SUMMARY

Factoring Fact-Checks:

Structured Information Extraction from Fact-Checking Articles

The research paper showcases an interesting approach towards the task of fact-checking. Authors discuss the experiments done to automate the summarization of fact checks into three factors: claim, claimant, and the claim verdict. Analysis of the data shows that the top 19% of fact-checkers publish 94% of the fact-checks. 6,216 English fact-check articles with 1,038 words and 24.4 paragraphs on average were used for data exploration.

As claims can be paraphrased in different ways, a fuzzy string matching concept was used where minimum window sub-string was used to find the paraphrased factors. A baseline model was developed that uses general observations, and tools like ClaimBuster as a baseline(for claim tagging) to tag the factors. Maintaining an expensive model for rules based on linguistic patterns was not feasible, so this problem was classified as a sequence tagging problem and a new-tuned BERT based model was made with some modifications. To utilize the knowledge of the relative position of factors, the uniform [CLS] token is replaced with a paragraph position token, in order to represent the relative position of a paragraph to help the BERT model learn better representations for each paragraph based on its location. Results reported show that this approach results in better performance of the model.

The baseline methods always return nonempty predictions, therefore the tight score is strictly equal to the loose score. The BERT model could tag overall 69%-75% claims, 86-90% claimants, and 96% - 97% verdicts for well-known fact-checkers. Overall, the BERT model performs much better than the intuitive baseline model. This demonstrates the usefulness of probability-based algorithms(like sequence tagging) over intuition and general observations.

The performance of tagging claimants and verdicts is improved after half of the test set is mixed with the training set and the model is retrained which shows that the performance of the model can keep improving when the model is trained on more fact check patterns

Once trained on lots of data and brought to a high accuracy, we can use this as a baseline to develop models to summarise readable resources. This technology can also be extended to be used in online checking where checking according to the rubric is automated.

The paper was well written, organized, and easy to understand, and all arguments were supported with the required statistics. It provided us a valuable insight into how BERT models work, how one can analyze article based data, and the importance of training models on sufficiently large datasets. If successful, this will bring a positive impact on e-journalism, as the fact-checkers will not need to spend large amounts of time using tools like ClaimReview to annotate their fact checks and the readers who want to save time, can still be well informed through the summary of long fact-check articles. However, BERT models are computationally intensive and will be costly for use on a large scale. This limitation was not mentioned in the paper.