Learn over Past, Evolve for Future: Forecasting Temporal Trends for Fake News Detection

Beizhe Hu1,2, Qiang Sheng1,2, Juan Cao1,2, Yongchun Zhu1,2, Danding Wang1, Zhengjia Wang1,2, Zhiwei Jin3
1Institute of Computing Technology, Chinese Academy of Sciences  2University of Chinese Academy of Sciences  3ZhongKeRuijian Technology Co., Ltd

Introduction

- **Temporal shift** in real-world fake news detection scenarios:
  The rapidly-evolving nature of news leads to the distributional difference between offline and online data, namely temporal shift, which causes significant performance degradation to the fake news detection model.

- **Observation:** The appearance of news events on the same topic presents diverse temporal patterns, which indicate the different importance of news samples in the training set for detection in future quarters.

### Table: Former Data vs Future Data

<table>
<thead>
<tr>
<th>Former Data</th>
<th>Time t-1</th>
<th>Time t</th>
<th>Future Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offline Data</td>
<td>Train</td>
<td>Distribution Gap</td>
<td>Online Data</td>
</tr>
<tr>
<td>Fake News Detector</td>
<td>Predict</td>
<td>Performance Degradation</td>
<td></td>
</tr>
</tbody>
</table>

#### Step 4: Forecast-Based Adaptation

- **Based on the topic-wise forecasts of frequency distribution in Quarter Q**, we apply instance reweighting to the training set.
- We calculate and then normalize the ratio between the forecasted frequency of Topic i and the sum of all forecasted frequencies of the preserved topics:

\[
\rho_{i,Q} = \text{Bound} \left( \frac{p_i(f_{i,Q})}{\sum_{i\in Q} p_i(f_{i,Q})} \right)
\]

#### Step 5: Fake News Detector Training

- We use the new weights based on the forecasted temporal distribution to increase or decrease the impact of instances during the training process:

\[ L = -\frac{1}{N} \sum_{i=1}^{N} w_{i,Q} \text{CrossEntropy}(y_i, \hat{y}_i) \]

Experiments & Case Study

↑ FTT outperforms the baseline and four other methods across all quarters in terms of most of the metrics

- Our design makes the model not only more familiar with news items in existing topics but also more generalizable to news items in new topics.
- After training on the reweighted set, the detector flips its previously incorrect predictions.

Conclusion

- **Problem:** To the best of our knowledge, we are the first to incorporate the feature of topic-level temporal patterns for fake news detection.
- **Method:** We propose the FTT framework which forecasts temporal trends to tackle temporal generalization issue in fake news detection.
- **Industrial Value:** We experimentally show that our FTT overall outperforms compared methods while maintaining good compatibility with any neural network-based fake news detector.

### Table: Metrics Comparison

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Existing Topics</th>
<th>New Topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>0.6272</td>
<td>0.5842</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.7689</td>
<td>0.7369</td>
</tr>
<tr>
<td>Precision</td>
<td>0.7689</td>
<td>0.7369</td>
</tr>
<tr>
<td>Recall</td>
<td>0.5612</td>
<td>0.5612</td>
</tr>
</tbody>
</table>

**GitHub Repo**

↑ FTT outperforms the baseline and four other methods across all quarters in terms of most of the metrics

Our design makes the model not only more familiar with news items in existing topics but also more generalizable to news items in new topics.

After training on the reweighted set, the detector flips its previously incorrect predictions.