Results and Ablation Study of our proposed
BotDetection model

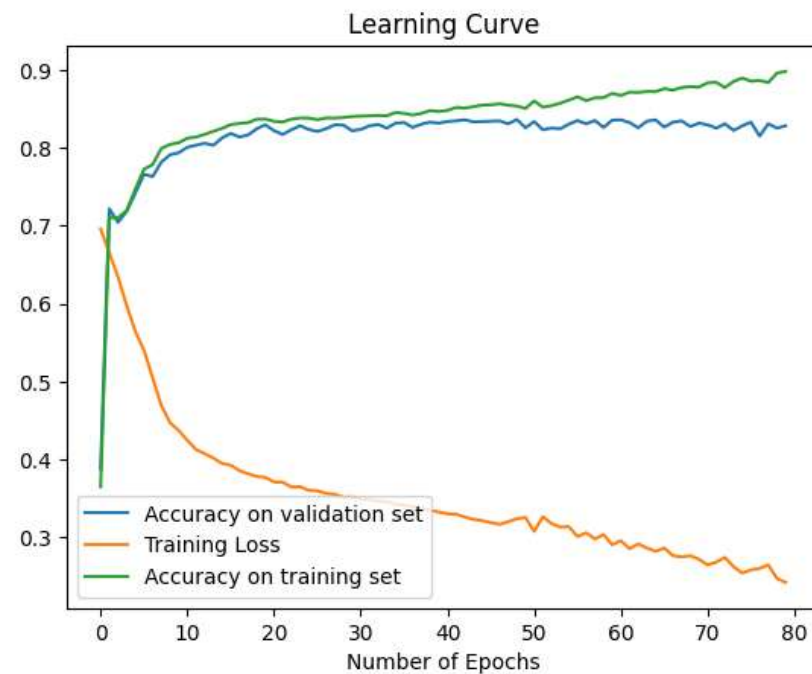Author:        Tsingelis Konstantinos

Supervisor:    Prof. Askoynis Dimitrios(NTUA)

# Results

- Evaluation metrics of our model.

| Evaluation Metrics | Values(%) |
|:---:|:---:|
| Accuracy | 85.55 |
| Precision | 81.31 |
| Recall | 95.16 |
| F1-score | 87.69 |



Learning Curve

# Results

- Our proposed model demonstrates competitive performance, surpassing most state-of-the-art architectures .

- Significant improvement compared to the BotRGCN, upon which our methodology is based (+2.3% accuracy).

| Method | Accuracy | F1-Score | Precision | Recall |
|---|---|---|---|---|
| **Miller et al.** | 64.50 | 74.81 | 60.71 | 97.44 |
| **Cresci et al.** | 47.76 | 13.69 | 7.66 | 64.47 |
| **BotOrNot** | 53.09 | 55.13 | 55.67 | 50.82 |
| **Kudugunta et al.** | 59.59 | 47.26 | 80.40 | 33.47 |
| **Alhossini et al.** | 59.92 | 72.09 | 57.83 | 95.72 |
| **BiLSTM** | 70.23 | 53.61 | 62.74 | 46.83 |
| **BotRGCN** | 83.27 | 85.26 | 81.39 | 89.53 |
| **SATAR** | 84.02 | 86.07 | 81.50 | 91.22 |
| **RGT** | 86.57 | 88.01 | 85.15 | 91.06 |
| **BIC** | 87.36 | 88.88 | 84.76 | 93.44 |
| **My_Model** | 85.55 | 87.69 | 81.31 | 95.16 |

# Ablation Study

- We experiment with different numbers of RGCN layers and assess their impact on the overall model performance.

- The optimal performance is achieved with 2 layers.



Accuracy — Number of RGCN layers
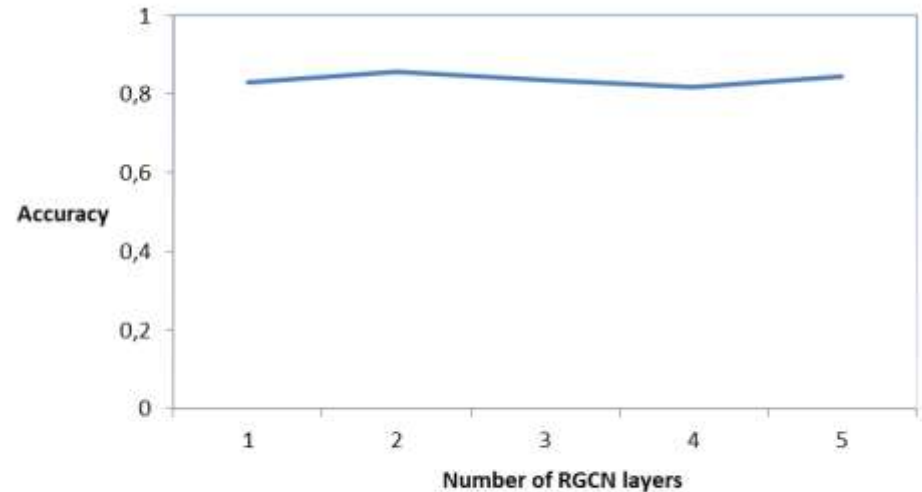
Accuracy of our model depending on the number of RGCN layers



F1-score — Number of RGCN layers
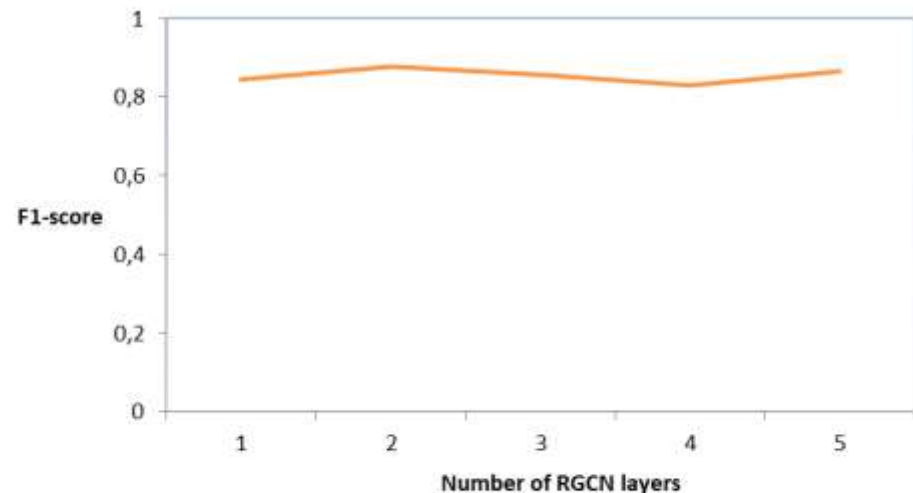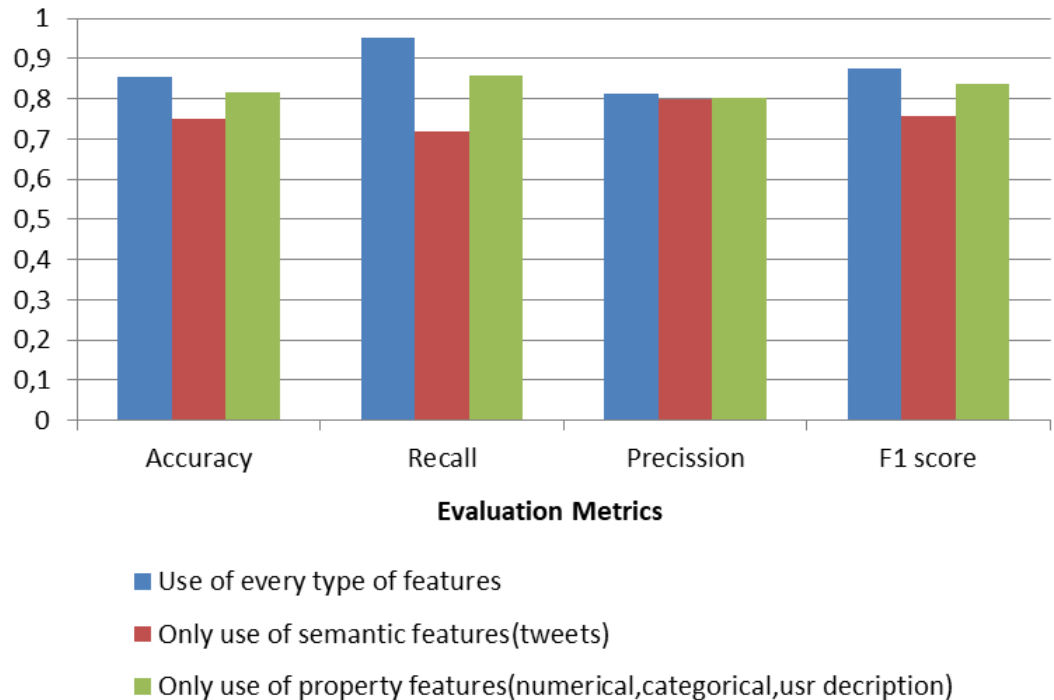
F1-score of our model depending on the number of RGCN layers'
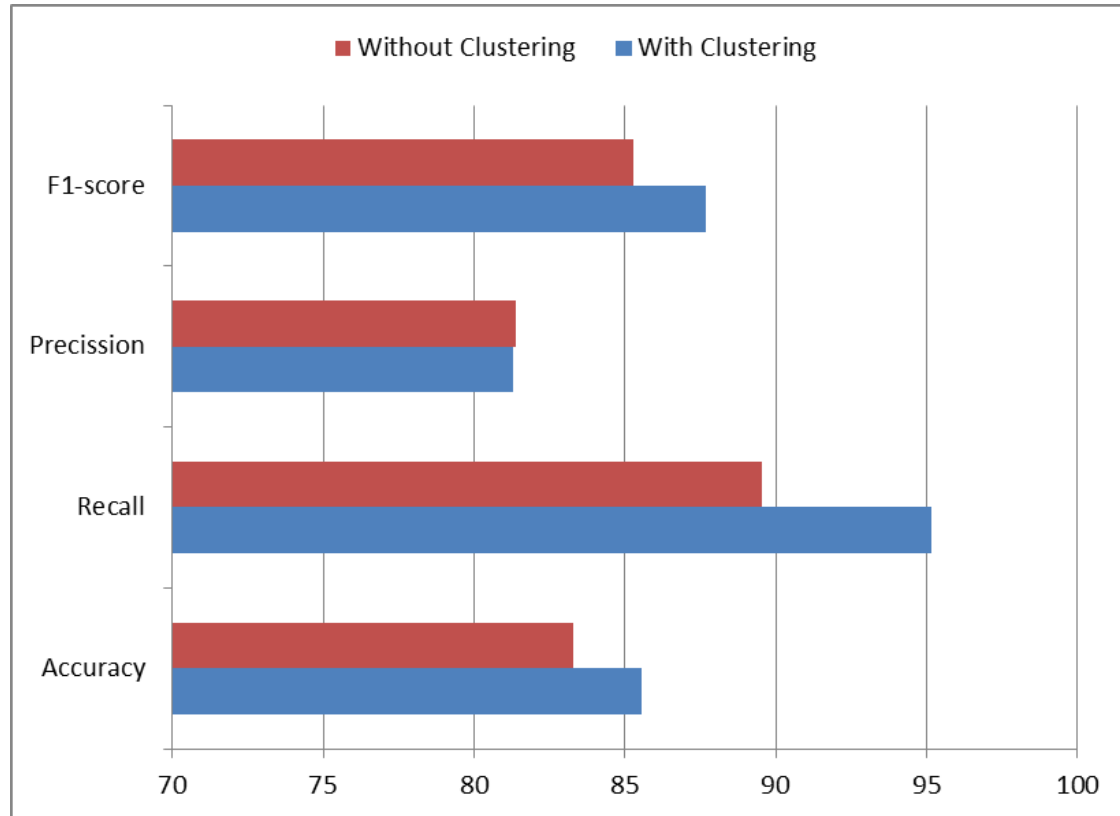
# Ablation Study

- We experiment with different combinations of features and assess their impact on the overall model performance.

- It is confirmed that each piece of user information is essential for developing high-performance models.
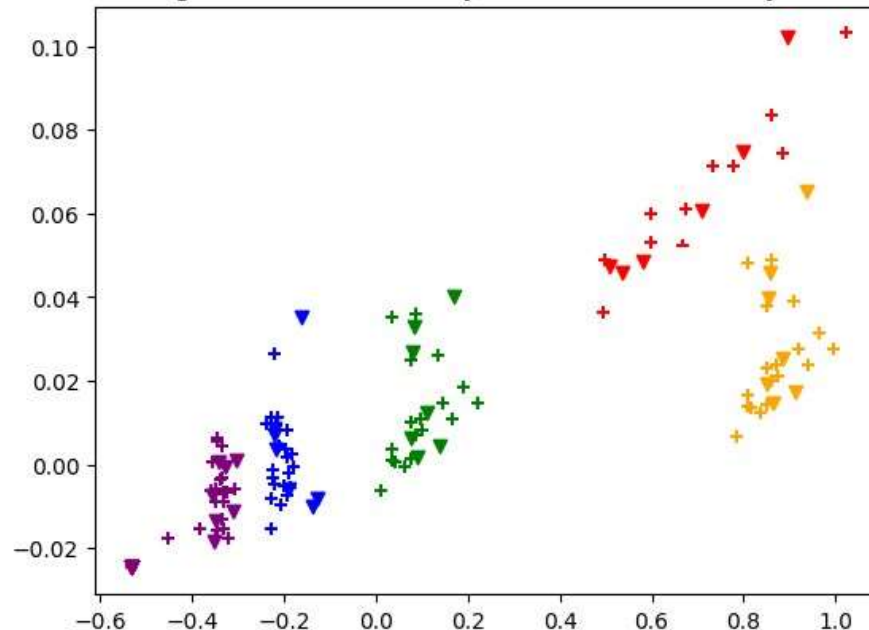
# Ablation Study

- We compare the performance of our model with the corresponding performance we would have if we had removed the clustering.

- It is confirmed that the combination of clustering and GNN significantly improves the model's performance.

# Ablation Study



Visualization of clustering in tweets where + symbolizes bots and v symbolizes genuine acounts

- We visualize the functionality of clustering.

- Intuitively, through clustering, we divide the set of tweets into 5 thematic classes, and for each user, we find their average sentiment for each thematic group. The inconsistency that characterizes bots, as they usually confine themselves to a specific thematic range in order to achieve their malicious objectives, makes the above method efficient