# Experiments
## Performance

| Approach | Accuracy | | False Claims | | | True Claims | | | AUC |
|---|---|---|---|---|---|---|---|---|---|
| | Micro | Macro | Prec. | Rec. | F1 | Prec. | Rec. | F1 | |
| SVM | 0.793 | 0.796 | 0.739 | 0.748 | 0.791 | 0.754 | 0.843 | 0.796 | 0.884 |
| LSTM | 0.735 | 0.726 | 0.690 | 0.734 | 0.762 | 0.829 | 0.558 | 0.667 | 0.835 |
| CRF [9] | 0.618 | 0.610 | 0.630 | 0.740 | 0.680 | 0.594 | 0.460 | 0.520 | 0.611 |
| Reddit Doc2Vec | 0.657 | 0.647 | 0.651 | 0.740 | 0.692 | 0.667 | 0.567 | 0.613 | 0.686 |
| News Article CNN | 0.802 | 0.798 | 0.777 | 0.873 | 0.822 | **0.838** | 0.724 | 0.777 | 0.885 |
| **Joint (News Article + Reddit)** | **0.826** | **0.825** | **0.837** | **0.895** | **0.866** | 0.814 | **0.817** | **0.812** | **0.909** |

- There is a performance gain by concatenating the two models, a reason for it being so small is that we do not have Reddit comments from all the articles either by there not existing or limitations of the Reddit API.

- If take away all the articles not containing Reddit comments, the performance of the joint model over the news article model increases by 2.5%.

# Conclusion and Future Work

- Release a large-scale dataset for fake news detection using the popular social media forum Reddit.

- Presented a simple neural network model to jointly consider the claim text, news article body and the social media comments to classify fake news.

- Will do more analysis of the Reddit dataset and exploit network features to improve fake news detection.