

1. Learn CRF model

— Understand basic algorithms for CRF & Implement yourself

— ~~Learning Problem~~

$$W \leftarrow W + \lambda (\Delta F)$$

$$F \rightarrow g \rightarrow W_i f_i$$

Learning is done using inference algorithm

1. Inference algorithm - viterbi

2. Gradient of log-linear model

3. Forward backward vectors

4. SGD

5. Collins Perceptron

6. Gibbs Sampling

7. Contrastive divergence

} 2 of them

— Tags

— No punctuation SPACE

— START & STOP tags

— Design and implement a process for generating features functions automatically from high level specs such as words with -ing, interrogative words, exclamation words.

— Feature function set should be large enough. \Rightarrow Covers almost all english. A small function should implement it.

— Feature functions $\&$ should have value zero for most word-tag pairs.

— Use this property in both data generation & CRF training

— Implement & do experiments with

— Collins Perceptron — Requires viterbi Only

— SGD, L-BFGS, CD — Requires Gradient

Prove that algs are correct

* For each of these, CRF should be implemented.

* Prove

- Algorithms are correct

- α, β by using z .

- Gradients are correct

- * Use a small random subset of sentences to maximize the speed of the code.

- * Use Matlab Profiler to make code fast.

- * Use a validation set (and not cross validation)

- * Use word level accuracy as a performance metric
i.e. fraction of words for which prediction was correct.

* Validation

* Overfitting

- Regularization

- * Feature scaling - FFs should output a value between $\frac{1}{e}$ & 1