Fake News Detection

Archita Pathak : 50207253
Jruvika Bhimani : 50207613
Group no. : 24

References:
1. Rubin, V. L., Chen, Y. and Conroy, N. J. (2015), Deception detection for news: Three types of fakes. Proc. Assoc. Info. Sci. Tech., 52: 1–4. doi:10.1002/pra2.2015.145052010083
2. Ankesh Anand, Tanmoy Chakraborty, Noseong Park (2016), We used Neural Networks to Detect Clickbaits: You won't believe what happened Next!
3. Martin Potthast, Johannes Kiesel, Kevin Reinartz, Janek Bevendorff, and Benno Stein (2017), A stylometric inquiry into hyperpartisan and fake news.
4. Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. Journal of Machine Learning Research 15(1) (2014) 1929–1958
5. Verónica Pérez-Rosas, Bennett Kleinberg, Alexandra Lefevre, Rada Mihalcea(2017), Automatic Detection of Fake News

TASK AND EXPERIENCE:

Task: Designing and training a neural network architecture to classify news articles as fake or real based on linguistic features.

Three types of fake news [1]:

- Serious Fabrication
- Hoaxes
- Satire

Features to be focused on:

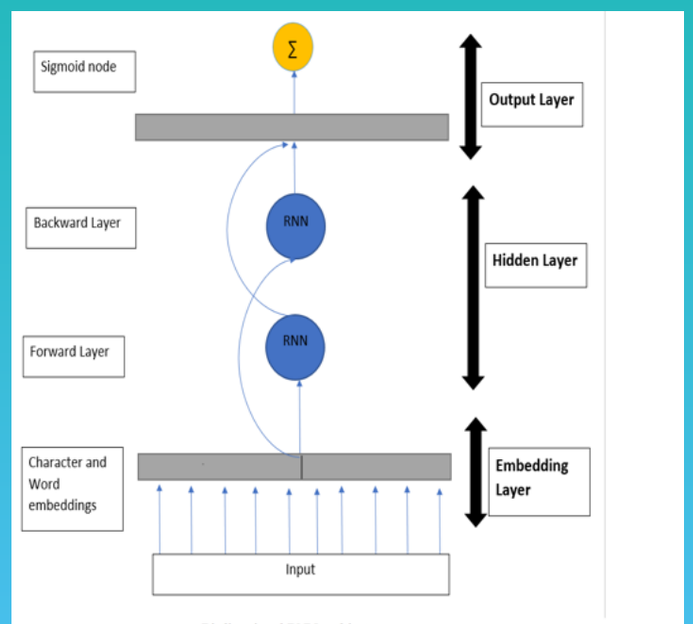
- Sentiment
- Grammar
- Punctuations
- Click baits

Dataset: Dataset from previous research:

- Clickbait detection: 15,000 [2]
- Buzzed dataset: 1627 news articles including 299 fake news articles [3]
- Word List: 800 profanities

MODEL AND ARCHITECTURE:

Model: Gated Recurrent Unit
Training: Mini-batch gradient descent
Cost Function: Binary Cross Entropy Loss
Optimization: ADAM optimizer for parameter updates
Regularization to avoid overfitting: Dropout technique [4]



Bi-directional RNN archiecture

PERFORMANCE MEASURE:

10-fold cross validation using:

- Accuracy
- Precision
- Recall
- F1 measure

Comparison baseline (previous results):

- Bag of words approach: 0.52 [3]
- Stylometric approach: 0.55 [3]
- NLP approach using ngrams, punctuations, syntax, readability etc: 75% [5]

Expected Result/ Aim:

- Accuracy >= 75%