

ML Week 10 - Attention + Guest

Recap

Weeks 1-9 ML - NN, other, data preprocessing

Attention

Attention - weights \times data

Soft & hard Attention

$$A = g(a(k, Q)) \times V$$

a = Alignment function

g = Distribution function

Xu et al 2015 - Show, ~~text~~ Attend & tell

uses Model structure:

CNN encoder \rightarrow LSTMNN Decoder

Features From images \rightarrow text sequences from features