**Recurrent Neural Network for Text Classiﬁcation with Multi-Task Learning**

In most works the Neural Network models are learned based on single-task supervised objectives, which often suffer from insufﬁcient training data. In this paper, they use the multi task learning framework to jointly learn across multiple related tasks. Based on recurrent neural network, 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.

The primary role of these models is to represent the variable-length sentence or document as a ﬁxedlength vector. A good representation of the variable-length text should fully capture the semantics of natural language. The deep neural networks (DNN) based methods usually need a large-scale corpus due to the large number of parameters, it is hard to train a network that generalizes well with limited data. However, the costs are extremely expensive to build the large scale resources for some NLP tasks. To deal with this problem, these models often involve an unsupervised pre-training phase. The ﬁnal model is ﬁne-tuned with respect to a supervised training criterion with a gradient basedoptimization.

This Paper also puts forwards Multitasking learning where the learning utilizes the correlation between related tasks to improve classiﬁcation by learning tasks in parallel.

Three different ways are put forward by using the RNN Model.

**What is a recurrent neural network?**

A **recurrent neural network** (**RNN**) is a class of [artificial neural networks](https://en.wikipedia.org/wiki/Artificial_neural_network) where connections between nodes form a [directed graph](https://en.wikipedia.org/wiki/Directed_graph) along a temporal sequence. This allows it to exhibit temporal dynamic behavior. Unlike [feedforward neural networks](https://en.wikipedia.org/wiki/Feedforward_neural_networks), RNNs can use their internal state (memory) to process sequences of inputs. This makes them applicable to tasks such as unsegmented, connected [handwriting recognition](https://en.wikipedia.org/wiki/Handwriting_recognition)[[1]](https://en.wikipedia.org/wiki/Recurrent_neural_network#cite_note-1) or [speech recognition](https://en.wikipedia.org/wiki/Speech_recognition). **Recurrent Neural Networks**. **Recurrent Neural Network** remembers the past and it’s decisions are influenced by what it has learnt from the past.

## What is Long Short-term Memory?

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single data points (such as images), but also entire sequences of data (such as speech or video). For example, LSTM is applicable to tasks such as unsegmented, connected handwriting recognition, speech recognition and anomaly detection in network traffic or IDS's (intrusion detection systems).

The Three Models

1. In Model-I, the different tasks share a same LSTM layer and an embedding layer besides their own embedding layers.
2. 2. In Model-II, we assign a LSTM layer for each task, which can use the information for the LSTM layer of the other task.
3. Model-IIIalsoassigns a separate LSTM layer for each task, but introduces a bidirectional LSTM layer to capture the shared information for all the tasks.

Or in a simple manner

The ﬁrst model uses just one shared layer for all the tasks. The second model uses different layers for different tasks, but each layer can read information from other layers. The third model not only assigns one speciﬁc layer for each task, but also builds a shared layer for all the tasks. Besides, a gating mechanism is also introduced to enable the model to selectively utilizethe shared information. The entire network is trained jointly on all these tasks.

The there is about Hyperparameter training, Error Analysis, Dataset and all

Overall this model was able to perform very efficiently,it was in par with other state of art neural networks. The Only exception is in the case of Tree LSTM.

https://github.com/RaRe-Technologies/gensim