



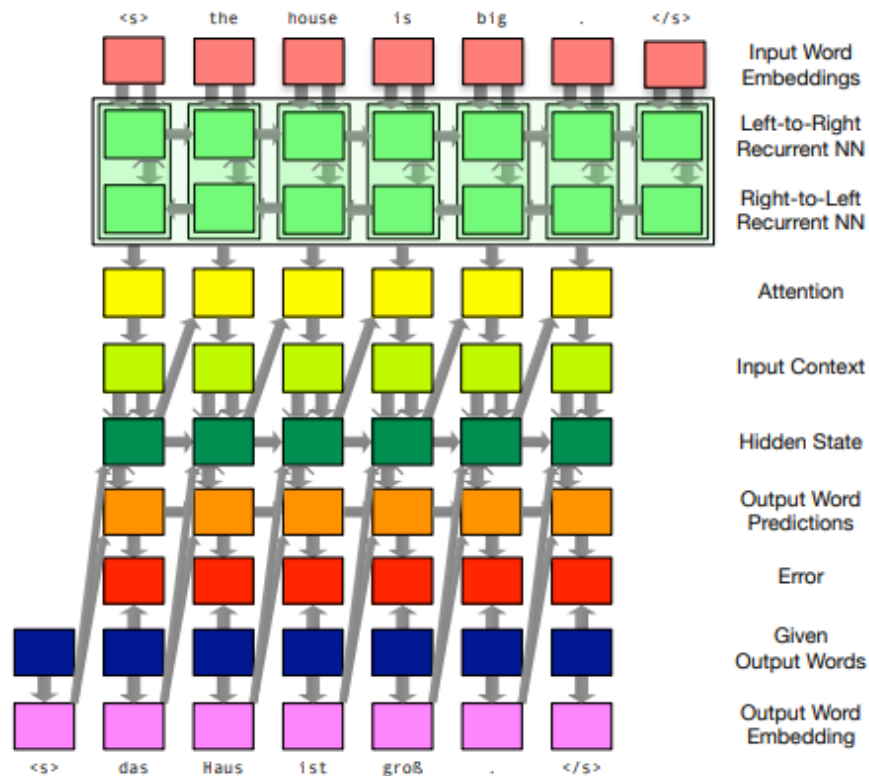
**Data Science
Academy**

www.datascienceacademy.com.br

Processamento de Linguagem Natural

Mecanismo de Atenção

Atualmente, temos duas extremidades soltas. O codificador nos deu uma sequência de representações de palavras $h_j = (\leftarrow - h_j, - \rightarrow h_j)$ e o decodificador espera um contexto c_i em cada passo i . Agora descrevemos o mecanismo de atenção que une essas extremidades. O mecanismo de atenção é difícil de visualizar usando nossos gráficos típicos de rede neural, mas a figura abaixo dá pelo menos uma ideia do que são as relações de entrada e saída.



O mecanismo de atenção é informado por todas as representações de palavras de entrada ($\leftarrow - h_j, - \rightarrow h_j$) e pelo estado oculto anterior do decodificador s_{i-1} , e produz um estado de contexto c_i . A motivação é que queremos calcular uma associação entre o estado do decodificador (que contém informações onde estamos na produção da sentença de saída) e cada palavra de entrada. Com base em quão forte é essa associação, ou em outras palavras, o quanto cada palavra de entrada é relevante para produzir a próxima palavra de saída, queremos ponderar o impacto da representação de palavras.