Recurrent Neural Network

RNN’s are used in speech to text.

RNN’s use sequential information to make predictions.

RNN’s are multiple Feed forward neural networks passing information through one another.

They cannot process words. Words must be converted into vectors. Used word embeddings for this.

Example of RNN:

Once we have obtained the correct weights, predicting the next word in the sentence “Napoleon was the Emperor of…” is quite straightforward. Plugging each word at a different time step of the RNN would produce h\_1, h\_2, h\_3, h\_4. We can derive y\_5 using h\_4 and x\_5 (vector of the word “of”). If our training was successful, we should expect that the index of the largest number in y\_5 is the same as the index of the word “France” in our vocabulary.

Drawback of the above example (simplest RNN problem) is vanishing gradient problem.

The network experiences difficulty in memorising words from far away in the sequence and makes predictions based on only the most recent ones.