design.

5. Goals for the coming week
    - Finish our respective part of the progress report.
    - Meeting with our communication tutor.
6. Meeting adjournment and next meeting
    - The meeting was on time.
    - The next meeting will be at 8:00pm on October 12, 2020.

6.9 Minutes of the 9^{th} Project Meeting

1. Arrangement
    - Date: October 30, 2020
    - Time: 3:30pm
    - Place: Online via Zoom

- Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun, Prof. Raymond Chi-Wing Wong
- Absent: None
- Recorder: Shang-Ling Hsu

2. Approval of minutes
   - The minutes of last meeting were approved without amendment.
3. Report on progress
   - We have submitted finished our responsible part of the progress report.
   - Prof. Wong and another reader have graded our progress report.
   - We have met our communication tutor before submitting our individual essays.
4. Discussion items
   - Challenge 1: The encoding model. Although XLM-RoBERTa is multilingual, it is too large to load the model in to our server RAM, so we have to resort to use two separate but smaller models (the ordinary BERT) for English and Chinese respectively.
   - Challenge 2: The API limit makes our data collecting process very slow, but we can use the Weibo dataset without the profile photos first, which is the part of the dataset that we already have in hand.
   - The intuition of using the profile photos in the Weibo datasets is that social bots usually does not set a profile photo.
5. Goals for the coming week
   - Use the ordinary BERT for text encoding.
   - Check the memory occupation of XLM-RoBERTa.
   - Keep gathering the datasets.
6. Meeting adjournment and next meeting
   - The meeting was on time.
   - The next meeting will be at 9:00pm on November 8, 2020.

6.10 Minutes of the 10^{th} Project Meeting (Internal)

1. Arrangement
   - Date: November 8, 2020
   - Time: 9:00am

- Place: Online via Zoom
- Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
- Absent: Prof. Raymond Chi-Wing Wong
- Recorder: Yanjia Li

2. Approval of minutes
   - The minutes of last meeting were approved without amendment.
3. Report on progress
   - We have updated our progress on image pre-processing on pixar frequency domain
   - We have updated our progress on text encoding by Par2Vec
4. Discussion items
   - We have discussed the structure of our HetGNN graph and the connection between re-post users, authors, news posts, and the re-posts.
   - We have discussed the features that should be contained in a re-post node and a re-post users and how those features should be aggregated.
5. Goals for the coming week
   - To obtain the output of Par2Vec model and ResNet model(i.e.the feature vectors of text and images)
   - To finish the script of Bi-LSTM model
   - To run Bi-LSTM model on the feature vectors
6. Meeting adjournment and next meeting
   - The meeting was on time.
   - The next meeting will be at 8:00pm on Nov 11, 2020.

6.11 Minutes of the 11^{th} Project Meeting (Internal)

1. Arrangement
   - Date: November 11, 2020
   - Time: 8:00pm
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong

- Recorder: Shang-Ling Hsu

2. Approval of minutes

   - The minutes of last meeting were approved without amendment.

3. Report on progress

   - The par2vec encoding is almost finished and will be harvested the day after tomorrow.
   - The image encoding part has been done.
   - The feature encoding part of a node has been finished.

4. Discussion items

   - The text encoding part is being trained and will be finished by the day after tomorrow.
   - We distributed our works:
     - Random walk with restart: Shang-Ling
     - Encoder of the homogeneous neighbors: Tianle
     - Encoder of the heterogeneous neighbors: Yushi
     - The final FCNN and loss function: Yanjia

5. Goals for the coming week

   - Finish the respective assigned part to each of us.

6. Meeting adjournment and next meeting

   - The meeting was on time.
   - The next meeting will be at 8:00pm on November 19, 2020.

6.12 Minutes of the 12^{th} Project Meeting (Internal)

1. Arrangement

   - Date: November 19, 2020
   - Time: 8:00pm
   - Place: Online via WeChat
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong
   - Recorder: Shang-Ling Hsu

2. Approval of minutes

   - The minutes of last meeting were approved without amendment.

3. Report on progress
   - Each member has finished their assigned parts.
4. Discussion items
   - Tanle and Yushi explain the details of their parts of implementation.
   - We reported our progress on the final part and the random walk as well.
5. Goals for the coming week
   - Tianle and Yushi will connect the components we coded.
   - Yanjia and Kate will code the part for data loading and batchification.
   - Finish the monthly report together in a shared document.
6. Meeting adjournment and next meeting
   - The meeting was on time.
   - The next meeting will be at 8:00pm on November 24, 2020.

6.13 Minutes of the 13^{th} Project Meeting (Internal)

1. Arrangement
   - Date: December 22, 2020
   - Time: 10:00am
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong
   - Recorder: Shang-Ling Hsu
2. Approval of minutes
   - The minutes of last meeting were approved without amendment.
3. Report on progress
   - Tianle and Yushi have connected most of the components we coded.
4. Discussion items
   - How to read and integrate the multi modal inputs
5. Goals for the coming week
   - Integrate the inputs and the different components of the model

6. Meeting adjournment and next meeting
    - The meeting was on time.
    - The next meeting will be at 8:00pm on December 26, 2020.

6.14 Minutes of the 14^{th} Project Meeting (Internal)

1. Arrangement
    - Date: December 26, 2020
    - Time: 8:00pm
    - Place: Online via Zoom
    - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
    - Absent: Prof. Raymond Chi-Wing Wong
    - Recorder: Shang-Ling Hsu
2. Approval of minutes
    - The minutes of last meeting were approved without amendment.
3. Report on progress
    - Completed training pipeline
4. Discussion items
    - Future work may be adding source (e.g. from URL) to make the graph denser.
5. Goals for the coming week
    - Inspect the data: Is the graph sparse or dense? Only the latter one makes sense to use HetGNN.
6. Meeting adjournment and next meeting
    - The meeting was on time.
    - The next meeting will be at 9:00pm on January 8, 2021.

6.15 Minutes of the 15^{th} Project Meeting (Internal)

1. Arrangement
    - Date: January 8, 2021
    - Time: 9:00pm
    - Place: Online via Zoom

- Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
- Absent: Prof. Raymond Chi-Wing Wong
- Recorder: Yushi Sun

2. Approval of minutes
    - The minutes of last meeting were approved without amendment.
3. Report on progress
    - Yushi and Yanjia trained the adapted HetGNN model and achieved 95% accuracy on test set
    - Tianle and Shang-Ling inspected the explanability part
4. Discussion items
    - How to enlarge the dataset to improve the genrative power of the model-Adapt Twitter dataset for our model
    - Finetune the parameters of our model in random walk part
    - Discussed the implementation of baseline models to compare the performance
5. Goals for the coming week
    - Use the RoBERTa embedding in our model to see if the performance change
    - Further finetune the parameters in hidden layer and the random walk part
    - Explore and prepare the training data (test embedding) from Twitter dataset
    - Yushi will look for data augmentation and baseline models if time permits
6. Meeting adjournment and next meeting
    - The meeting was adjourned at 10:00 pm
    - The next meeting will be at 9:00pm on January 14, 2021.

6.16 Minutes of the 16^{th} Project Meeting

1. Arrangement
    - Date: January 14, 2021
    - Time: 6:00pm
    - Place: Online via Zoom
    - Present: Prof. Raymond Chi-Wing Wong, Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun

- Absent: None
- Recorder: Yushi Sun

2. Approval of minutes

- The minutes of last meeting were approved without amendment.

3. Report on progress

- Found out that the performance of RoBERTa embedding is similar to that of Par2Vec.
- Tried to apply different length of users and posts neighbor lists.
- Looked for some explanability methodologies.

4. Discussion items

- How to evaluate the performance of our model and compare with others'.
- How to adapt our model to Twitter dataset.
- How to enhance explanability in our model.

5. Goals for the coming week

- Yushi and Yanjia will look for some baselines to compare the performance.
- Tianle and Kate will implement the co-attention mechanism in our current model and see if it's possible to replace the Bi-RNN structure in our current model with attention.

6. Meeting adjournment and next meeting

- The meeting was adjourned at around 7:00 pm
- The next meeting will be at 8:00 pm on January 20, 2021.

6.17 Minutes of the 17^{th} Project Meeting (Internal)

1. Arrangement

- Date: January 20, 2021
- Time: 9:00pm
- Place: Online via Zoom
- Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
- Absent: Prof. Raymond Chi-Wing Wong
- Recorder: Yushi Sun

2. Approval of minutes

- The minutes of last meeting were approved without amendment.

3. Report on progress

   - Yushi and Yanjia looked for several baseline models.
   - Tianle and Kate explored the co-attention mechanism

4. Discussion items

   - How to extend the current model to allow batch operations?
   - How to utilize the similarities between users and posts?
   - How to modify the co-attention mechanism to fit in our model?

5. Goals for the coming week

   - Yushi and Tianle modify the code to allow batch operations.
   - Yanjia continues on implementing one of the baselines.
   - Kate computes the similarity scores for users and posts.

6. Meeting adjournment and next meeting

   - The meeting was adjourned at 10:30 pm
   - The next meeting will be at 9:00pm on January 25, 2021.

6.18 Minutes of the 18^{th} Project Meeting (Internal)

1. Arrangement

   - Date: January 25, 2021
   - Time: 9:00pm
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong
   - Recorder: Shang-Ling Hsu

2. Approval of minutes

   - The minutes of last meeting were approved without amendment.

3. Report on progress

   - Yushi and Tianle have modified the code to allow batch operations.
   - Yanjia has finished implementing one of the baselines.

- Kate has finished coding for computing the similarity scores for users and posts.

4. Discussion items

   - Prioritize: The most important goal: Beat the 96% baseline!
   - (If time permits) Parse paragraph to KG? (Call large API built from super large dataset instead of building our own)

5. Goals for the coming week

   - Post (content) similarity: add an edge with similarity weight (4664 x 4664 should not be too large) → Kate
   - Change Bi-GRU to attention → Yanjia
   - Co-attention → Tianle, Yushi

6. Meeting adjournment and next meeting

   - The meeting was on time.
   - The next meeting will be at 9:00pm on January 31, 2021.

6.19 Minutes of the 19^{th} Project Meeting (Internal)

1. Arrangement

   - Date: January 31, 2021
   - Time: 9:00pm
   - Place: Online via Zoom
   - Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
   - Absent: Prof. Raymond Chi-Wing Wong
   - Recorder: Shang-Ling Hsu

2. Approval of minutes

   - The minutes of last meeting were approved without amendment.

3. Report on progress

   - Kate is still computing the scores on server.
   - Yanjia has changed Bi-GRU to attention.
   - Tianle and Yushi have finished implementing the co-attention.

4. Discussion items

   - We can make use of meta-path.

- Try to use the random-walk neighbors directly first.
- The FakeNewsDetection dataset (Twitter) has questionable quality, so we will conduct research on it after other experiments.

5. Goals for the coming week

- w2v RoBERTa's word-by-word embeddings (2-dim) for co-attention → Kate
- Image data input → Yushi
- Test RoBERTa for co-attention → Yushi
- Post-to-post link with the similarity computed: Run random walk for it → Kate
- Change all Bi-RNN to attentions → Yanjia
- Progress report (update results) (due Feb 12)
  - Methodology: methods tested → Yushi
  - Gannt + meeting minutes etc → Kate
  - Design → Tianle
  - Evaluation & results → Yushi & Tianle
- Send baseline results to Yushi & Tianle → Yanjia

6. Meeting adjournment and next meeting

- The meeting was on time.
- The next meeting will be at 9:00am on February 9, 2021.

6.20 Minutes of the 20^{th} Project Meeting (Internal)

1. Arrangement

- Date: February 9, 2021
- Time: 9:00am
- Place: Online via Zoom
- Present: Shang-Ling Hsu, Tianle Li, Yanjia Li, Yushi Sun
- Absent: Prof. Raymond Chi-Wing Wong
- Recorder: Shang-Ling Hsu

2. Approval of minutes

- The minutes of last meeting were approved without amendment.

3. Report on progress

- All of us finished our respective parts of the progress report.
- Yushi and Yanjia finished conducting experiments on the co-attention.
- Kate finished w2v & RoBERTa's word-by-word embeddings (2-dim) and random walk on the weighted graph.

4. Discussion items

- Accuracy for the attention version (changed the Bi-LSTM of same type agg to attention mechanisms): 0.933
- IARNet is just HetGAN (totally the same model used in a new application, i.e. fake news detection), and we are using HetGNN. What is our innovation?
  - Co-attention?
  - The only contribution of IARNet is to be the first study of using HetGNN in fake news detection, so we cannot repeat it.

5. Goals for the coming week

- Finalized and submit the progress report.
- Think about the next steps and our innovation. (Review our initial idea.)
- Experiment on the similarity-weighted graph.
  - Convert the weighted-version neighbor list to node file → Yanjia.
- Think about / survey on the possible formulations and previous studies about knowledge graph in fake news detection.
- Experiments on Twitter dataset (the one other than the FakeNewsNet.)
- Finetune BERT/XLM-RoBERTa on Weibo and Twitter → Kate.

6. Meeting adjournment and next meeting

- The meeting was on time.
- The next meeting will be in February after the report submission, 2021.