hw2-hanz-report

Our goal of this project basically is try to combine different annotators (engines) together and get better F1 scores compared to single engine system. What we supposed to do is to strengthen the strength of each engine and avoid the weakness by using other engine. What I did was that I tested the F1 score on each engine first. And then used the engine with the highest F1 score as base. And then I added other engines like feature selection task. I added other engine one by one, and if the engine had positive result along with the base system, I will keep the engine.

Annotators I tried: Lingpipe, with different models. And Abner.

Single Lingpipe is already very good : it has 80%+ F1 scores. They have both CRF and HMM model and can also add other plugin tools like dictionary and gazetteer.

Abner is a statistical machine learning system using linear-chain conditional random fields (CRFs) with a variety of orthographic and contextual features. I tried to test if Abner can be combined to Lingpipe.

In this project I found out that Abner sometimes will generate long names, which will always hurt the performance. So I tried to constrain on the length of the tokens.