

ASSIGNMENT SET XI

Amir Mehrpanah

02/11/2020

Assignment 11 of
Deep Learning Course Given by
Prof. H. Haji Abol Hassan

written in L^AT_EX

Observations After First Project on NLP

building a dictionary of words is the first step in an NLP project, which involves choosing a hyperparameter, considerably impacts the results but is not considered among model hyperparameters. The size of dictionary is not necessarily equal to the document corpus dictionary. I obtained good classification accuracy using not more than 1024 or 2048 words.

There are some practical subtleties that usually remains unnoticed during lectures for example sentence lengths should be equal in a batch hence some kind of padding is necessary. The second step is word embedding which is essentially a transformation from words' space to a vector space.

Recurrent models turned out to be very memory consuming and it may become problematic when working with data other than text.