In this homework, I tried to model n-gram language model and calculate their perplexity. After calculation of their perplexity without smoothing, I used add-k smoothing and linear interpolation methods for smoothing and calculate their perplexity, over again and results are given below in the tables. As it is seen from the Table1, best result is obtained by linear interpolation on test data set; however, when applied add-k-smoothing on development set on bigram model, I obtained lower score.

By Table2 and Table3, it can be said that when we set *k* with lower values, we get lower perplexity. Also, bigram models give lower perplexity while unigram models yield higher perplexity. So, we can say that when we choose *n>1,* we can have more meaningful sentences.

The result of the linear interpolation on development set is given. By comparing the results, I chose the best (λ_1, λ_2, λ_3) value and I used these value on test set. The result of it is given in Table 5.

Table1: Perplexity values with the best score

Perplexity	unsmoothed	Add-k-smoothing	Linear interpolation on test data
Unigram	1026.242	1026.24	276.58617
Bigram		0.201	
Trigram		2625.332	

Table2: Perplexity values with different k values

Perplexity	Add-k-smoothing (k values)						
	1e-07	1e-06	1e-05	0.0001	0.01	0.1	1.0
Unigram	1026.24	873806628.18	86058614310.12	8632585725.93	86672503.13	8706285.56	895203.12
Bigram	0.201	0.485	2.193	9.912	202.208	900.795	3634.510
Trigram	2625.332	2627.345	2646.229	2776.273	4511.466	6873.002	10339.404

Table3: Perplexity values with different k values

Perplexity	Add-k-smoothing (k values)						
	2e-07	2e-06	2e-05	0.0002	0.02	0.2	2.0
Unigram	1026.24	858014534.61	43080058386.51	4318721088.01	43373216.86	4368315.27	460611.66
Bigram	0.264	0.764	3.453	15.609	317.907	1397.064	5232.007
Trigram	2625.578	2629.574	2665.111	2665.111	5087.449	7838.735	11364.277

Table4: Perplexity values with different (λ_1, λ_2, λ_3) on development set

Linear Interpolation (λ_1, λ_2, λ_3)	Perplexity
(0.001, 0.009, 0.99)	449.16572
(0.3, 0.3, 0.4)	280.12654
(0.6, 0.3, 0.1)	442.65786
(0.99, 0.009, 0.001)	7858.52449
(0.1, 0.2, 0.7)	275.14091
(0.2, 0.2, 0.6)	276.36418

Table5: Perplexity values with different $(\lambda_1, \lambda_2, \lambda_3)$ on test set

Linear Interpolation $(\lambda_1, \lambda_2, \lambda_3)$	Perplexity
(0.3, 0.3, 0.4)	280.60572
(0.1, 0.2, 0.7)	276.58617
(0.2, 0.2, 0.6)	277.56260

P.S. I had some personal problems, that's why I couldn't do this homework well. There are lots of missing parts in the homework. But I chose to send it in this way rather than sending nothing. I apologize for that, I will do my best for the other two homework.