A)

Token: MILLION Uni-Prob: 0.002073

Token: MORE Uni-Prob: 0.001709

Token: MR. Uni-Prob: 0.001442

Token: MOST Uni-Prob: 0.000788

Token: MARKET Uni-Prob: 0.00078

Token: MAY Uni-Prob: 0.00073

Token: M. Uni-Prob: 0.000703

Token: MANY Uni-Prob: 0.000697

Token: MADE Uni-Prob: 0.00056

Token: MUCH Uni-Prob: 0.000515

Token: MAKE Uni-Prob: 0.000514

Token: MONTH Uni-Prob: 0.000445

Token: MONEY Uni-Prob: 0.000437

Token: MONTHS Uni-Prob: 0.000406

Token: MY Uni-Prob: 0.0004

Token: MONDAY Uni-Prob: 0.000382

Token: MAJOR Uni-Prob: 0.000371

Token: MILITARY Uni-Prob: 0.000352

Token: MEMBERS Uni-Prob: 0.000336

Token: MIGHT Uni-Prob: 0.000274

Token: MEETING Uni-Prob: 0.000266

Token: MUST Uni-Prob: 0.000267

Token: ME Uni-Prob: 0.000264

Token: MARCH Uni-Prob: 0.00026

Token: MAN Uni-Prob: 0.000253

Token: MS. Uni-Prob: 0.000239

Token: MINISTER Uni-Prob: 0.00024

Token: MAKING Uni-Prob: 0.000212

Token: MOVE Uni-Prob: 0.00021

Token: MILES Uni-Prob: 0.000206

B)

| **probability** | **words** |
| --- | --- |
| **0** | 0.615020 | <UNK> |
| **1** | 0.013372 | U. |
| **2** | 0.011720 | FIRST |
| **3** | 0.011659 | COMPANY |
| **4** | 0.009451 | NEW |
| **5** | 0.008672 | UNITED |
| **6** | 0.006803 | GOVERNMENT |
| **7** | 0.006651 | NINETEEN |
| **8** | 0.006287 | SAME |
| **9** | 0.006161 | TWO |

C)

Unigram log probability: - 64.50944034364878

Bigram log probability: - 40.91813213

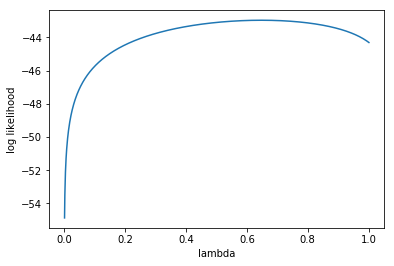
D)

Unigram log probability: -44.291934473132606

Bigram log probability: -inf

Pairs not seen: 'SIXTEEN OFFICIALS' 'SOLD FIRE' – when pairs are not seen it send the log probability to negative infinity

e)



Optimal Lambda: 0.648 -> log prob -42.96

4.4

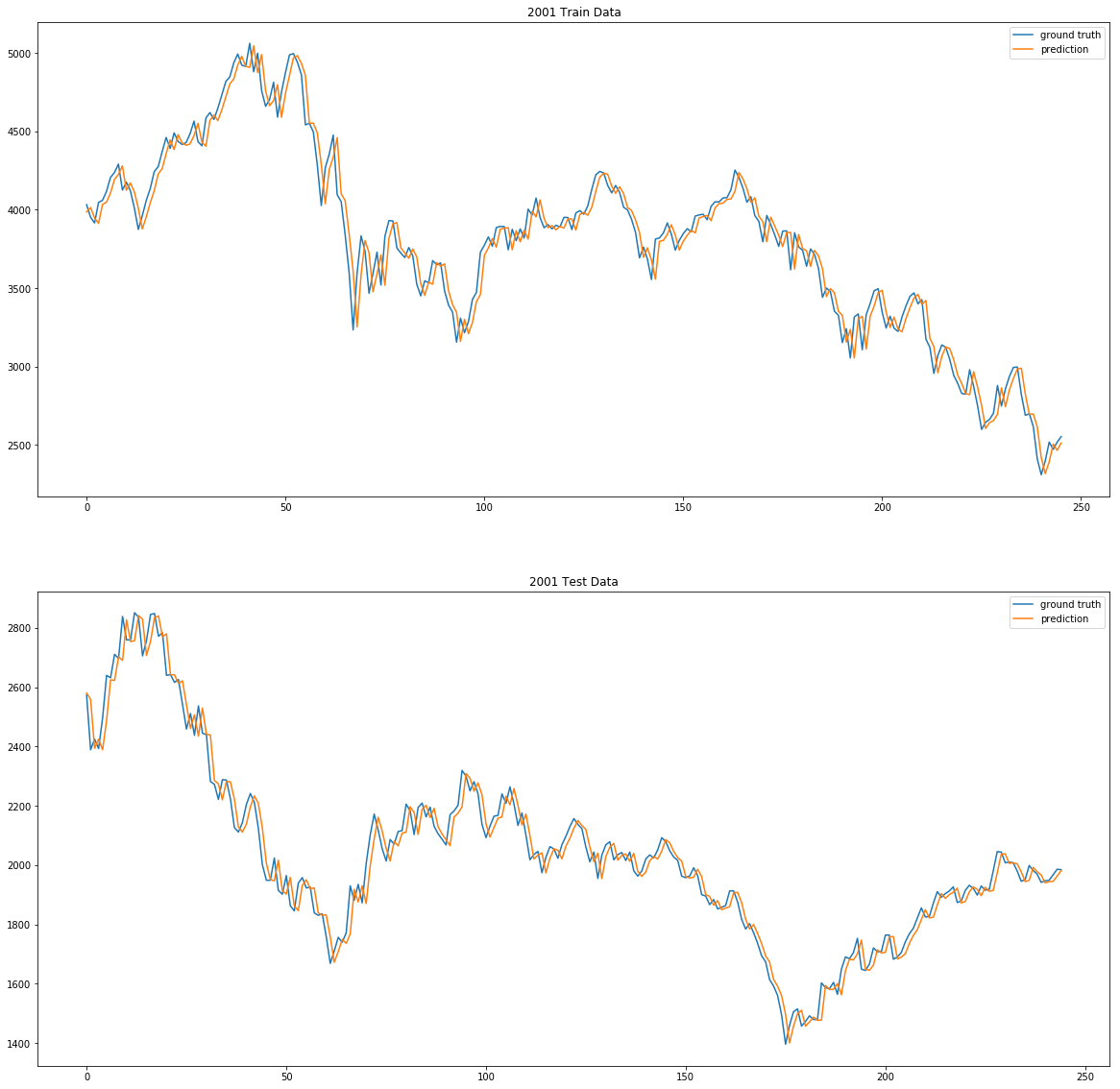


a1 : 0.031896

a2 : 0.015601

a3 : 0.950673

b) MSE Test: 2985.1 MSE Train: 13895.86



The model is essentially a time lag model that predicts the value at the next time step to be equivalent to the current time step, as evidenced by the heavy weight on a3. This is a poor model because it essentially admits that it does not know if the price will go up or down, so the safest bet is to predict the same as the previous step in order to minimize the prediction loss.