Deep Learning

Assignment # 2

Introduction:

**Literature Review:**

Our goal of this research is to automatically identify whether the author of a speech or a tweet is in favor or against of that topic. It is widely discussed that LSTM and RNN are extremely useful in NLP and for stance detection. An exploratory analysis was conducted by Zarrella and Marsh where they maximize the available training data by use of transfer learning and create a high performing system. They used a recurrent neural network whose features were learned through distant supervision on two big unlabeled datasets. They also used word2vec skip gram method to train the embeddings of words and phrases and then used these features to train and learn those sentences using hashtag predication (auxiliary task). They fined tuned this model on a big labeled data of hundreds of examples for the stance detection. The author firstly discussed that how stance identification is different from sentiment analysis and then what are the challenges that comes up related to data collection and training for the problem and what makes them use transfer learning for the problem. Firstly, they use unsupervised learning on a large text dataset with a goal for their model to learn sequences and useful representation of words. Then they used a recurrent neural network consist of four layers. Theinput data to the model used one-hot encoding, these input tokens were projected via embedding layer which in turns feeds to recurrent layer of 128 LSTM units. At the end, into a three-dimensional output layer of three classes. In this research author trained a large amount of unsupervised learning which computationally very expensive and takes a lot of time which is difficult to replicate in normal scenarios because it needs a lot of resources to train much a huge model.

Another research was carried out by Jiachen Du and team related to target specific stance detection. Their research comprises of combining two parts, first in which a fully connected network which finds out about the target using an attention selector and second where they proposed a recurrent neural network (RNN) [Elman, 1990] model with LSTM as a feature extractor for text which take target specific information to classify the stance. They used two data sets i.e. English and Chinese in this research for better evaluation of their model. In start of paper the author discussed the importance of sentiment and stance analysis in NLP.The author believes that they must consider both the information of text content and target related features. They proposed a new model named Target specific Attention Neural Network (TAN) which is based on RNN with the modification to classify stance using target specific information. The author then explains their whole model and then compare the result of their model on the given two datasets with the standard LSTM model and shows how TAN outperformed the LSTM both in accuracy and time complexity using numbers and visual representations. In this research author used two languages to check the model performance which is better to analyze model performance than first approach we discuss above.

**Data Exploration and Visualization:**

**Evaluation Framework:**

Performance metrics

Our system is evaluated by macro-average of F1 score of all class (favor, against and none).

**Approach:**

Model Architecture:

Hyper Parameter tuning

Performance

**Ultimate judgment and Analysis:**

limitation

**Appendix:**

**References:**

Guido Zarrella and Amy Marsh. Mitre at semeval-2016 task 6: Transfer learning for stance detection. arXiv preprint arXiv:1606.03784, 2016.

[Elman, 1990] Jeffrey L Elman. Finding structure in time. Cognitive science, 14(2):179–211, 1990.