REPORT

	EuroParl corpus		Medical Abstracts corpus	
	Test	Train	Test	Train
4-gram LM + Kneyser-Ney smoothing	192.595047448851	27.6364217256442	407.533385373540	5.0401232088332
4-gram LM + Witten-Bell smoothing	54.898864996017	51.4024367838669	57.0019975820193	3.613325311967627

The Witten bell smoothing algorithm generally seems to be better at predicting a valid sentence. Due to its method of handling new words by predicting the probability based on the words appearing with a context rather than contexts appearing with words(which cannot be found out for new words) as seen in Kneyser-Ney Witten bell model seems to perform significantly better.
Witten bell smoothing seems to perform bad when the sentence length is low. Due to its hig dependency on context count the model gives relatively high score on the EuroParl train data due to presence of large number of single word sentences such as “Applause”, “Cheers” and “.” – Punctation sentences. However when short sentences are absent as seen in Medical Abstracts corpus the model performs extremely good. Knesyser Ney on the other hand seems to perform better when the length of the sentences is small.