

### **Average perplexity of smoothing models:**

LM1:

“Pride and Prejudice - Jane Austen.txt” + kneser ney smoothing  
train:3.87704456037714  
test: 680.364300526354

LM2:

“Pride and Prejudice - Jane Austen.txt” + witten bell smoothing  
train:2.886633347800083  
test:76.53710156395842

LM3:

“corpus/Ulysses - James Joyce.txt”+ kneser ney smoothing  
train:4.573136986370531  
test:205.30112811534445

LM4:

“corpus/Ulysses - James Joyce.txt”+witten bell smoothing  
train:3.6422831162616753  
test:102.40780908526048

It can be observed that witten bell smoothing performs better than the kneser ney smoothing. We also observe less perplexity values on test for models trained on large set of data.

### **Average perplexity on neural language model:**

LM5:

“Pride and Prejudice - Jane Austen.txt”+neural language model  
train:0.9968153407574312  
test:0.9687357257097478

LM6:

“corpus/Ulysses - James Joyce.txt”+neural language model

train:0.9980476811337683

Test:0.9998586245388407

We can observe that neural language model gives better results(less perplexity) for both train and test data compared to smoothing techniques