



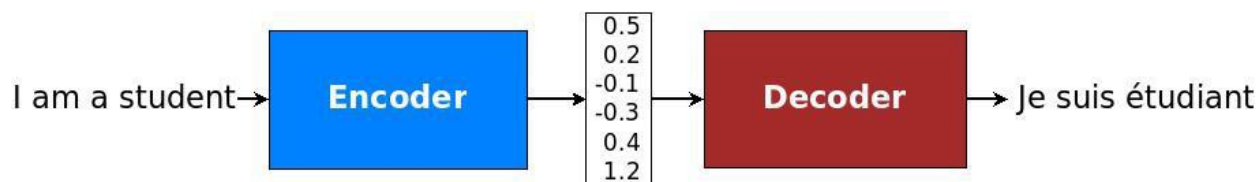
**Data Science
Academy**

www.datascienceacademy.com.br

Processamento de Linguagem Natural

Compreendendo o Encoder/Decoder

Vamos compreender o papel do Encoder e do Decoder no NMT. Esse conceito foi estudado em detalhes no curso Deep Learning II, no capítulo sobre Autoencoders.



Encoder

A tarefa do codificador é fornecer uma representação da sentença de entrada. A sentença de entrada é uma sequência de palavras, para a qual primeiro consultamos a embedding matrix. Então, como no modelo de linguagem básica descrito anteriormente, processamos essas palavras com uma rede neural recorrente. Isso resulta em estados ocultos que codificam cada palavra com seu contexto esquerdo, ou seja, todas as palavras precedentes. Para também obter o contexto certo, também construímos uma rede neural recorrente que corre da direita para a esquerda, ou mais precisamente, do final da frase para o começo. Ter duas redes neurais recorrentes sendo executadas em duas direções é chamado de rede neural recorrente bidirecional.

Decoder

O decodificador é uma rede neural recorrente. É necessária alguma representação do contexto de entrada (mais sobre isso no próximo item sobre o mecanismo de atenção) e o estado oculto anterior e a previsão da palavra de saída, e gera um novo estado de decodificador oculto e uma nova previsão de palavra de saída.

Se usarmos LSTMs para o codificador, então também usamos LSTMs para o decodificador. Do estado oculto, agora prevemos a palavra de saída. Essa previsão assume a forma de uma distribuição de probabilidade sobre todo o vocabulário de saída. Se temos um vocabulário de, digamos, 50.000 palavras, então a predição é um vetor dimensional de 50.000, cada elemento correspondendo à probabilidade prevista para uma palavra no vocabulário.