Problem 1: M1 Word absence & presence
Accuracy is 0.86%
1 -5.956848974266556 the
1 -5.956848974266556 of
1 -5.957959468550583 to
1 -5.957959468550583 is
1 -5.957959468550583 and
0 -5.889864371418765 the
0 -5.889864371418765 of
0 -5.889864371418765 and
0 -5.890974865702792 to
0 -5.890974865702792 is
The accuracy is quite high. However, the most informative words are almost exactly the same for both classes. These words seems to be stop words, that are common is any documents.

Problem 1: M2 Term frequency
Accuracy is 0.835%
1 -2.840738587277439 the
1 -3.5756175119023332 and
1 -3.6397703179736975 of
1 -3.759997481071675 to
1 -3.898152823893758 is
0 -2.899320163163642 the
0 -3.7056424046697813 and
0 -3.719386357592498 of
0 -3.7196026127744553 to
0 -4.0058678632326075 is
The accuracy is also quite high. However, once again, the most informative words are almost exactly the same for both classes. These words seems to be stop words, that are common is any documents.

Problem 1: M3 Adj/Adv POS tagging
Accuracy is 0.87%
1 -4.839375545880551 not
1 -4.978294782472198 more
1 -5.005486790167449 when
1 -5.065441651637363 so
1 -5.122313162591969 most
0 -4.784832166746287 not
0 -4.916657559525709 when
0 -4.930387752337611 so
0 -4.930387752337611 more
0 -4.979177916507043 only
The accuracy is the highest so far. However, once again, the most informative words are almost exactly the same for both classes. These words seems to be stop words, that are common is any documents.

Problem 1: M4 Sublinear TF-IDF

Accuracy is 0.855%
1 -6.950690881656748 movie
1 -7.117846360899243 like
1 -7.217263992577482 story
1 -7.219986250430932 life
1 -7.229668354616171 good
0 -6.738589295181579 movie
0 -6.97872693412327 like
0 -7.098268371686496 bad
0 -7.183224842908503 good
0 -7.211225327985219 plot
The accuracy is also quite high. However, once again, the most informative words are almost exactly the same for both classes, even when the stop words are removed. The only noticeable difference is that the negative word 'bad' is only most informative for the negative class.

Problem 1: M5 Bigrams
Accuracy is 0.835%
Bigram accuracy is the lowest among all. Seems like neigbouring words do not help the classifier to identify classes.

Problem 2: M1
Accuracy is 0.865%
1 -5.874126057465444 the
1 -5.874126057465444 of
1 -5.875236551749471 to
1 -5.875236551749471 is
1 -5.875236551749471 and
0 -5.811011388055521 the
0 -5.811011388055521 of
0 -5.811011388055521 and
0 -5.8121218823395475 to
0 -5.8121218823395475 is

Problem 2: M2
Accuracy is 0.835%
1 -2.8218423300926645 the
1 -3.5565865292205547 and
1 -3.6204997425983514 of
1 -3.7408313906412616 to
1 -3.879199411704203 is
0 -2.878436116876138 the
0 -3.6844726304885285 and
0 -3.6981445087418567 of
0 -3.6984328385932024 to
0 -3.984986093661419 is

Problem 2: M3
Accuracy is 0.865%

1 -4.816078359324657 not
1 -4.954997595916304 more
1 -4.982189603611555 when
1 -5.042144465081469 so
1 -5.099015976036076 most
0 -4.7584993363994705 not
0 -4.892977250628024 when
0 -4.9067074434399265 so
0 -4.9067074434399265 more
0 -4.955497607609358 onli

Problem 2: M4
Accuracy is 0.85%
1 -6.851678248100096 charact
1 -6.873881312329068 like
1 -6.913031948970374 make
1 -6.93374621915882 time
1 -6.975088959919418 stori
0 -6.725568462109111 like
0 -6.811139260741474 charact
0 -6.881642269038005 bad
0 -6.890574116575978 make
0 -6.939384912776044 time

Problem 2: M5
Accuracy is 0.835%

The accuracy did not significantly increase and sometimes decreased so it is not worth stemming.

Problem 3: Pos/Neg Ratio
Accuracy is 0.5%

Problem 4: Improvements
1. Numbers can be converted to just placeholders. There are many numbers that became parts of the dictionary. However, the different numbers do not have different sentiments attached unless they are used as ways to rate things.
2. Removing stop words and words that are generic for the subject (e.g. movie, story, plot). As is seen previously, removing stop words helped with model performance. We can go one step further to remove generic words that refers to the subject in discussion. Previously, the informative words were things like movie, story, and plot. However, they are just words that are synonyms to the subject 'movie'. Thus getting rid of them might make the more sentiment-telling words surface more easily.
3. Only keeping words that are not neutral (i.e. sentiment-telling words). Even though the pos/neg ratio did not work too well with predicting sentiment, one thing that it can be useful is as a filter. After only keeping the non-neutral words, if will be more clear what

the sentiment of the text is.