**Name: Kalintha Arun Sagar**

**Roll no: 15MI433**

1. Pengfei Liu Xipeng Qiu Xuanjing Huang, *Recurrent Neural Network for Text Classification with Multi-Task Learning ,* Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)

In this paper, multitask learning framework is used to jointly learn across multiple related tasks. Based on recurrent neural network, we propose three different mechanisms of sharing information to model text with task-specific and shared layers. The entire network is trained jointly on all these tasks. Experiments on four benchmark text classification tasks show that our proposed models can improve theperformance of a task with the help of other related tasks. The final model is fine-tuned with respect to a supervised training criterion with a gradient based optimization.

# 2.) [Zachary C. Lipton](https://arxiv.org/search/cs?searchtype=author&query=Lipton%2C+Z+C), [John Berkowitz](https://arxiv.org/search/cs?searchtype=author&query=Berkowitz%2C+J), [Charles Elkan](https://arxiv.org/search/cs?searchtype=author&query=Elkan%2C+C), A Critical Review of Recurrent Neural Networks for Sequence Learning

Countless learning tasks require dealing with sequential data. Image captioning, speech synthesis, and music generation all require that a model produce outputs that are sequences. In other domains, such as time series prediction, video analysis, and musical information retrieval, a model must learn from inputs that are sequences. Interactive tasks, such as translating natural language, engaging in dialogue, and controlling a robot, often demand both capabilities. Recurrent neural networks (RNNs) are connectionist models that capture the dynamics of sequences via cycles in the network of nodes. Unlike standard feedforward neural networks, recurrent networks retain a state that can represent information from an arbitrarily long context window. Although recurrent neural networks have traditionally been difficult to train, and often contain millions of parameters, recent advances in network architectures, optimization techniques, and parallel computation have enabled successful large-scale learning with them. In recent years, systems based on long short-term memory (LSTM) and bidirectional (BRNN) architectures have demonstrated ground-breaking performance on tasks as varied as image captioning, language translation, and handwriting recognition. In this survey, there is a review and synthesize the research that over the past three decades first yielded and then made practical these powerful learning models.