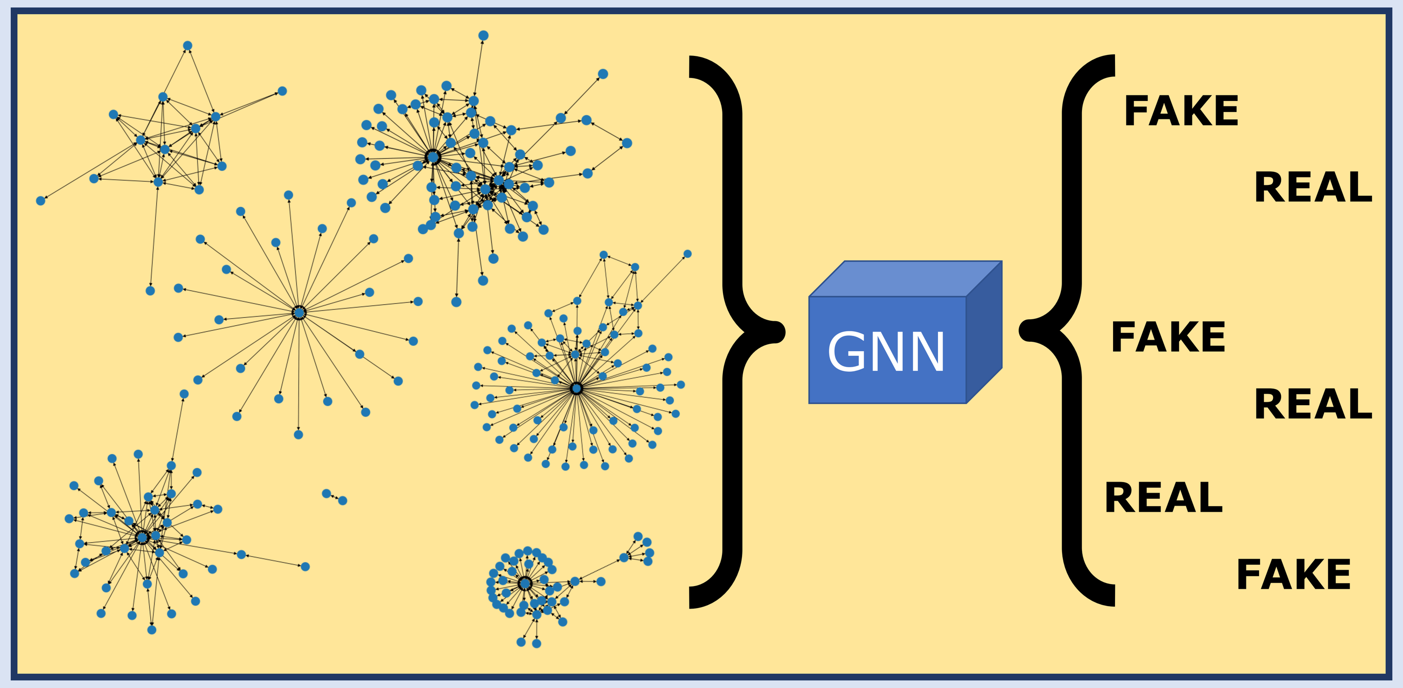
COLORADO COLLEGE
Department of Computer Science

# Finding Fake News Without The News:
## Structural Detection of Misinformation Using Machine Learning

Simay Cural and Max Perozek, supervised by Cory B. Scott

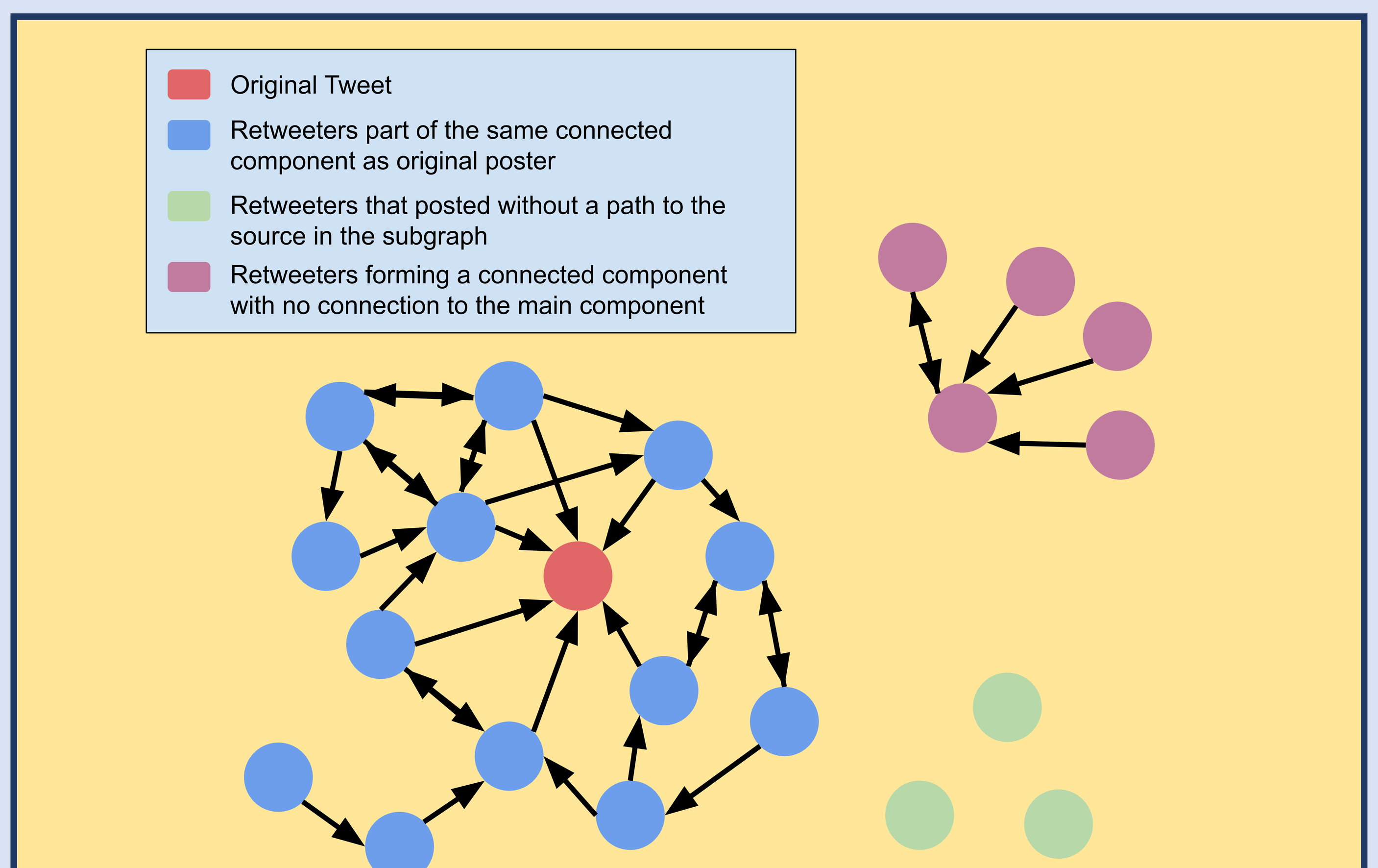
s_cural@coloradocollege.edu
m_perozek@coloradocollege.edu
cbs@coloradocollege.edu




## Background and Motivation

- False information spreads easily on social media.
- Traditionally, machine learning (ML) methods for misinformation prediction have used natural language processing (NLP) tools.
- However, NLP tools struggle at this task, since

1. tweets may have too little text to classify confidently,
2. for rare events like the COVID pandemic, there may not be much data on which to train a model.



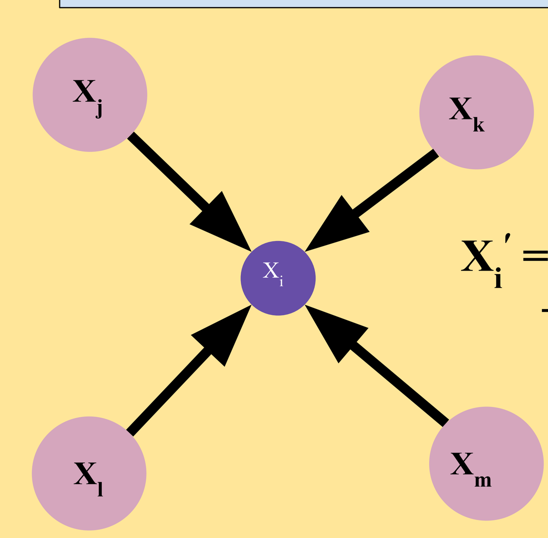

## Dataset

- 3,511 human labeled graphs representing social networks associated with misinformation and non-misinformation tweets.
- The graphs are induced from the twitter follower graph

## Graph Neural Network Detection of 5G Conspiracy Tweets

- Graph neural networks transform input data via message passing and aggregation using order-invariant functions.
- GNN architecture generates an embedding of a given graph which can be used for classification.
- There are many different GNN architectures, we focused most of our experiments on the Graph Isomorphism Network (GIN) architecture, as it has been proven to be equivalent to the Weisfeiler-Lehman isomorphism test in its discriminating power.
- Performed hyperparameter search experiments, where models were trained using cross entropy loss and the Adam optimizer.
- Models were evaluated using 5-fold cross validation
- F1 score used as primary criteria to evaluate model accuracy due to imbalanced nature of the dataset.

### GNN Message Passing Layer




$$X_i' = (1+\varepsilon)\cdot X_i + (X_j + X_k + X_l + X_m)$$

Graph Neural networks (GNNs) process semi-structured data by having each node exchange information with its neighbors, followed by processing with a small neural network.

## Analysis of GIN Hyperparameter Search

- Wide range of hyperparameters for embedding dimension and number of hidden layers.
- All top models had a batch size of 64.
- Dropout did not seem to have a significant effect on model performance. We interpret this to be linked to the small size of models that were trained and the corresponding inability to overfit the training dataset.

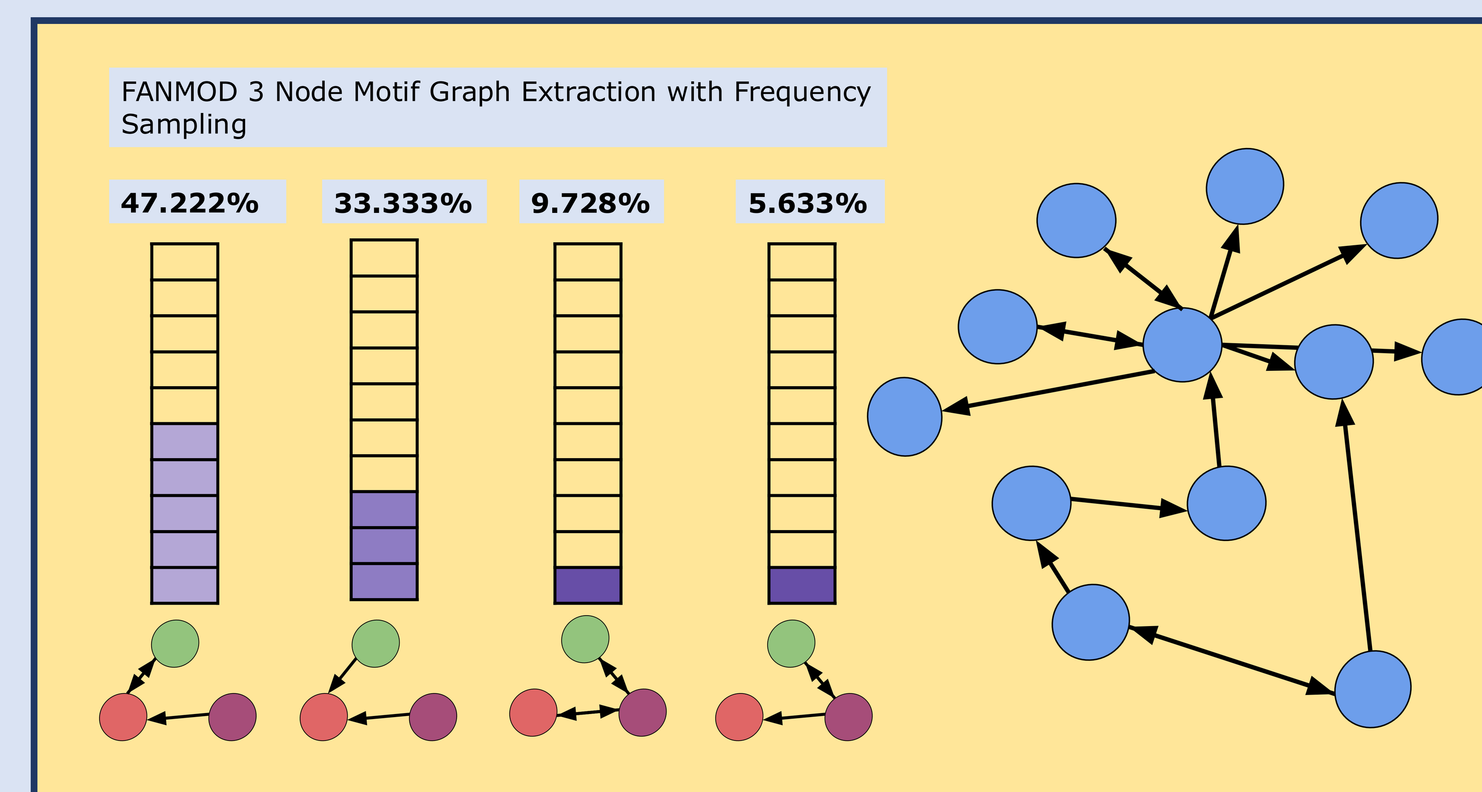
| F1 Score | Evaluation Accuracy | Embedding Dimension | Layers | Aggregation Function | Batch Size | Dropout |
|---|---|---|---|---|---|---|
| 0.696 ± .02 | 60.21% ± 5.2% | 32 | 3 | sum | 64 | 0.0 |
| 0.684 ± .02 | 67.35% ± 2.5% | 64 | 3 | mean | 64 | 0.0 |
| 0.675 ± .04 | 63.96% ± 9.3% | 64 | 4 | mean | 64 | 0.0 |
| 0.662 ± .06 | 59.23% ± 3.2% | 32 | 4 | mean | 64 | 0.5 |
| 0.661 ± .03 | 56.43% ± 3.6% | 64 | 3 | mean | 64 | 0.5 |

## Motif Finding Using GNNs

Data was processed in two different structures. The first structure made use of the general features of each of the graphs (ex. number of nodes, edges, etc). The second structure looked for specific motifs that were present in the graphs, hoping to see that certain motifs were more likely to be found in fake news, possibly indicating a sign of bot activity.

## Structure A: Graph Feature Extraction

- This is a more straightforward approach to observing whether if misinformation graphs have any distinguishable features.
- Features like average clustering number, number of source/sink nodes, page rank features (standard deviation, average), and cyclic triangular clusters were extracted.
- The data was then manipulated into tensors with labels specifying false/true news to be fed into a basic NN.
- The accuracy rating on this data ranged from around 60% to 70%.




## Structure B: Motif Research

- For this structure a pre-existing motif sorting algorithm, FANMOD, was used. FANMOD is a fast motif detection tool that takes in an graphs in adjacency matrix form.
- For the research 3 and 4 node motifs were extracted.
- Similar to structure A, motifs were fed into a basic NN resulting in an accuracy rate ranging from 30% to 40%.