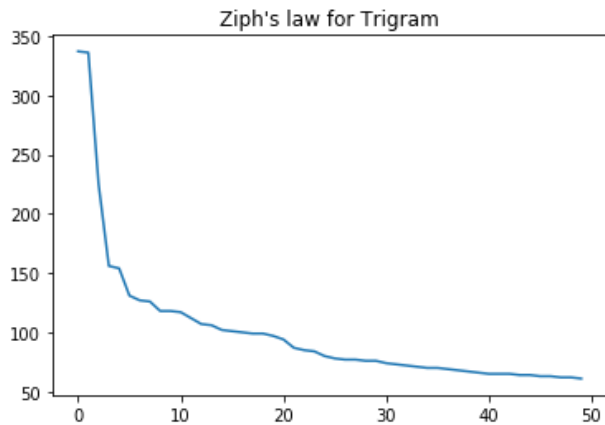
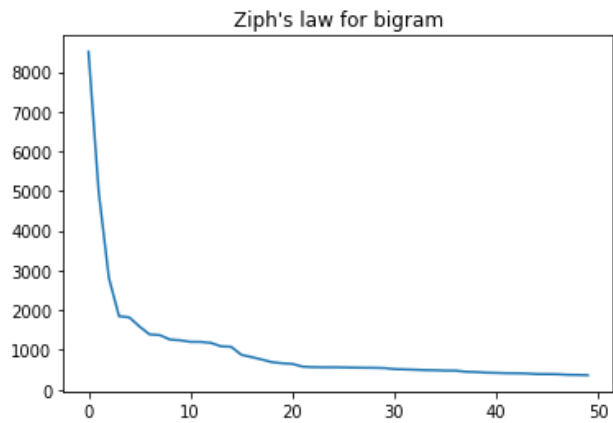
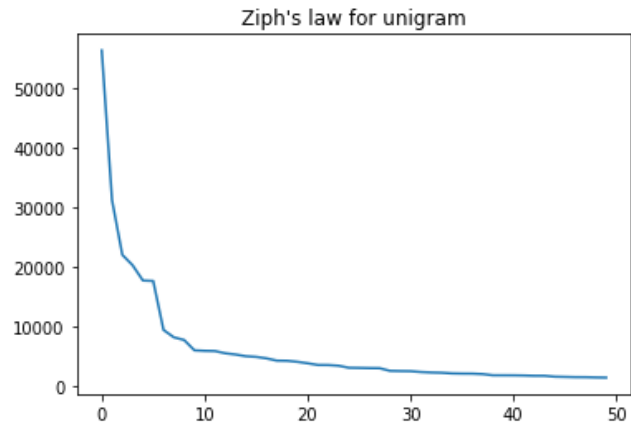


Language Modelling (Assignment 1)

Talluri Surya Teja
15EE35028

Task 1

- Ziph's law:



- Top 10 unigrams:

the	56448
</s>	40000
<s>	40000
of	31276
and	22092
to	20341
a	17780
in	17705
is	9474
that	8240

Top 10 Bigrams:

of the	8508
<s> the	5798
in the	4985
to the	2819
and the	1848
on the	1821
for the	1591
<s> in	1585
<s> it	1816
it is	1390

Top 10 Trigrams:

<s> <s> the	5798
<s> <s> in	1585
<s> <s> it	1516
<s> <s> he	1377
<s> <s> this	1052
<s> <s> but	1038
<s> <s> a	955
<s> <s> and	831
<s> <s> i	675
<s> <s> they	590

- Results:

For Unigram

For sentence.. he lived a good life

probability: 6.31579394802e-15 log-likelihood: -14.1995720472 Perplexity: 691.69463863

For sentence.. the man was happy

probability: 3.83848564498e-11 log-likelihood: -10.4158400792 Perplexity: 401.753824571

For sentence.. the person was good

probability: 7.51041309361e-11 log-likelihood: -10.1243361749 Perplexity: 339.691002502

For sentence.. the girl was sad

probability: 1.50611135001e-12 log-likelihood: -11.8221429186 Perplexity: 902.683968749

For sentence.. he won the war

probability: 2.00987358508e-11 log-likelihood: -10.6968312575 Perplexity: 472.288982341

For Bigram

For sentence.. he lived a good life

probability: 2.37783822206e-12 log-likelihood: -11.6238176963 Perplexity: 211.233861956

For sentence.. the man was happy

probability: 2.88943840323e-10 log-likelihood: -9.53918655935 Perplexity: 242.547409099

For sentence.. the person was good

probability: 1.70995361896e-11 log-likelihood: -10.7670156693 Perplexity: 491.760812118

For sentence.. the girl was sad

probability: 1.70995361896e-11 log-likelihood: Inf Perplexity:Inf

For sentence.. he won the war

probability: 8.28700074926e-10 log-likelihood: -9.08160262189 Perplexity: 186.380578535

For Trigram

For sentence.. he lived a good life

probability: 8.28700074926e-10 log-likelihood: Inf Perplexity:Inf

For sentence.. the man was happy

probability: 8.28700074926e-10 log-likelihood: Inf Perplexity:Inf

For sentence.. the person was good

probability: 8.28700074926e-10 log-likelihood: Inf Perplexity:Inf

For sentence.. the girl was sad

probability: 8.28700074926e-10 log-likelihood: Inf Perplexity:Inf

For sentence.. he won the war

probability: 8.28700074926e-10 log-likelihood: Inf Perplexity:Inf

Task 2:

For Unigram

For k= 0.0001

For sentence.. he lived a good life

probability: 6.3156424219e-15 log-likelihood: -14.1995824668 Perplexity: 691.697957653

For sentence.. the man was happy

probability: 3.83841299156e-11 log-likelihood: -10.4158482995 Perplexity: 401.755725655

For sentence.. the person was good

probability: 7.51026336065e-11 log-likelihood: -10.1243448334 Perplexity: 339.692695604

For sentence.. the girl was sad

probability: 1.50609039321e-12 log-likelihood: -11.8221489617 Perplexity: 902.687108878

For sentence.. he won the war

probability: 2.00983541843e-11 log-likelihood: -10.6968395047 Perplexity: 472.29122451

For k= 0.001

For sentence.. he lived a good life

probability: 6.3142788826e-15 log-likelihood: -14.1996762402 Perplexity: 691.727828743

For sentence.. the man was happy

probability: 3.83775919093e-11 log-likelihood: -10.4159222796 Perplexity: 401.772835338

For sentence.. the person was good

probability: 7.50891593486e-11 log-likelihood: -10.1244227577 Perplexity: 339.707933486

For sentence.. the girl was sad

probability: 1.50590180011e-12 log-likelihood: -11.8222033476 Perplexity: 902.71536978

For sentence.. he won the war

probability: 2.00949196076e-11 log-likelihood: -10.6969137269 Perplexity: 472.311403944

For k= 0.01

For sentence.. he lived a good life

probability: 6.30066346575e-15 log-likelihood: -14.2006137165 Perplexity: 692.026528598

For sentence.. the man was happy

probability: 3.83122910912e-11 log-likelihood: -10.4166618762 Perplexity: 401.943924965

For sentence.. the person was good

probability: 7.49545858027e-11 log-likelihood: -10.1252017914 Perplexity: 339.860308933

For sentence.. the girl was sad

probability: 1.50401765742e-12 log-likelihood: -11.822747065 Perplexity: 902.997953916

For sentence.. he won the war

probability: 2.00606155941e-11 log-likelihood: -10.697655744 Perplexity: 472.513189843

For k= 0.1

For sentence.. he lived a good life

probability: 6.16647766171e-15 log-likelihood: -14.2099628374 Perplexity: 695.01242304

For sentence.. the man was happy

probability: 3.76671290546e-11 log-likelihood: -10.42403748 Perplexity: 403.654101792

For sentence.. the person was good

probability: 7.36255821705e-11 log-likelihood: -10.1329712581 Perplexity: 341.383726303

For sentence.. the girl was sad

probability: 1.48535363493e-12 log-likelihood: -11.8281701366 Perplexity: 905.821311355

For sentence.. he won the war

probability: 1.97217093176e-11 log-likelihood: -10.7050554466 Perplexity: 474.53020513

For k= 1

For sentence.. he lived a good life

probability: 5.000612657e-15 log-likelihood: -14.3009767842 Perplexity: 724.76190333

For sentence.. the man was happy

probability: 3.19281231757e-11 log-likelihood: -10.4958266098 Perplexity: 420.684637146

For sentence.. the person was good

probability: 6.18516058945e-11 log-likelihood: -10.20864902 Perplexity: 356.584330571

For sentence.. the girl was sad

probability: 1.31511720174e-12 log-likelihood: -11.8810355416 Perplexity: 933.810784827

For sentence.. he won the war

probability: 1.67080125936e-11 log-likelihood: -10.7770752061 Perplexity: 494.616729534

For Bigram

For k= 0.0001

For sentence.. he lived a good life

probability: 1.58008582943e-21 log-likelihood: -20.8013193218 Perplexity: 14463.182436

For sentence.. the man was happy

probability: 1.2294491527e-19 log-likelihood: -18.910289428 Perplexity: 53403.824129

For sentence.. the person was good

probability: 1.13480298836e-20 log-likelihood: -19.9450795292 Perplexity: 96887.9757234

For sentence.. the girl was sad

probability: 7.57238354186e-23 log-likelihood: -22.1207673971 Perplexity: 338993.873368

For sentence.. he won the war

probability: 6.43868815942e-19 log-likelihood: -18.1912026083 Perplexity: 35302.109042

For k= 0.001

For sentence.. he lived a good life

probability: 2.22934374909e-22 log-likelihood: -21.6518229612 Perplexity: 21397.5767224

For sentence.. the man was happy

probability: 8.24147478332e-20 log-likelihood: -19.0839950659 Perplexity: 59019.9403814

For sentence.. the person was good

probability: 3.74391660904e-21 log-likelihood: -20.4266738332 Perplexity: 127840.499336

For sentence.. the girl was sad

probability: 2.19866352726e-23 log-likelihood: -22.6578412279 Perplexity: 461806.782402

For sentence.. he won the war

probability: 1.41883121111e-19 log-likelihood: -18.8480692666 Perplexity: 51524.9188731

For k= 0.01

For sentence.. he lived a good life

probability: 3.24713855941e-24 log-likelihood: -23.4884991791 Perplexity: 49853.9800796

For sentence.. the man was happy

probability: 1.3374480405e-20 log-likelihood: -19.8737230812 Perplexity: 92988.8271985

For sentence.. the person was good

probability: 3.18534759617e-22 log-likelihood: -21.496843169 Perplexity: 236706.831581

For sentence.. the girl was sad

probability: 2.59082348343e-24 log-likelihood: -23.5865621751 Perplexity: 788207.465164

For sentence.. he won the war

probability: 9.86706241699e-21 log-likelihood: -20.0058121245 Perplexity: 100335.133102

For k= 0.1

For sentence.. he lived a good life

probability: 3.93006333315e-27 log-likelihood: -26.4056004509 Perplexity: 191038.144011

For sentence.. the man was happy

probability: 2.48623706383e-22 log-likelihood: -21.6044574635 Perplexity: 251834.001302

For sentence.. the person was good

probability: 5.57308739996e-24 log-likelihood: -23.2539041461 Perplexity: 650842.697677

For sentence.. the girl was sad

probability: 1.82997270951e-25 log-likelihood: -24.7375553869 Perplexity: 1528934.38105

For sentence.. he won the war

probability: 1.47551588312e-22 log-likelihood: -21.8310561112 Perplexity: 286922.117836

For k= 1

For sentence.. he lived a good life

probability: 2.46538077359e-30 log-likelihood: -29.6081159952 Perplexity: 834878.349871

For sentence.. the man was happy

probability: 9.0780623742e-25 log-likelihood: -24.0420068376 Perplexity: 1024475.81268

For sentence.. the person was good

probability: 3.35027830407e-26 log-likelihood: -25.4749191151 Perplexity: 2337382.44952

For sentence.. the girl was sad

probability: 6.32850487942e-27 log-likelihood: -26.1986988807 Perplexity: 3545477.39012

For sentence.. he won the war

probability: 5.75917370155e-25 log-likelihood: -24.2396398226 Perplexity: 1147915.59399

For Trigram

For $k=0.0001$

For sentence.. he lived a good life

probability: 5.62434468386e-38 log-likelihood: -37.2499280718 Perplexity: 28182895.7618

For sentence.. the man was happy

probability: 2.74225843469e-31 log-likelihood: -30.561891619 Perplexity: 43699141.4719

For sentence.. the person was good

probability: 1.3432921425e-35 log-likelihood: -34.8718295256 Perplexity: 522344927.24

For sentence.. the girl was sad

probability: 1.1336777668e-31 log-likelihood: -30.9455103705 Perplexity: 54497627.062

For sentence.. he won the war

probability: 1.07198068772e-27 log-likelihood: -26.9698130386 Perplexity: 5526539.26124

For $k=0.001$

For sentence.. he lived a good life

probability: 1.6346322777e-38 log-likelihood: -37.7865799297 Perplexity: 36084108.9918

For sentence.. the man was happy

probability: 5.86743990639e-32 log-likelihood: -31.2315513497 Perplexity: 64252175.5729

For sentence.. the person was good

probability: 1.3432921425e-35 log-likelihood: -34.8718295256 Perplexity: 522344927.24

For sentence.. the girl was sad

probability: 1.32020942703e-32 log-likelihood: -31.8793571705 Perplexity: 93290902.1001

For sentence.. he won the war

probability: 1.31349038456e-29 log-likelihood: -28.8815731021 Perplexity: 16610904.2764

For $k=0.01$

For sentence.. he lived a good life

probability: 2.03714919086e-39 log-likelihood: -38.6909771642 Perplexity: 54726217.5885

For sentence.. the man was happy

probability: 6.63375112737e-33 log-likelihood: -32.1782408256 Perplexity: 110805216.437

For sentence.. the person was good

probability: 1.3432921425e-35 log-likelihood: -34.8718295256 Perplexity: 522344927.24

For sentence.. the girl was sad

probability: 1.35421953896e-33 log-likelihood: -32.8683109244 Perplexity: 164845741.011

For sentence.. he won the war

probability: 1.36691937578e-31 log-likelihood: -30.8642571004 Perplexity: 52007296.1063

For $k=0.1$

For sentence.. he lived a good life

probability: 2.27210310541e-40 log-likelihood: -39.6435719647 Perplexity: 84862220.821

For sentence.. the man was happy

probability: 6.84087549214e-34 log-likelihood: -33.1648883139 Perplexity: 195533907.179

For sentence.. the person was good

probability: 1.3432921425e-35 log-likelihood: -34.8718295256 Perplexity: 522344927.24

For sentence.. the girl was sad

probability: 1.47734802349e-34 log-likelihood: -33.8305171845 Perplexity: 286833119.469

For sentence.. he won the war

probability: 1.62498261755e-33 log-likelihood: -32.7891512803 Perplexity: 157502633.621

For $k = 1$

For sentence.. he lived a good life

probability: 4.1410682511e-41 log-likelihood: -40.3828876116 Perplexity: 119282716.858

For sentence.. the man was happy

probability: 8.05858492971e-35 log-likelihood: -34.0937412127 Perplexity: 333760781.134

For sentence.. the person was good

probability: 1.3432921425e-35 log-likelihood: -34.8718295256 Perplexity: 522344927.24

For sentence.. the girl was sad

probability: 2.68653457976e-35 log-likelihood: -34.570807565 Perplexity: 439240008.481

For sentence.. he won the war

probability: 5.37303602316e-35 log-likelihood: -34.2697802477 Perplexity: 369355918.035

Task 4:

For $\lambda = 0.2$

For sentence.. he lived a good life

probability: 1.0732086659e-13 log-likelihood: -12.9693158292 Perplexity: 392.521243338

For sentence.. the man was happy

probability: 1.21661485434e-09 log-likelihood: -8.9148468853 Perplexity: 169.321349562

For sentence.. the person was good

probability: 2.41687507101e-10 log-likelihood: -9.61674579782 Perplexity: 253.621723478

For sentence.. the girl was sad

probability: 1.05573937824e-09 log-likelihood: -8.9764432793 Perplexity: 175.432810028

For sentence.. he won the war

probability: 3.76308204424e-09 log-likelihood: -8.42445631291 Perplexity: 127.67741412

For $\lambda = 0.5$

For sentence.. he lived a good life

probability: 1.19344118558e-12 log-likelihood: -11.9231989788 Perplexity: 242.459829682

For sentence.. the man was happy

probability: 6.91848307481e-09 log-likelihood: -8.15998911717 Perplexity: 109.64713271

For sentence.. the person was good

probability: 1.42574518904e-09 log-likelihood: -8.8459580852 Perplexity: 162.738204953

For sentence.. the girl was sad

probability: 6.34406984773e-09 log-likelihood: -8.19763204411 Perplexity: 112.049006727

For sentence.. he won the war

probability: 2.41164040854e-08 log-likelihood: -7.61768744787 Perplexity: 80.2457170216

For $\lambda = 0.8$

For sentence.. he lived a good life

probability: 2.62707518982e-12 log-likelihood: -11.580527497 Perplexity: 207.064429144

For sentence.. the man was happy

probability: 5.52207825154e-09 log-likelihood: -8.25789744341 Perplexity: 116.004372942

For sentence.. the person was good

probability: 1.14583988759e-09 log-likelihood: -8.94087606369 Perplexity: 171.877495165

For sentence.. the girl was sad

probability: 4.17517182855e-09 log-likelihood: -8.37932564649 Perplexity: 124.403159854

For sentence.. he won the war

probability: 1.95152712135e-08 log-likelihood: -7.70962540873 Perplexity: 84.6070109289