Sarp Misoglu Victor de Fontnouvelle

1.

Here are the outputs on the two sentences:

```
positive -2.2626488108728853       negative -2.739770065592548
positive -4.701388880801452        negative -4.701388880801452
```

They were both classified as positive. The first sentence was assigned a probability that was clearly larger than that of the negative sentence. However, the probabilities for the two labels were equal for the second sentence, so the program broke the tie by choosing the positive label. It makes sense that both labels would have the same probability, as the word “hated” and “loved” were both included once.

2.

```
Pos label prob: 0.5
Neg label prob: 0.5

Pos word probs:
that -- 0.15384615384615385
movie -- 0.07692307692307693
loved -- 0.23076923076923078
i -- 0.3076923076923077
it -- 0.15384615384615385
hated -- 0.07692307692307693

Neg word probs:
that -- 0.15384615384615385
movie -- 0.07692307692307693
loved -- 0.07692307692307693
i -- 0.3076923076923077
it -- 0.15384615384615385
hated -- 0.23076923076923078
```

3.

```
Top positive predictors:
superb -- 46.062439785890874
outstanding -- 30.63786132547788
touching -- 24.510289060382302
all-time -- 23.031219892945437
incredible -- 21.869094118530764
intense -- 15.847169651109246
genius -- 15.847169651109246
brilliantly -- 15.001987269716754
wonderfully -- 13.734213697628014
fabulous -- 13.100326911583645
```

```
Top negative predictors:
redeeming -- 80.4560074808535
unfunny -- 73.35694799724878
dreadful -- 66.25788851364405
waste -- 63.497143158908884
worst -- 61.525182191240916
dreck -- 56.79247586883776
horrid -- 54.42612270763619
wasted -- 48.510239804632256
horrendous -- 47.32706322403147
atrocious -- 47.32706322403147
```

These words make sense. The positive predictors are all words that are associated with praise, while the negative predictors are all associated with disgust.

4.

Negative - correctly classified:
worst movie ever . are you kidding me ?
Negative - incorrectly classified:
i could not wait to exit the theater ! ! ! it is incredible that they made a movie like this

Positive - correctly classified:
classic ! way better thanthe others ! i love what they do !
Positive - incorrectly classified:
my friend said this movie was unfunny , a dreadful waste of time , but i disagree ! i think it should be in the imdb top 250 . what a horrible friend

The incorrectly classified positive review uses many negative words, but the context of the sentence negates them. A similar phenomenon occurs for the negative sentence. This explains the misclassification.

5.

Lamda	Accuracy Dev	Accuracy Test
0.001	0.8997837058399423	0.8738284066330209
0.01	0.9098774333093006	0.8882480173035328
0.1	0.9206921413121846	0.9098774333093006
0.5	0.9271809661139149	0.9142033165104542
0.8	0.916366258111031	0.8940158615717375

For both the dev and test set, a lambda value of 0.5 works best. It is surprising that there is a significant difference in performance across the dev and test set. The accuracy for the test set is consistently worse.

It is also worth noting that the fraction of positive sentences in the dev set is 0.789, so if the classifier predicted positive every time, it would have an accuracy of 0.789. The classifier we made has an accuracy of 0.927 at best. This is definitely an improvement, but put into perspective, it is not as impressive as it might initially seem.