CS565: Intelligent Systems and Interfaces

Assignment 2

Topics: N-Gram Language Models, Smoothing, Vector Semantics

Roll no. - 170101022

Name – Devansh Gupta

1.N-gram language model

**Linear Interpolation Model Link:**

<https://colab.research.google.com/drive/14FXlU3ZEuNgWwSkSgdRkByUplxVjWdiI?usp=sharing>

**Discounting Model Link:**

<https://colab.research.google.com/drive/1BYK-vNJc0IGBM81yu9AKvmP2V29dUn1m?usp=sharing>

* Corpus of 8642 sentences was considered for both models.
* Training Data – 6999 sentences
* Development Data – 778 sentences
* Training Data – 865 sentences
* Starting Token – ‘S123T’
* Ending Token – ‘E321D’
* Unknown Token – ‘U345K’
* Words having less than 10 frequency in training set is mapped to ‘U345K’
  + - **For Discounting Model (For 5 different training and development sets)**
      * **1st Set of Training and Development**
      * For beta = 0.1
      * M = 21867
      * log likelihood = -139244.25262179424
      * Perplexity = 82.5834243600811
      * For beta = 0.2
      * M = 21867
      * log likelihood = -131652.71208278582
      * Perplexity = 64.92091907207453
      * For beta = 0.3
      * M = 21867
      * log likelihood = -127706.6817432287
      * Perplexity = 57.2877815219459
      * For beta = 0.4
      * M = 21867
      * log likelihood = -125322.54668532494
      * Perplexity = 53.11791672729417
      * For beta = 0.5
      * M = 21867
      * log likelihood = -123878.33281852056
      * Perplexity = 50.7410441162747
      * For beta = 0.6
      * M = 21867
      * log likelihood = -123142.45615992809
      * Perplexity = 49.571153275359926
      * For beta = 0.7
      * M = 21867
      * log likelihood = -123072.85356624042
      * Perplexity = 49.46190572644899
      * For beta = 0.8
      * M = 21867
      * log likelihood = -123830.2740051193
      * Perplexity = 50.66380488500688
      * For beta = 0.9
      * M = 21867
      * log likelihood = -126159.63454401199
      * Perplexity = 54.54622818075785
      * **2nd Set of Training and Development**
      * For beta = 0.1
      * M = 22021
      * log likelihood = -143361.60668433734
      * Perplexity = 91.15326099615137
      * For beta = 0.2
      * M = 22021
      * log likelihood = -135456.41121290493
      * Perplexity = 71.07340754010262
      * For beta = 0.3
      * M = 22021
      * log likelihood = -131336.87490870323
      * Perplexity = 62.42988416792307
      * For beta = 0.4
      * M = 22021
      * log likelihood = -128837.58719310687
      * Perplexity = 57.70679227509784
      * For beta = 0.5
      * M = 22021
      * log likelihood = -127311.58125464234
      * Perplexity = 55.000448392017645
      * For beta = 0.6
      * M = 22021
      * log likelihood = -126517.00796458471
      * Perplexity = 53.6419192024763
      * For beta = 0.7
      * M = 22021
      * log likelihood = -126407.85482625241
      * Perplexity = 53.4579339117612
      * For beta = 0.8
      * M = 22021
      * log likelihood = -127145.98047653053
      * Perplexity = 54.71450133793638
      * For beta = 0.9
      * M = 22021
      * log likelihood = -129488.9728338983
      * Perplexity = 58.90219271187139
      * **3rd Set of Training and Development**
      * For beta = 0.1
      * M = 21710
      * log likelihood = -139059.4329572454
      * Perplexity = 84.76031480517166
      * For beta = 0.2
      * M = 21710
      * log likelihood = -131409.70172422976
      * Perplexity = 66.39290818630587
      * For beta = 0.3
      * M = 21710
      * log likelihood = -127416.43540061083
      * Perplexity = 58.445519122384844
      * For beta = 0.4
      * M = 21710
      * log likelihood = -124987.1978493211
      * Perplexity = 54.08383875103826
      * For beta = 0.5
      * M = 21710
      * log likelihood = -123496.02847784705
      * Perplexity = 51.569273615715765
      * For beta = 0.6
      * M = 21710
      * log likelihood = -122707.8807801069
      * Perplexity = 50.28779457062341
      * For beta = 0.7
      * M = 21710
      * log likelihood = -122575.83955290727
      * Perplexity = 50.07623987874931
      * For beta = 0.8
      * M = 21710
      * log likelihood = -123250.8947106338
      * Perplexity = 51.16724090315917
      * For beta = 0.9
      * M = 21710
      * log likelihood = -125447.96214812965
      * Perplexity = 54.88535135151738
      * **4th Set of Training and Development**
* For beta = 0.1
* M = 21350
* log likelihood = -137076.90688750206
* Perplexity = 85.65490878872905
* For beta = 0.2
* M = 21350
* log likelihood = -129540.00829996995
* Perplexity = 67.06311985979318
* For beta = 0.3
* M = 21350
* log likelihood = -125607.67912652975
* Perplexity = 59.025385728361485
* For beta = 0.4
* M = 21350
* log likelihood = -123217.06872191168
* Perplexity = 54.617496245584086
* For beta = 0.5
* M = 21350
* log likelihood = -121751.195815899
* Perplexity = 52.07908233755648
* For beta = 0.6
* M = 21350
* log likelihood = -120978.3750607095
* Perplexity = 50.788657309265645
* For beta = 0.7
* M = 21350
* log likelihood = -120852.25428983604
* Perplexity = 50.58112195218155
* For beta = 0.8
* M = 21350
* log likelihood = -121521.27214142562
* Perplexity = 51.69177594104701
* For beta = 0.9
* M = 21350
* log likelihood = -123688.58311871924
* Perplexity = 55.460021036329415
  + - * **5th Set of Training and Development**
* For beta = 0.1
* M = 22396
* log likelihood = -145135.4745110167
* Perplexity = 89.28955237534049
* For beta = 0.2
* M = 22396
* log likelihood = -137071.94382860238
* Perplexity = 69.56916650032437
* For beta = 0.3
* M = 22396
* log likelihood = -132856.12354712424
* Perplexity = 61.059177645148345
* For beta = 0.4
* M = 22396
* log likelihood = -130284.09336968376
* Perplexity = 56.38708940643882
* For beta = 0.5
* M = 22396
* log likelihood = -128695.9660231957
* Perplexity = 53.682567393744776
* For beta = 0.6
* M = 22396
* log likelihood = -127842.77778585507
* Perplexity = 52.283586435388706
* For beta = 0.7
* M = 22396
* log likelihood = -127672.19582560609
* Perplexity = 52.00828557945955
* For beta = 0.8
* M = 22396
* log likelihood = -128336.4331707537
* Perplexity = 53.088531321907354
* For beta = 0.9
* M = 22396
* log likelihood = -130566.37676068401
* Perplexity = 56.88187666371085
  + - * **For all the 5 set I found minimum likelihood at beta = 0.7. For test considered**

**beta =0.7.**

* + - * **Test Set Analysis using discounting**
        + For beta = 0.7
        + M = 24240
        + log likelihood = -138103.1591
        + Perplexity = 51.8878
      * **Same Test using Laplace Smoothing**
        + M = 24240
        + log likelihood = -208644.5477
        + Perplexity = 390.03190
    - **For Interpolation Model (For 5 different training and development sets and at last Test Set)**
      * **1st Set of Training and Development**
* For lambda1 = 0.1 lambda2 = 0.1 lambda3 = 0.8
* M = 22047
* log likelihood = -135980.98231391786
* Perplexity = 71.89288663897669
* For lambda1 = 0.1 lambda2 = 0.2 lambda3 = 0.7
* M = 22047
* log likelihood = -132665.91877992352
* Perplexity = 64.7771752146211
* For lambda1 = 0.1 lambda2 = 0.3 lambda3 = 0.6
* M = 22047
* log likelihood = -130380.78739853362
* Perplexity = 60.28660046830073
* For lambda1 = 0.1 lambda2 = 0.4 lambda3 = 0.5
* M = 22047
* log likelihood = -128756.32271423172
* Perplexity = 57.284923191314185
* For lambda1 = 0.1 lambda2 = 0.5 lambda3 = 0.4
* M = 22047
* log likelihood = -127670.94548566578
* Perplexity = 55.36312316943219
* For lambda1 = 0.1 lambda2 = 0.6 lambda3 = 0.30000000000000004
* M = 22047
* log likelihood = -127128.93200884367
* Perplexity = 54.42769261514737
* For lambda1 = 0.1 lambda2 = 0.7 lambda3 = 0.20000000000000007
* M = 22047
* log likelihood = -127294.2416198627
* Perplexity = 54.71130359792055
* For lambda1 = 0.1 lambda2 = 0.8 lambda3 = 0.09999999999999998
* M = 22047
* log likelihood = -128831.87628467311
* Perplexity = 57.42115771996189
* For lambda1 = 0.2 lambda2 = 0.1 lambda3 = 0.7
* M = 22047
* log likelihood = -133379.26212694874
* Perplexity = 66.24635749067228
* For lambda1 = 0.2 lambda2 = 0.2 lambda3 = 0.6
* M = 22047
* log likelihood = -130570.84915396098
* Perplexity = 60.647918544172036
* For lambda1 = 0.2 lambda2 = 0.3 lambda3 = 0.5
* M = 22047
* log likelihood = -128704.19775409461
* Perplexity = 57.19112246070115
* For lambda1 = 0.2 lambda2 = 0.4 lambda3 = 0.3999999999999999
* M = 22047
* log likelihood = -127482.74650274507
* Perplexity = 55.036513069353994
* For lambda1 = 0.2 lambda2 = 0.5 lambda3 = 0.30000000000000004
* M = 22047
* log likelihood = -126857.02471475344
* Perplexity = 53.96439320203747
* For lambda1 = 0.2 lambda2 = 0.6 lambda3 = 0.19999999999999996
* M = 22047
* log likelihood = -126968.89961215638
* Perplexity = 54.15453600927498
* For lambda1 = 0.2 lambda2 = 0.7 lambda3 = 0.10000000000000009
* M = 22047
* log likelihood = -128473.3858491471
* Perplexity = 56.77761058820434
* For lambda1 = 0.3 lambda2 = 0.1 lambda3 = 0.6
* M = 22047
* log likelihood = -132017.32741601768
* Perplexity = 63.46965409235463
* For lambda1 = 0.3 lambda2 = 0.2 lambda3 = 0.5
* M = 22047
* log likelihood = -129527.83592732284
* Perplexity = 58.69141573488818
* For lambda1 = 0.3 lambda2 = 0.3 lambda3 = 0.4
* M = 22047
* log likelihood = -128001.9725958948
* Perplexity = 55.94231441309084
* For lambda1 = 0.3 lambda2 = 0.4 lambda3 = 0.30000000000000004
* M = 22047
* log likelihood = -127196.53821548473
* Perplexity = 54.543502086884985
* For lambda1 = 0.3 lambda2 = 0.5 lambda3 = 0.19999999999999996
* M = 22047
* log likelihood = -127193.08946747308
* Perplexity = 54.537588419458444
* For lambda1 = 0.3 lambda2 = 0.6 lambda3 = 0.10000000000000009
* M = 22047
* log likelihood = -128622.29455052932
* Perplexity = 57.04404489449347
* For lambda1 = 0.4 lambda2 = 0.1 lambda3 = 0.5
* M = 22047
* log likelihood = -131462.73353286248
* Perplexity = 62.37257875680155
* For lambda1 = 0.4 lambda2 = 0.2 lambda3 = 0.3999999999999999
* M = 22047
* log likelihood = -129239.82102772745
* Perplexity = 58.16236087995466
* For lambda1 = 0.4 lambda2 = 0.3 lambda3 = 0.30000000000000004
* M = 22047
* log likelihood = -128089.14243638908
* Perplexity = 56.095838981428685
* For lambda1 = 0.4 lambda2 = 0.4 lambda3 = 0.19999999999999996
* M = 22047
* log likelihood = -127880.52714940195
* Perplexity = 55.72912345582119
* For lambda1 = 0.4 lambda2 = 0.5 lambda3 = 0.09999999999999998
* M = 22047
* log likelihood = -129180.01409003326
* Perplexity = 58.05310085054638
* For lambda1 = 0.5 lambda2 = 0.1 lambda3 = 0.4
* M = 22047
* log likelihood = -131600.1261396222
* Perplexity = 62.64258346690845
* For lambda1 = 0.5 lambda2 = 0.2 lambda3 = 0.30000000000000004
* M = 22047
* log likelihood = -129678.20937759963
* Perplexity = 58.96954629729049
* For lambda1 = 0.5 lambda2 = 0.3 lambda3 = 0.19999999999999996
* M = 22047
* log likelihood = -129089.43058693825
* Perplexity = 57.88800649486367
* For lambda1 = 0.5 lambda2 = 0.4 lambda3 = 0.09999999999999998
* M = 22047
* log likelihood = -130168.0613348195
* Perplexity = 59.8847486061526
* For lambda1 = 0.6 lambda2 = 0.1 lambda3 = 0.30000000000000004
* M = 22047
* log likelihood = -132479.84798343256
* Perplexity = 64.39933646622049
* For lambda1 = 0.6 lambda2 = 0.2 lambda3 = 0.19999999999999996
* M = 22047
* log likelihood = -131035.87937277797
* Perplexity = 61.54112461597707
* For lambda1 = 0.6 lambda2 = 0.3 lambda3 = 0.10000000000000009
* M = 22047
* log likelihood = -131704.23365482243
* Perplexity = 62.847954239437776
* For lambda1 = 0.7 lambda2 = 0.1 lambda3 = 0.20000000000000007
* M = 22047
* log likelihood = -134368.2175858723
* Perplexity = 68.33846401105792
* For lambda1 = 0.7 lambda2 = 0.2 lambda3 = 0.10000000000000009
* M = 22047
* log likelihood = -134084.1571569094
* Perplexity = 67.7308691782336
* For lambda1 = 0.8 lambda2 = 0.1 lambda3 = 0.09999999999999998
* M = 22047
* log likelihood = -138174.68080911564
* Perplexity = 77.02624144079198
  + - * **2nd Set of Training and Development**
* For lambda1 = 0.1 lambda2 = 0.1 lambda3 = 0.8
* M = 22190
* log likelihood = -137572.99945177813
* Perplexity = 73.50521218334536
* For lambda1 = 0.1 lambda2 = 0.2 lambda3 = 0.7
* M = 22190
* log likelihood = -134262.85727047655
* Perplexity = 66.284605942246
* For lambda1 = 0.1 lambda2 = 0.3 lambda3 = 0.6
* M = 22190
* log likelihood = -131979.04481138763
* Perplexity = 61.720641607336255
* For lambda1 = 0.1 lambda2 = 0.4 lambda3 = 0.5
* M = 22190
* log likelihood = -130355.29335391166
* Perplexity = 58.668176037382025
* For lambda1 = 0.1 lambda2 = 0.5 lambda3 = 0.4
* M = 22190
* log likelihood = -129270.56698988062
* Perplexity = 56.71359353716312
* For lambda1 = 0.1 lambda2 = 0.6 lambda3 = 0.30000000000000004
* M = 22190
* log likelihood = -128729.16865494536
* Perplexity = 55.76253969195336
* For lambda1 = 0.1 lambda2 = 0.7 lambda3 = 0.20000000000000007
* M = 22190
* log likelihood = -128894.70630194565
* Perplexity = 56.051628187709966
* For lambda1 = 0.1 lambda2 = 0.8 lambda3 = 0.09999999999999998
* M = 22190
* log likelihood = -130430.7507612405
* Perplexity = 58.80662335734735
* For lambda1 = 0.2 lambda2 = 0.1 lambda3 = 0.7
* M = 22190
* log likelihood = -135065.97344789974
* Perplexity = 67.9685131094139
* For lambda1 = 0.2 lambda2 = 0.2 lambda3 = 0.6
* M = 22190
* log likelihood = -132256.29426020736
* Perplexity = 62.25748946625839
* For lambda1 = 0.2 lambda2 = 0.3 lambda3 = 0.5
* M = 22190
* log likelihood = -130385.90972304178
* Perplexity = 58.72431082945441
* For lambda1 = 0.2 lambda2 = 0.4 lambda3 = 0.3999999999999999
* M = 22190
* log likelihood = -129160.8826816541
* Perplexity = 56.51961394215413
* For lambda1 = 0.2 lambda2 = 0.5 lambda3 = 0.30000000000000004
* M = 22190
* log likelihood = -128531.9787767848
* Perplexity = 55.42012020798398
* For lambda1 = 0.2 lambda2 = 0.6 lambda3 = 0.19999999999999996
* M = 22190
* log likelihood = -128640.69183873292
* Perplexity = 55.6086391356538
* For lambda1 = 0.2 lambda2 = 0.7 lambda3 = 0.10000000000000009
* M = 22190
* log likelihood = -130140.55468382854
* Perplexity = 58.27596009362315
* For lambda1 = 0.3 lambda2 = 0.1 lambda3 = 0.6
* M = 22190
* log likelihood = -133769.1985765013
* Perplexity = 65.27031291855941
* For lambda1 = 0.3 lambda2 = 0.2 lambda3 = 0.5
* M = 22190
* log likelihood = -131275.0857299486
* Perplexity = 60.37824824823918
* For lambda1 = 0.3 lambda2 = 0.3 lambda3 = 0.4
* M = 22190
* log likelihood = -129742.54591819592
* Perplexity = 57.55592554953289
* For lambda1 = 0.3 lambda2 = 0.4 lambda3 = 0.30000000000000004
* M = 22190
* log likelihood = -128930.80510461923
* Perplexity = 56.11486850938417
* For lambda1 = 0.3 lambda2 = 0.5 lambda3 = 0.19999999999999996
* M = 22190
* log likelihood = -128921.3575902791
* Perplexity = 56.098310825487914
* For lambda1 = 0.3 lambda2 = 0.6 lambda3 = 0.10000000000000009
* M = 22190
* log likelihood = -130343.41670002187
* Perplexity = 58.646414764813535
* For lambda1 = 0.4 lambda2 = 0.1 lambda3 = 0.5
* M = 22190
* log likelihood = -133270.87172929765
* Perplexity = 64.26216938818695
* For lambda1 = 0.4 lambda2 = 0.2 lambda3 = 0.3999999999999999
* M = 22190
* log likelihood = -131040.20951032428
* Perplexity = 59.93688565566655
* For lambda1 = 0.4 lambda2 = 0.3 lambda3 = 0.30000000000000004
* M = 22190
* log likelihood = -129880.19570234022
* Perplexity = 57.80393472591666
* For lambda1 = 0.4 lambda2 = 0.4 lambda3 = 0.19999999999999996
* M = 22190
* log likelihood = -129662.61082365144
* Perplexity = 57.4123919400989
* For lambda1 = 0.4 lambda2 = 0.5 lambda3 = 0.09999999999999998
* M = 22190
* log likelihood = -130952.30354676757
* Perplexity = 59.77253004286687
* For lambda1 = 0.5 lambda2 = 0.1 lambda3 = 0.4
* M = 22190
* log likelihood = -133463.02761986016
* Perplexity = 64.64905385290818
* For lambda1 = 0.5 lambda2 = 0.2 lambda3 = 0.30000000000000004
* M = 22190
* log likelihood = -131529.64069283145
* Perplexity = 60.86026044662277
* For lambda1 = 0.5 lambda2 = 0.3 lambda3 = 0.19999999999999996
* M = 22190
* log likelihood = -130928.37249014468
* Perplexity = 59.72786483518923
* For lambda1 = 0.5 lambda2 = 0.4 lambda3 = 0.09999999999999998
* M = 22190
* log likelihood = -131993.99627301435
* Perplexity = 61.74947419766425
* For lambda1 = 0.6 lambda2 = 0.1 lambda3 = 0.30000000000000004
* M = 22190
* log likelihood = -134400.56166924478
* Perplexity = 66.57034067597034
* For lambda1 = 0.6 lambda2 = 0.2 lambda3 = 0.19999999999999996
* M = 22190
* log likelihood = -132940.10686257636
* Perplexity = 63.60162612905961
* For lambda1 = 0.6 lambda2 = 0.3 lambda3 = 0.10000000000000009
* M = 22190
* log likelihood = -133591.00070332474
* Perplexity = 64.90800467747758
* For lambda1 = 0.7 lambda2 = 0.1 lambda3 = 0.20000000000000007
* M = 22190
* log likelihood = -136354.43731834475
* Perplexity = 70.75988145387888
* For lambda1 = 0.7 lambda2 = 0.2 lambda3 = 0.10000000000000009
* M = 22190
* log likelihood = -136045.96603460275
* Perplexity = 70.08133542600282
* For lambda1 = 0.8 lambda2 = 0.1 lambda3 = 0.09999999999999998
* M = 22190
* log likelihood = -140241.25512849912
* Perplexity = 79.89429096366332
  + - * **3rd Set of Training and Development**
* For lambda1 = 0.1 lambda2 = 0.1 lambda3 = 0.8
* M = 20962
* log likelihood = -129318.36212954458
* Perplexity = 71.96285521447591
* For lambda1 = 0.1 lambda2 = 0.2 lambda3 = 0.7
* M = 20962
* log likelihood = -126170.74092145891
* Perplexity = 64.8494342756749
* For lambda1 = 0.1 lambda2 = 0.3 lambda3 = 0.6
* M = 20962
* log likelihood = -124002.07011011652
* Perplexity = 60.36183824257893
* For lambda1 = 0.1 lambda2 = 0.4 lambda3 = 0.5
* M = 20962
* log likelihood = -122460.00426804642
* Perplexity = 57.36106928085884
* For lambda1 = 0.1 lambda2 = 0.5 lambda3 = 0.4
* M = 20962
* log likelihood = -121428.32737326578
* Perplexity = 55.43723827200326
* For lambda1 = 0.1 lambda2 = 0.6 lambda3 = 0.30000000000000004
* M = 20962
* log likelihood = -120910.14156974822
* Perplexity = 54.49542590931786
* For lambda1 = 0.1 lambda2 = 0.7 lambda3 = 0.20000000000000007
* M = 20962
* log likelihood = -121059.04498663986
* Perplexity = 54.764410301070896
* For lambda1 = 0.1 lambda2 = 0.8 lambda3 = 0.09999999999999998
* M = 20962
* log likelihood = -122499.55962372157
* Perplexity = 57.43614497046414
* For lambda1 = 0.2 lambda2 = 0.1 lambda3 = 0.7
* M = 20962
* log likelihood = -127036.26520448623
* Perplexity = 66.73224994213669
* For lambda1 = 0.2 lambda2 = 0.2 lambda3 = 0.6
* M = 20962
* log likelihood = -124357.98479206335
* Perplexity = 61.0764319903127
* For lambda1 = 0.2 lambda2 = 0.3 lambda3 = 0.5
* M = 20962
* log likelihood = -122577.02298542878
* Perplexity = 57.583454452455534
* For lambda1 = 0.2 lambda2 = 0.4 lambda3 = 0.3999999999999999
* M = 20962
* log likelihood = -121409.2320321742
* Perplexity = 55.40224499238083
* For lambda1 = 0.2 lambda2 = 0.5 lambda3 = 0.30000000000000004
* M = 20962
* log likelihood = -120806.01928086526
* Perplexity = 54.308121034903074
* For lambda1 = 0.2 lambda2 = 0.6 lambda3 = 0.19999999999999996
* M = 20962
* log likelihood = -120899.70366135651
* Perplexity = 54.476620130210435
* For lambda1 = 0.2 lambda2 = 0.7 lambda3 = 0.10000000000000009
* M = 20962
* log likelihood = -122305.24243649094
* Perplexity = 57.06827467612124
* For lambda1 = 0.3 lambda2 = 0.1 lambda3 = 0.6
* M = 20962
* log likelihood = -125864.04973052473
* Perplexity = 64.19509959100226
* For lambda1 = 0.3 lambda2 = 0.2 lambda3 = 0.5
* M = 20962
* log likelihood = -123482.50934173631
* Perplexity = 59.33366510652047
* For lambda1 = 0.3 lambda2 = 0.3 lambda3 = 0.4
* M = 20962
* log likelihood = -122019.51299940915
* Perplexity = 56.53162305696834
* For lambda1 = 0.3 lambda2 = 0.4 lambda3 = 0.30000000000000004
* M = 20962
* log likelihood = -121241.03366051416
* Perplexity = 55.09496468581941
* For lambda1 = 0.3 lambda2 = 0.5 lambda3 = 0.19999999999999996
* M = 20962
* log likelihood = -121221.65490715828
* Perplexity = 55.05967146044733
* For lambda1 = 0.3 lambda2 = 0.6 lambda3 = 0.10000000000000009
* M = 20962
* log likelihood = -122553.0945406588
* Perplexity = 57.53791024721636
* For lambda1 = 0.4 lambda2 = 0.1 lambda3 = 0.5
* M = 20962
* log likelihood = -125430.90181475664
* Perplexity = 63.2821970865157
* For lambda1 = 0.4 lambda2 = 0.2 lambda3 = 0.3999999999999999
* M = 20962
* log likelihood = -123297.49249802984
* Perplexity = 58.9717744458125
* For lambda1 = 0.4 lambda2 = 0.3 lambda3 = 0.30000000000000004
* M = 20962
* log likelihood = -122185.73445342263
* Perplexity = 56.84320003063945
* For lambda1 = 0.4 lambda2 = 0.4 lambda3 = 0.19999999999999996
* M = 20962
* log likelihood = -121967.67452610163
* Perplexity = 56.43480324494791
* For lambda1 = 0.4 lambda2 = 0.5 lambda3 = 0.09999999999999998
* M = 20962
* log likelihood = -123173.29222634959
* Perplexity = 58.73007926734154
* For lambda1 = 0.5 lambda2 = 0.1 lambda3 = 0.4
* M = 20962
* log likelihood = -125641.24858897683
* Perplexity = 63.723890978458094
* For lambda1 = 0.5 lambda2 = 0.2 lambda3 = 0.30000000000000004
* M = 20962
* log likelihood = -123788.18204662073
* Perplexity = 59.93642906106061
* For lambda1 = 0.5 lambda2 = 0.3 lambda3 = 0.19999999999999996
* M = 20962
* log likelihood = -123203.70702774065
* Perplexity = 58.78917508203698
* For lambda1 = 0.5 lambda2 = 0.4 lambda3 = 0.09999999999999998
* M = 20962
* log likelihood = -124196.19625777926
* Perplexity = 60.75055555360996
* For lambda1 = 0.6 lambda2 = 0.1 lambda3 = 0.30000000000000004
* M = 20962
* log likelihood = -126550.00382987146
* Perplexity = 65.66783386517122
* For lambda1 = 0.6 lambda2 = 0.2 lambda3 = 0.19999999999999996
* M = 20962
* log likelihood = -125143.32637680072
* Perplexity = 62.68328535472616
* For lambda1 = 0.6 lambda2 = 0.3 lambda3 = 0.10000000000000009
* M = 20962
* log likelihood = -125740.6872846732
* Perplexity = 63.93376788395778
* For lambda1 = 0.7 lambda2 = 0.1 lambda3 = 0.20000000000000007
* M = 20962
* log likelihood = -128414.86264502283
* Perplexity = 69.84470026445165
* For lambda1 = 0.7 lambda2 = 0.2 lambda3 = 0.10000000000000009
* M = 20962
* log likelihood = -128095.95503853388
* Perplexity = 69.11203902423333
* For lambda1 = 0.8 lambda2 = 0.1 lambda3 = 0.09999999999999998
* M = 20962
* log likelihood = -132102.93812618393
* Perplexity = 78.90362931642275
  + - * **4th Set of Training and Development**
* For lambda1 = 0.1 lambda2 = 0.1 lambda3 = 0.8
* M = 22185
* log likelihood = -135999.87709076988
* Perplexity = 70.04754867266492
* For lambda1 = 0.1 lambda2 = 0.2 lambda3 = 0.7
* M = 22185
* log likelihood = -132657.31514921974
* Perplexity = 63.10117669422472
* For lambda1 = 0.1 lambda2 = 0.3 lambda3 = 0.6
* M = 22185
* log likelihood = -130330.78017327274
* Perplexity = 58.6770840292505
* For lambda1 = 0.1 lambda2 = 0.4 lambda3 = 0.5
* M = 22185
* log likelihood = -128657.31483541589
* Perplexity = 55.68793856377327
* For lambda1 = 0.1 lambda2 = 0.5 lambda3 = 0.4
* M = 22185
* log likelihood = -127515.49747383496
* Perplexity = 53.736296645807535
* For lambda1 = 0.1 lambda2 = 0.6 lambda3 = 0.30000000000000004
* M = 22185
* log likelihood = -126906.81848355959
* Perplexity = 52.72402048223815
* For lambda1 = 0.1 lambda2 = 0.7 lambda3 = 0.20000000000000007
* M = 22185
* log likelihood = -126987.71051146998
* Perplexity = 52.85744302151083
* For lambda1 = 0.1 lambda2 = 0.8 lambda3 = 0.09999999999999998
* M = 22185
* log likelihood = -128398.80324036039
* Perplexity = 55.23996282386069
* For lambda1 = 0.2 lambda2 = 0.1 lambda3 = 0.7
* M = 22185
* log likelihood = -133317.13283892983
* Perplexity = 64.41552869329257
* For lambda1 = 0.2 lambda2 = 0.2 lambda3 = 0.6
* M = 22185
* log likelihood = -130478.61547459748
* Perplexity = 58.94873801572459
* For lambda1 = 0.2 lambda2 = 0.3 lambda3 = 0.5
* M = 22185
* log likelihood = -128567.42987588377
* Perplexity = 55.53176608209703
* For lambda1 = 0.2 lambda2 = 0.4 lambda3 = 0.3999999999999999
* M = 22185
* log likelihood = -127291.95131662033
* Perplexity = 53.36228494030713
* For lambda1 = 0.2 lambda2 = 0.5 lambda3 = 0.30000000000000004
* M = 22185
* log likelihood = -126601.05113468316
* Perplexity = 52.222726124077596
* For lambda1 = 0.2 lambda2 = 0.6 lambda3 = 0.19999999999999996
* M = 22185
* log likelihood = -126629.43950667081
* Perplexity = 52.26906640590316
* For lambda1 = 0.2 lambda2 = 0.7 lambda3 = 0.10000000000000009
* M = 22185
* log likelihood = -128007.83892079604
* Perplexity = 54.56929616380456
* For lambda1 = 0.3 lambda2 = 0.1 lambda3 = 0.6
* M = 22185
* log likelihood = -131880.16535674088
* Perplexity = 61.58745412241415
* For lambda1 = 0.3 lambda2 = 0.2 lambda3 = 0.5
* M = 22185
* log likelihood = -129355.79656696033
* Perplexity = 56.91659289553919
* For lambda1 = 0.3 lambda2 = 0.3 lambda3 = 0.4
* M = 22185
* log likelihood = -127779.20498517057
* Perplexity = 54.18087326952674
* For lambda1 = 0.3 lambda2 = 0.4 lambda3 = 0.30000000000000004
* M = 22185
* log likelihood = -126910.05018103027
* Perplexity = 52.72934434892808
* For lambda1 = 0.3 lambda2 = 0.5 lambda3 = 0.19999999999999996
* M = 22185
* log likelihood = -126823.73210873362
* Perplexity = 52.587329212487354
* For lambda1 = 0.3 lambda2 = 0.6 lambda3 = 0.10000000000000009
* M = 22185
* log likelihood = -128126.86017051793
* Perplexity = 54.77260049964106
* For lambda1 = 0.4 lambda2 = 0.1 lambda3 = 0.5
* M = 22185
* log likelihood = -131253.83080882087
* Perplexity = 60.39395477497045
* For lambda1 = 0.4 lambda2 = 0.2 lambda3 = 0.3999999999999999
* M = 22185
* log likelihood = -128988.09035120878
* Perplexity = 56.266442897972
* For lambda1 = 0.4 lambda2 = 0.3 lambda3 = 0.30000000000000004
* M = 22185
* log likelihood = -127775.75586434473
* Perplexity = 54.17503482663644
* For lambda1 = 0.4 lambda2 = 0.4 lambda3 = 0.19999999999999996
* M = 22185
* log likelihood = -127484.69821485871
* Perplexity = 53.68461157848009
* For lambda1 = 0.4 lambda2 = 0.5 lambda3 = 0.09999999999999998
* M = 22185
* log likelihood = -128657.67379444125
* Perplexity = 55.688563124227656
* For lambda1 = 0.5 lambda2 = 0.1 lambda3 = 0.4
* M = 22185
* log likelihood = -131319.02014889376
* Perplexity = 60.51708890226749
* For lambda1 = 0.5 lambda2 = 0.2 lambda3 = 0.30000000000000004
* M = 22185
* log likelihood = -129341.31111962876
* Perplexity = 56.89083925908714
* For lambda1 = 0.5 lambda2 = 0.3 lambda3 = 0.19999999999999996
* M = 22185
* log likelihood = -128670.68629462383
* Perplexity = 55.71120858051495
* For lambda1 = 0.5 lambda2 = 0.4 lambda3 = 0.09999999999999998
* M = 22185
* log likelihood = -129621.84645622023
* Perplexity = 57.39168113672517
* For lambda1 = 0.6 lambda2 = 0.1 lambda3 = 0.30000000000000004
* M = 22185
* log likelihood = -132120.90940766802
* Perplexity = 62.05244903882577
* For lambda1 = 0.6 lambda2 = 0.2 lambda3 = 0.19999999999999996
* M = 22185
* log likelihood = -130598.24079590998
* Perplexity = 59.169475207333655
* For lambda1 = 0.6 lambda2 = 0.3 lambda3 = 0.10000000000000009
* M = 22185
* log likelihood = -131137.56676017
* Perplexity = 60.17496876001152
* For lambda1 = 0.7 lambda2 = 0.1 lambda3 = 0.20000000000000007
* M = 22185
* log likelihood = -133916.00523237634
* Perplexity = 65.63216369248737
* For lambda1 = 0.7 lambda2 = 0.2 lambda3 = 0.10000000000000009
* M = 22185
* log likelihood = -133501.7150646392
* Perplexity = 64.7880914078024
* For lambda1 = 0.8 lambda2 = 0.1 lambda3 = 0.09999999999999998
* M = 22185
* log likelihood = -137584.64576309378
* Perplexity = 73.6032161202169
  + - * **5th Set of Training and Development**
* For lambda1 = 0.1 lambda2 = 0.1 lambda3 = 0.8
* M = 21394
* log likelihood = -131647.31120803198
* Perplexity = 71.18339249179311
* For lambda1 = 0.1 lambda2 = 0.2 lambda3 = 0.7
* M = 21394
* log likelihood = -128405.579589897
* Perplexity = 64.08627524600296
* For lambda1 = 0.1 lambda2 = 0.3 lambda3 = 0.6
* M = 21394
* log likelihood = -126160.21685629408
* Perplexity = 59.589680128902145
* For lambda1 = 0.1 lambda2 = 0.4 lambda3 = 0.5
* M = 21394
* log likelihood = -124551.97878991328
* Perplexity = 56.564234046259244
* For lambda1 = 0.1 lambda2 = 0.5 lambda3 = 0.4
* M = 21394
* log likelihood = -123461.53706917327
* Perplexity = 54.60074379294788
* For lambda1 = 0.1 lambda2 = 0.6 lambda3 = 0.30000000000000004
* M = 21394
* log likelihood = -122890.50517416588
* Perplexity = 53.59986575370515
* For lambda1 = 0.1 lambda2 = 0.7 lambda3 = 0.20000000000000007
* M = 21394
* log likelihood = -122992.49002547062
* Perplexity = 53.77726449689657
* For lambda1 = 0.1 lambda2 = 0.8 lambda3 = 0.09999999999999998
* M = 21394
* log likelihood = -124396.3941197725
* Perplexity = 56.27982196162713
* For lambda1 = 0.2 lambda2 = 0.1 lambda3 = 0.7
* M = 21394
* log likelihood = -129146.54542639852
* Perplexity = 65.64338810985693
* For lambda1 = 0.2 lambda2 = 0.2 lambda3 = 0.6
* M = 21394
* log likelihood = -126392.41502241087
* Perplexity = 60.039665024687054
* For lambda1 = 0.2 lambda2 = 0.3 lambda3 = 0.5
* M = 21394
* log likelihood = -124547.68157707773
* Perplexity = 56.55635938262043
* For lambda1 = 0.2 lambda2 = 0.4 lambda3 = 0.3999999999999999
* M = 21394
* log likelihood = -123323.3330945044
* Perplexity = 54.35680514826146
* For lambda1 = 0.2 lambda2 = 0.5 lambda3 = 0.30000000000000004
* M = 21394
* log likelihood = -122668.92640266746
* Perplexity = 53.21645228697299
* For lambda1 = 0.2 lambda2 = 0.6 lambda3 = 0.19999999999999996
* M = 21394
* log likelihood = -122716.61307343656
* Perplexity = 53.29873563014661
* For lambda1 = 0.2 lambda2 = 0.7 lambda3 = 0.10000000000000009
* M = 21394
* log likelihood = -124085.42130468969
* Perplexity = 55.715635535162406
* For lambda1 = 0.3 lambda2 = 0.1 lambda3 = 0.6
* M = 21394
* log likelihood = -127822.69394458446
* Perplexity = 62.88736274133869
* For lambda1 = 0.3 lambda2 = 0.2 lambda3 = 0.5
* M = 21394
* log likelihood = -125373.99630028704
* Perplexity = 58.09092978243172
* For lambda1 = 0.3 lambda2 = 0.3 lambda3 = 0.4
* M = 21394
* log likelihood = -123853.66085629824
* Perplexity = 55.29884257260932
* For lambda1 = 0.3 lambda2 = 0.4 lambda3 = 0.30000000000000004
* M = 21394
* log likelihood = -123023.58241550103
* Perplexity = 53.83146521604039
* For lambda1 = 0.3 lambda2 = 0.5 lambda3 = 0.19999999999999996
* M = 21394
* log likelihood = -122957.51751023854
* Perplexity = 53.71636509563869
* For lambda1 = 0.3 lambda2 = 0.6 lambda3 = 0.10000000000000009
* M = 21394
* log likelihood = -124250.71577722425
* Perplexity = 56.014815370165756
* For lambda1 = 0.4 lambda2 = 0.1 lambda3 = 0.5
* M = 21394
* log likelihood = -127269.14789640225
* Perplexity = 61.76956811322603
* For lambda1 = 0.4 lambda2 = 0.2 lambda3 = 0.3999999999999999
* M = 21394
* log likelihood = -125073.11136625406
* Perplexity = 57.52738670641366
* For lambda1 = 0.4 lambda2 = 0.3 lambda3 = 0.30000000000000004
* M = 21394
* log likelihood = -123907.29761297471
* Perplexity = 55.39502360396854
* For lambda1 = 0.4 lambda2 = 0.4 lambda3 = 0.19999999999999996
* M = 21394
* log likelihood = -123640.47383243377
* Perplexity = 54.91820532124313
* For lambda1 = 0.4 lambda2 = 0.5 lambda3 = 0.09999999999999998
* M = 21394
* log likelihood = -124805.11764653273
* Perplexity = 57.03005212570151
* For lambda1 = 0.5 lambda2 = 0.1 lambda3 = 0.4
* M = 21394
* log likelihood = -127376.5669508136
* Perplexity = 61.9849183298038
* For lambda1 = 0.5 lambda2 = 0.2 lambda3 = 0.30000000000000004
* M = 21394
* log likelihood = -125462.78383849803
* Perplexity = 58.25827704628633
* For lambda1 = 0.5 lambda2 = 0.3 lambda3 = 0.19999999999999996
* M = 21394
* log likelihood = -124825.54328526315
* Perplexity = 57.06780557134139
* For lambda1 = 0.5 lambda2 = 0.4 lambda3 = 0.09999999999999998
* M = 21394
* log likelihood = -125772.81135533538
* Perplexity = 58.846408404856206
* For lambda1 = 0.6 lambda2 = 0.1 lambda3 = 0.30000000000000004
* M = 21394
* log likelihood = -128192.14597305274
* Perplexity = 63.64464317684304
* For lambda1 = 0.6 lambda2 = 0.2 lambda3 = 0.19999999999999996
* M = 21394
* log likelihood = -126724.38398954463
* Perplexity = 60.6889073726681
* For lambda1 = 0.6 lambda2 = 0.3 lambda3 = 0.10000000000000009
* M = 21394
* log likelihood = -127269.89511961502
* Perplexity = 61.77106353481571
* For lambda1 = 0.7 lambda2 = 0.1 lambda3 = 0.20000000000000007
* M = 21394
* log likelihood = -129967.2726821566
* Perplexity = 67.41231645218595
* For lambda1 = 0.7 lambda2 = 0.2 lambda3 = 0.10000000000000009
* M = 21394
* log likelihood = -129584.6591093174
* Perplexity = 66.58180803191198
* For lambda1 = 0.8 lambda2 = 0.1 lambda3 = 0.09999999999999998
* M = 21394
* log likelihood = -133560.19087875812
* Perplexity = 75.73460535540505
  + - * **For all the 5 set I found minimum likelihood at lambda1 = 0.2, lambda2 = 0.5 and lambda3 = 0.3.**
      * **Test Set Analysis using Linear Interpolation**
        + For lambda1 = 0.2 lambda2 = 0.5 lambda3 = 0.3
        + M = 23820
        + log likelihood = -137150.5706
        + Perplexity = 54.10976
      * **Same Test using Laplace Smoothing**
        + M = 23820
        + log likelihood = -205915.6957
        + Perplexity = 400.22169
  + **Result: Both Linear Interpolation and Discounting Smoothing performed approximately equal in my case. Laplace Smoothing gave poor results as compared to both Linear Interpolation and Discounting Smoothing Techniques.**

**2. Vector Semantics: GloVE implementation**

**Colab Link:** <https://colab.research.google.com/drive/13wiKCr-hc2iwmK6qcu7zYfk-V_tLrZEU?usp=sharing>

**Pre-Trained Vector Embedding Used For Comparision:** <https://www.kaggle.com/fnugget/glove-wikipedia-2014-gigaword-5>

* **Comparision: My vector embedding had given mostly random similar words.**
  + print(mostSimilar(['two'], ['one']))

Similar Words:-

('however', 0.6165489530455645), ('including', 0.6139982091478733), ('would', 0.588528429289489), ('cord', 0.5857220645489676), ('who', 0.5712874340646518), ('Arabic', 0.5697040374630845), ('A', 0.5667756873945056), ('about', 0.5635937513598844), ('Wisconsin', 0.5621473727901255), ('if', 0.5553539994099479)

* + print(mostSimilar(['was', 'has'], ['is']))

Similar Words: -

('have', 0.9828102417065027), ('used', 0.9735759264042814), ("'s", 0.9723950099335225), ('that', 0.9709112792483112), ('Derleth', 0.970736058307157), ('an', 0.970728567356585), ('his', 0.9704711833085259), ('are', 0.9696385243192851), ('Lovelace', 0.9695295696677386), ('had', 0.9684613148424044)

* + print(mostSimilar(['by', 'the'], ['to']))

Most Similar:-

('an', 0.9925601985939683), ('ataxia', 0.991052540464115), ('his', 0.9908705132546708), ('her', 0.9905561385538236), ('are', 0.9902644731804793), ('that', 0.9902021010258358), ('and', 0.9897249363663992), ('in', 0.9887695283377845), ('was', 0.9886601018051455), ("'s", 0.9875598838893258)

* + print(mostSimilar(['husband', 'father'], ['wife']))

Most Similar:-

('important', 0.4454047487508519), ('could', 0.43799456374126017), ('common', 0.4364977003915271), ('potentials', 0.42733658286178866), ('MS', 0.42609557980060203), ('treatment', 0.4256605814279289), ('did', 0.4239630878646214), ('associated', 0.42305440550270407), ('seen', 0.42259954869152294), ('Himmler', 0.42223604473621146)

* **Pre-Trained Embedding had given better results than mine.**
* **Expression of derivatives required in optimization algorithm:**

