Executive Summary

An Adaptive Fusion of Statistical and Textual Features-Based Rumour Detection Method

Abhilash Singh

People choose social media platforms in the Internet era because of their capabilities for information sharing and free speech. However, a lot of inaccurate information and even rumours are spread on social media platforms due to the benefits of the abundance of information and quick transmission. Wilfully allowing rumours to proliferate would be extremely harmful to society. In the early research on rumour identification, features were mostly created manually, and rumours were classified using decision trees, SVMs, random forests, and other machine learning techniques. These techniques rely heavily on feature engineering, demand a lot of labour, and are unable to automatically identify rumours. The objective of rumour detection is to determine whether something is a rumour based on the pertinent content information posted by users (such as textual content, comments, communication mode, etc).

Based on the BERT model, the CNN+Attention module or the Bi LSTM+Attention module's effect was improved to some extent; in contrast to the text's sequence characteristics, its semantic features were able to more directly mine the text. They combined the CNN+Attention and Bi LSTM+Attention modules to account for the textual semantic features and sequence features to achieve the best results. As a result, the effect of adding the CNN+Attention module alone on the three datasets was better than adding the Bi LSTM+Attention module alone.

Based on an adaptive synthesis of statistical data and textual features, they suggested a rumour detection system. Statistical features were encoded using a VAE together with an attention mechanism that can both capture and highlight locational information. Both semantic and sequence features were considered while extracting textual features using the CNN+Attention and Bi LSTM+Attention parallel networks. Finally, to overcome the over-fitting issue brought on by the excessive usage of statistical feature bands, an adaptive valve component was utilized to merge meaningful statistical information with textual information. Three available datasets using this strategy produced outstanding results.

To further improve the model's capacity for learning, future work will incorporate a rumour transmission mechanism.

Source: [Information](#)

KEYWORDS
Rumour detection; adaptive gate; statistical features; textual features