Conditional Random Fields



Andrei Pruteanu
PHD

@andrei_pruteanu

https://sites.google.com/site/andreipruteanu

Overview



Data Preparation

Model Training & Evaluation

Hyperparameter Tuning

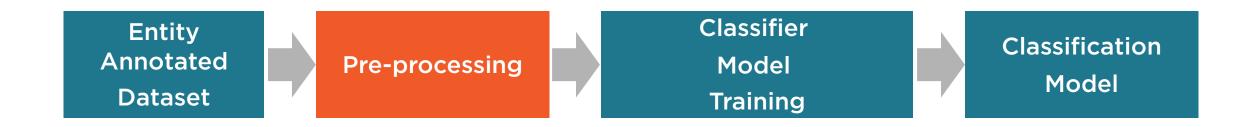
Model Explainability



Data Preparation

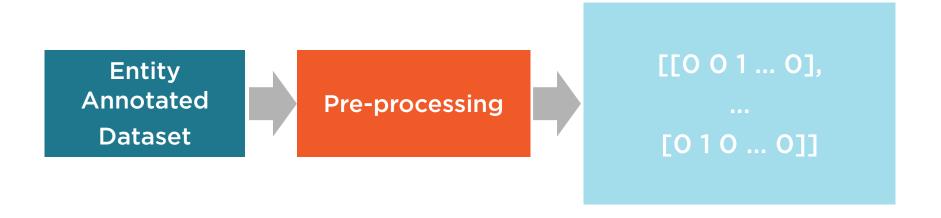


Preprocessing



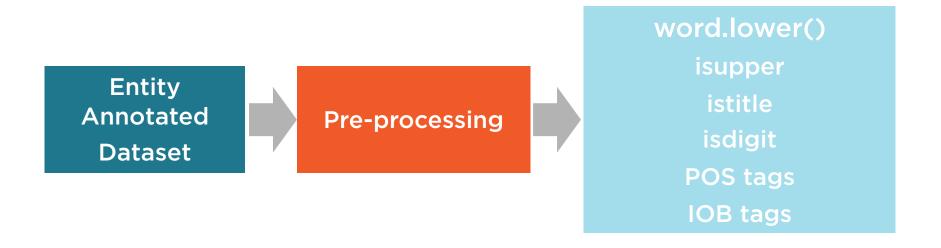


Preprocessing - Standard





Preprocessing - CRF

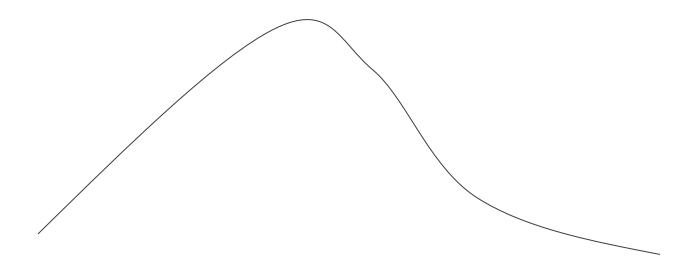




Model Training & Evaluation



BFGS





L-BFGS

Stores only a few vectors

Is the default option in Scikit-learn CRF suite

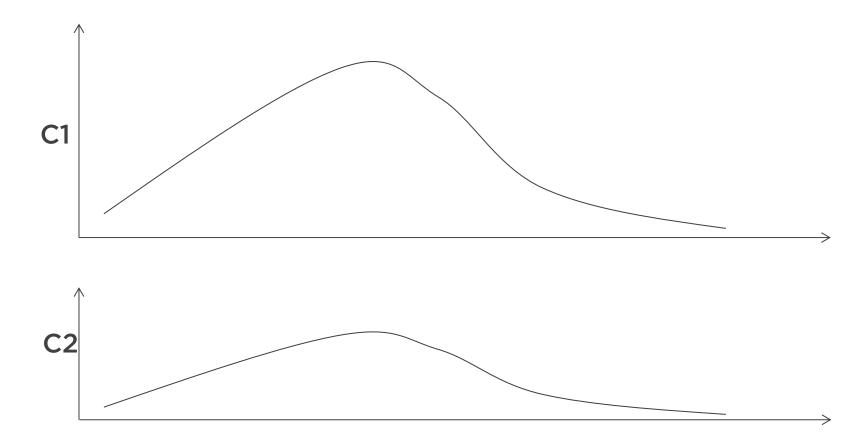
Uses a limited amount of memory Well suited for models with many variables



Hyperparameter Tuning - Parameter Search



Tuning





Hyperparameter Optimization

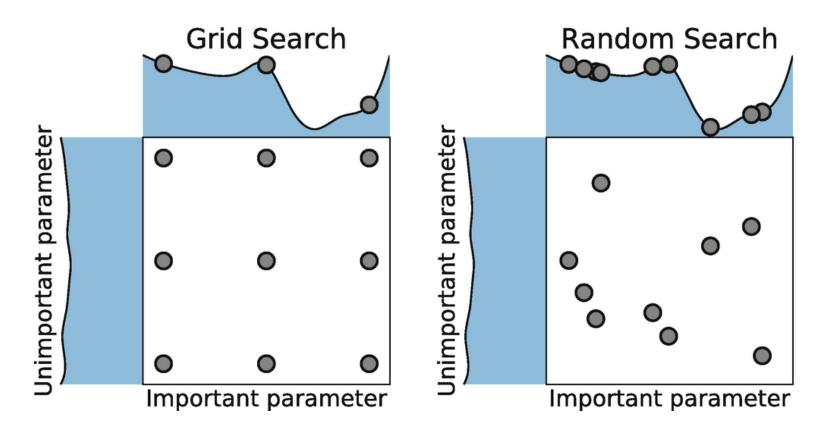


Image Source: Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization.

Journal of Machine Learning Research 13, 281-305 (2012)



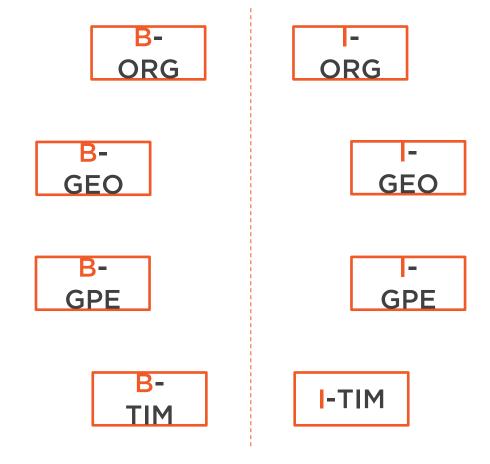
Hyperparameter Tuning - Performance Comparison



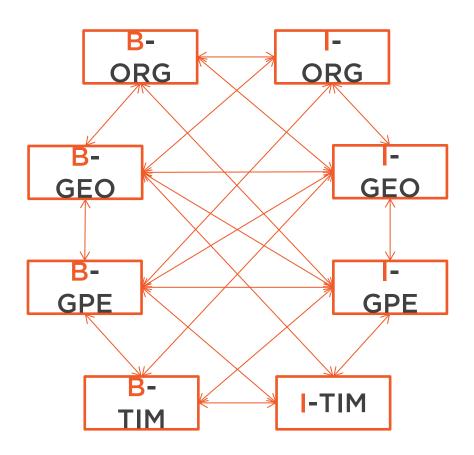
Model Explainability



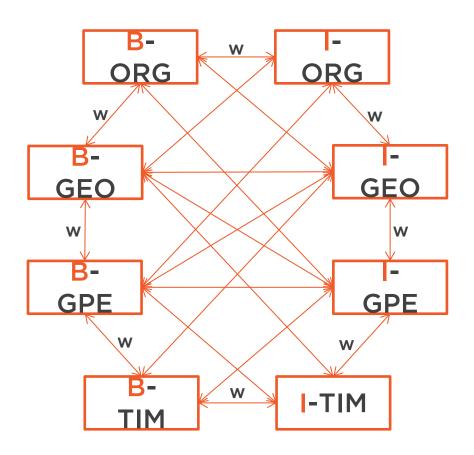
Entity Transitions



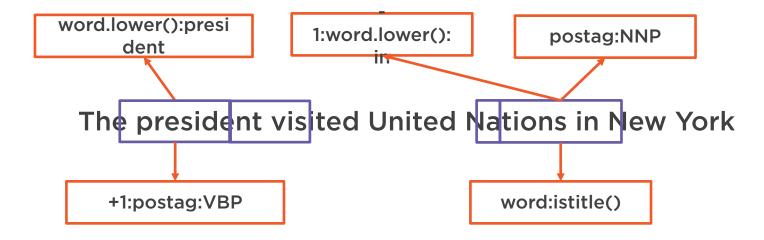
Entity Transitions



Entity Transitions



Features





Model Explainability Visualization



ELI-5

sklearn-crfsuite -> ELI5 allows to check weights of sklearn_crfsuite.CRF models



Conditional Random Fields Remarks

Better performance compared to classic approaches

Hyperparameter tuning brings more performance

The cost for marginal improvement is debatable

Model explainability is important

ELI5 is a powerful model debugging library



Summary



Data Preparation

Model Training & Evaluation

Hyperparameter Tuning

Model Explainability

