

Cedric Fausey
LING 5802
Homework 3
Part 1 results, without normalization

All results have batch size and pass count of 100.

Turkish (1 layer)

Accuracy	0.47155337
Precision	0.81950517
Recall	0.73698644
F-score	0.77605841

Turkish (2 layers)

Accuracy	0.49840187
Precision	0.81406583
Recall	0.74987064
F-score	0.78065072

Turkish (3 layers)

Accuracy	0.47560196
Precision	0.81063841
Recall	0.73652074
F-score	0.77180425

Turkish (4 layers)

Accuracy	0.43831238
Precision	0.81212947
Recall	0.70883783
F-score	0.75697629

Cedric Fausey
LING 5802
Homework 3
Part 1 results, with normalization

Turkish (1 layer)

Accuracy	0.55273812
Precision	0.82165213
Recall	0.76068508
F-score	0.78999408

Turkish (2 layers)

Accuracy	0.49371404
Precision	0.79640842
Recall	0.74930145
F-score	0.77217586

Turkish (3 layers)

Accuracy	0.49264862
Precision	0.80046298
Recall	0.73357135
F-score	0.76555876

Turkish (4 layers)

Accuracy	0.51587470
Precision	0.79313006
Recall	0.74434057
F-score	0.76795771

Cedric Fausey

LING 5802

Homework 3

Part 1 results for other languages (all batch size 100, pass count 50, 2 hidden layers without normalization)

Spanish

Accuracy	0.85170765
Precision	0.93661574
Recall	0.92152920
F-score	0.92901123

English

Accuracy	0.75458540
Precision	0.87594279
Recall	0.82444050
F-score	0.84941168

Russian

Accuracy	0.46942492
Precision	0.85066149
Recall	0.76144712
F-score	0.80358574

Chinese

Accuracy	0.49052651
Precision	0.83520985
Recall	0.54711273
F-score	0.66029530

German

Accuracy	0.43984857
Precision	0.82652977
Recall	0.71679754
F-score	0.76776262

Part 2 – analysis

Here are the results of Homework 2's program (edited to use Adam instead of SGD) on the Turkish data sets, for comparison:

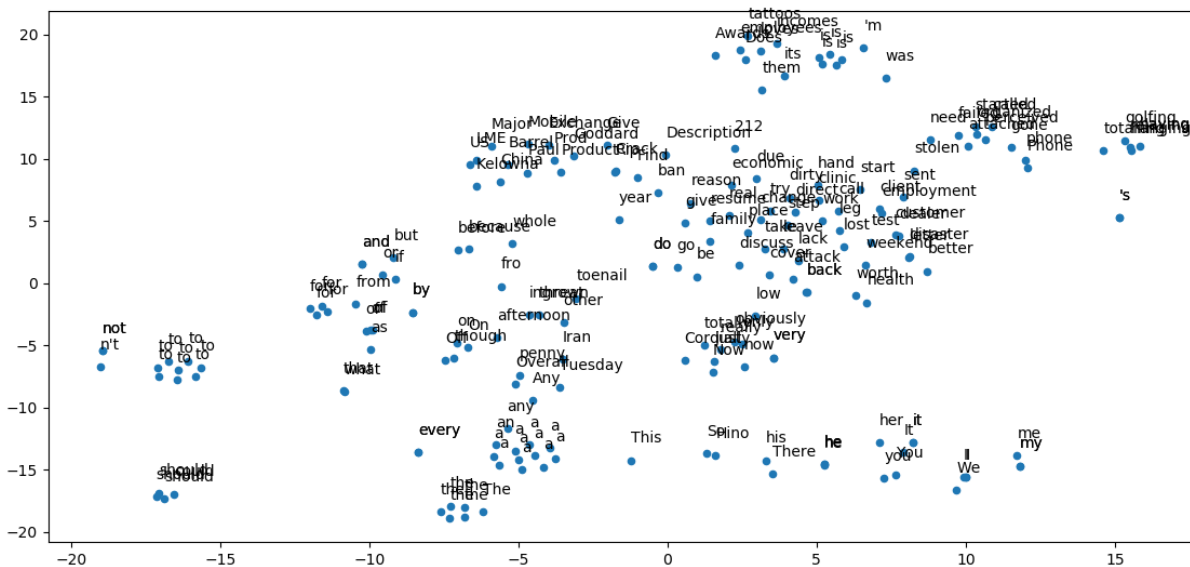
Accuracy	0.11080332
Precision	0.43434028
Recall	0.64948773
F-score	0.52055987

Hidden layers are a significant improvement on the prior Turkish results, which is good because the Turkish data file did not output good results with Homework 2's code. The number of hidden layers and the presence of normalization layers did not have much of an additional affect: any number of hidden layers, with or without normalization, had about the same effect. The same goes for all the languages in homework 2; every single one had significantly better results with hidden layers added.

A notable issue with the classifier is that it typically fails to guess tags that occur less frequently, and assumes they don't occur at all, and similarly outputs a lot of false positives with the tags that occur the most frequently. A possible way to fix this issue would be to adjust the probabilities' values so that the highest probability of the tag is converted to 1 and the lowest is converted to 0, which will change the probability of a tag depending how frequent it is in the data set. The program does this naturally if run with a larger number of passes, but it is inefficient compared to the methods in the second homework.

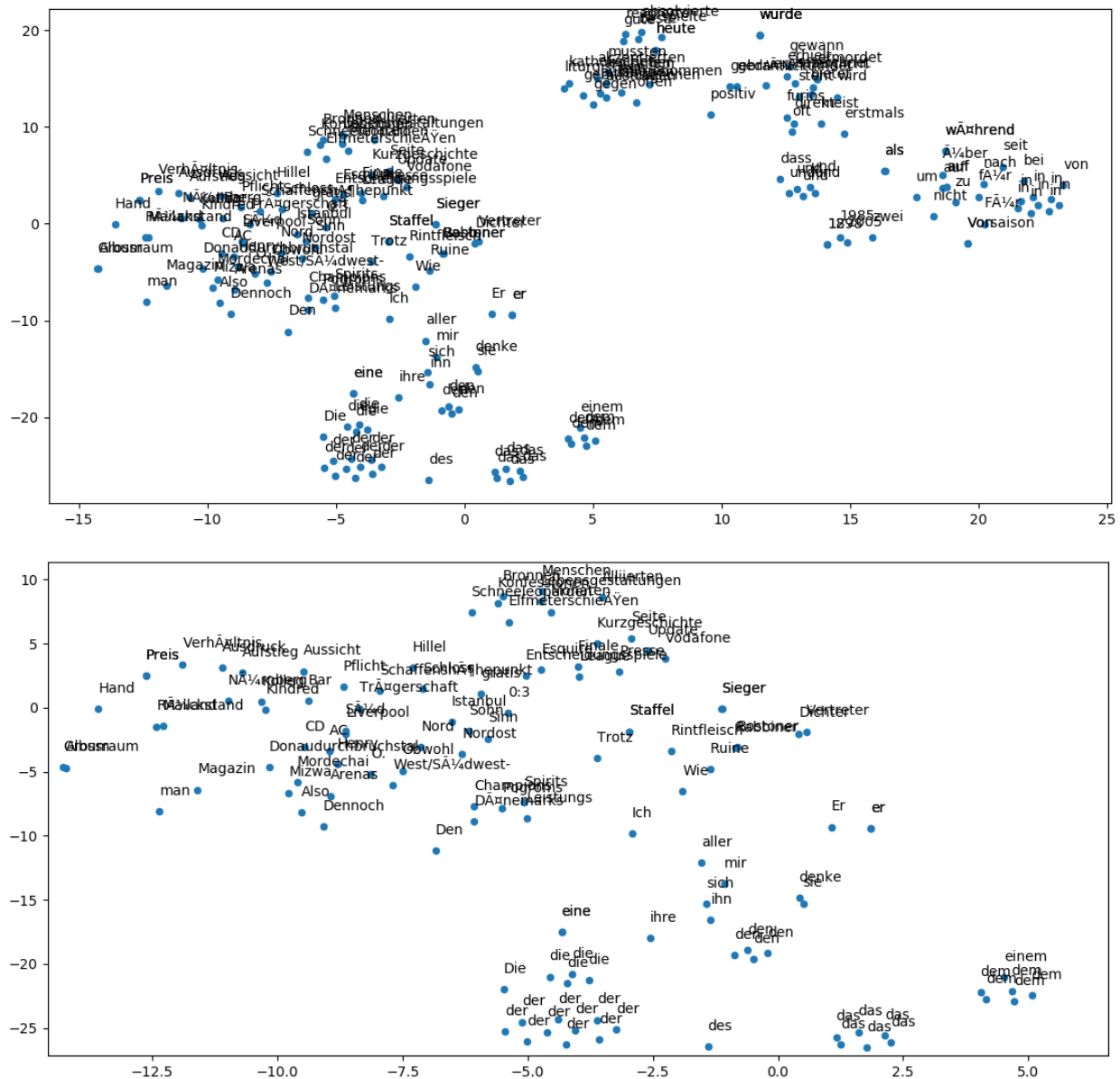
Part 3 – visualization

English data, batch size 100, pass count 50, 2 hidden layers:

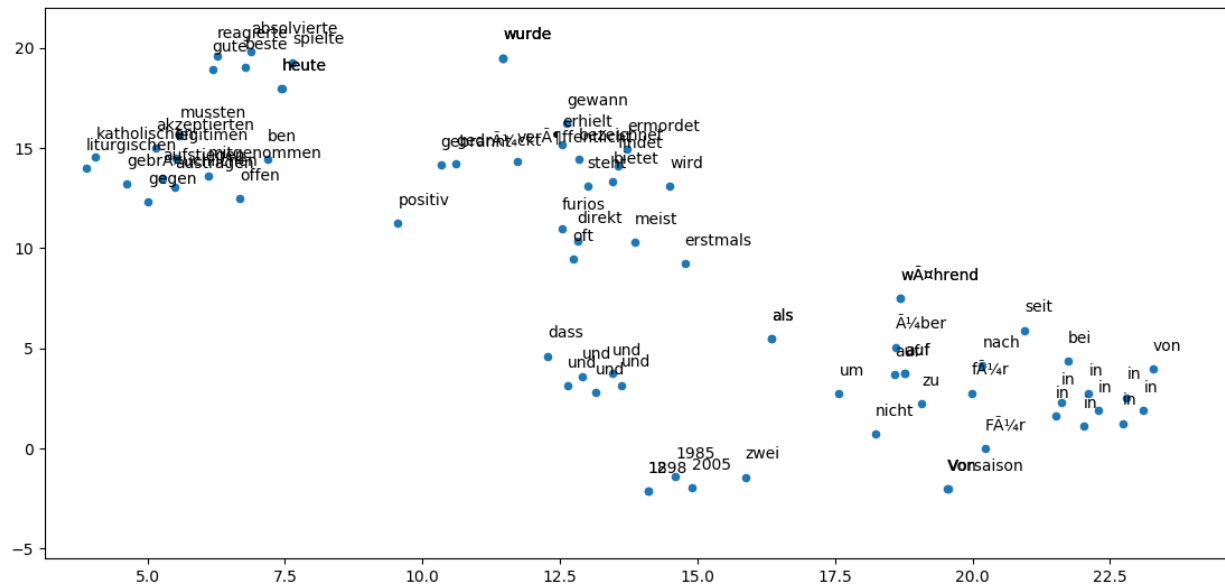


Homework 3

German data, batch size 100, pass count 50, 2 hidden layers (full chart, then two zoom ins):



Cedric Fausey
LING 5802
Homework 3



These charts demonstrate that German suits itself quite a bit better to this form of neural network analysis than English does, because of the consistent conjugation and inflection of its words. The words are more concentrated into clusters in the German diagram than the English one.