

WRITEUP

Daksh Sinha

DELIVERABLE 3.1

	they	can	can	fish	END
Start	n/a	n/a	n/a	n/a	n/a
Noun	-2	+0	+0	-15	n/a
Verb	-13	-6	+1	-6	n/a
End	n/a	n/a	n/a	n/a	-17

DELIVERABLE 8.1

Bake-off Design

Embeddings: Polyglot embeddings

Model: Embedding layer -> LSTM layers (1 hidden layer for English, 3 for Norwegian) -> Linear layer -> Softmax

I found increasing the number of LSTM hidden layers or linear layers decreased the dev set accuracy. Also tried several activation functions between the LSTM layer and the linear layer, but they also reduced the accuracy.

Hyperparameters: Learning rate = 0.09, Momentum = 0.0

Special Rules: If the first letter of the word is uppercase, set the emission probability from NOUN to 1 (except for the first word in the sentence)

DELIVERABLE 8.2

Started on the weekend after the assignment was released. Spent approximately 60 hours on it, including the time spent watching the models train.

Difficulty wise, I'd say the assignment was a 3.5-4. I was stuck on the forward algorithm for CRF for a while, but it was a careless oversight rather than a conceptual mistake. I learned a lot from this assignment and coding up the Viterbi algorithm and the BiLSTM-CRF model was a challenging and intriguing task!