Neural Networks: NLP Applications

COSC 6336: Natural Language Processing
Spring 2020

Some content in these slides has been adapted from Greg Durret.

UNIVERSITY of HOUSTON

Previous Lecture

- ★ Neural network intro
  - ○ Perceptron training
  - ○ Backpropagation

Today

- ★ Example application of using NNs for NLP problems

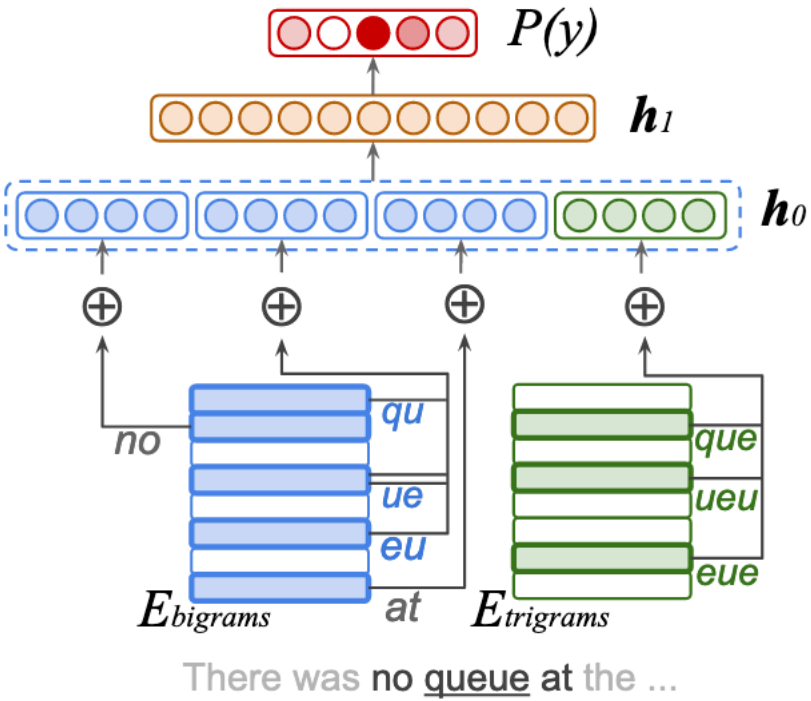
Example Tasks using NNs

- ★ Natural Language Processing with Small Feed-Forward Networks (Botha et al., 2017)
  - Motivation: large models are not suitable for light weight devices

Tricks to reduce memory and computational

- Small feature embeddings trained from scratch
- Quantization of parameters
- Word clusters using bloom maps (Talbot and Talbot, 2008)

Model and results



Model	Micro F1	Size
Baldwin and Lui (2010): NN	90.2	-
Baldwin and Lui (2010): NP	87.0	-
Small FF, 6 dim	87.3	334 KB
Small FF, 16 dim	88.0	800 KB
Small FF, 16 dim, *quantized*	88.0	302 KB

Table 1: Language Identification. Quantization allows trading numerical precision for larger embeddings. The two models from Baldwin and Lui (2010) are the nearest neighbor (NN) and nearest prototype (NP) approaches.

Model	Acc.	Wts.	MB	Ops.
Gillick et al. (2016)	95.06	900k	-	6.63m
Small FF	94.76	241k	0.6	0.27m
+Clusters	95.56	261k	1.0	0.31m
$\frac{1}{2}$ Dim.	95.39	143k	0.7	0.18m

Table 2: POS tagging. Embedded word clusters improves accuracy and allows the use of smaller embedding dimensions.