Shen Gao   *A20377269*

CS585 Assignment 1
metrics and evaluation

Classification and Evaluation:
Evaluation results:

ALPHA	0.1	0.5	1.0	5.0	10.0
Accuracy	81.920%	82.896%	83.216%	83.196%	83.032%

Probability Prediction:
Probability estimated for the first 10 reviews in the test data:

Real label	Predicted label	Positive Probability	Negative Probability
-1.0	-1.0	0.00368	0.99632
-1.0	-1.0	3.978E-08	0.99999
-1.0	1.0	0.97916	0.02084
-1.0	-1.0	0.03322	0.96678
-1.0	-1.0	0.00245	0.99755
-1.0	-1.0	6.107E-19	1.0
-1.0	-1.0	1.426E-08	0.99999
1.0	1.0	1.0	3.021E-16
-1.0	-1.0	5.435E-08	0.99999
-1.0	-1.0	0.38128	0.61872

Precision and Recall:
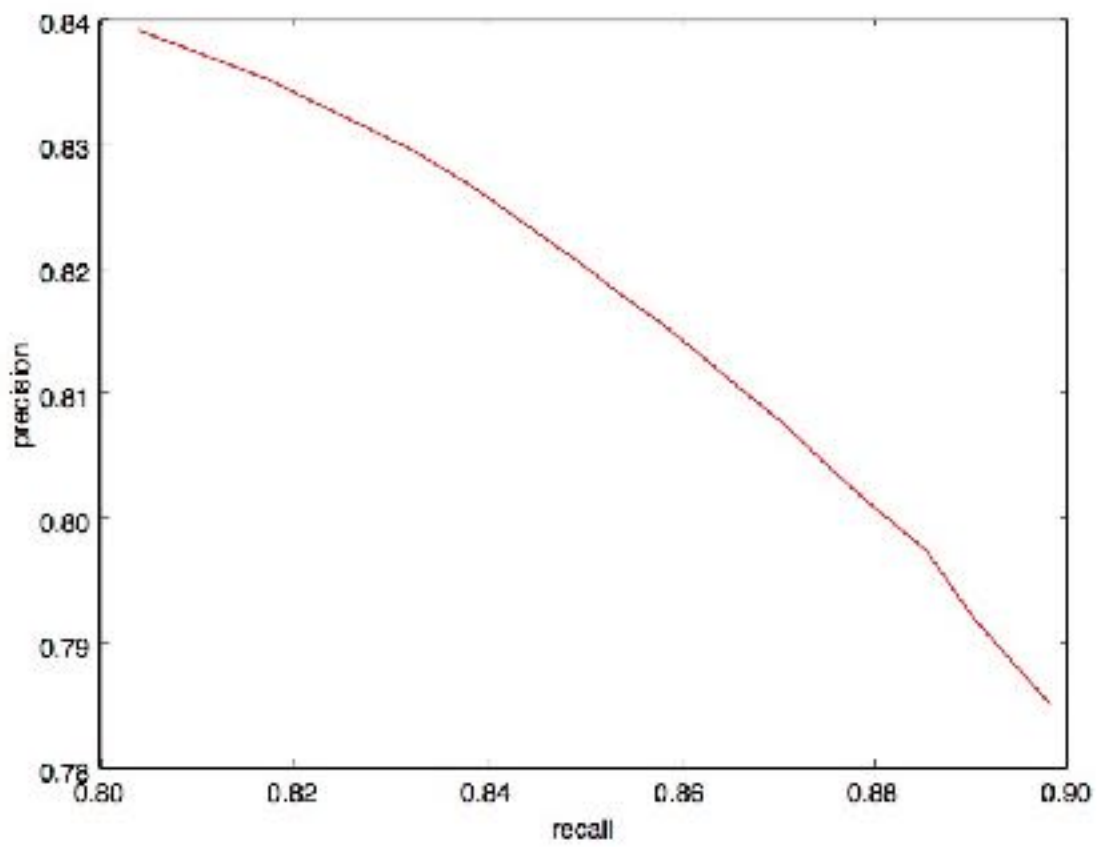Precision of positive class = 0.81560
Recall of positive class = 0.8584
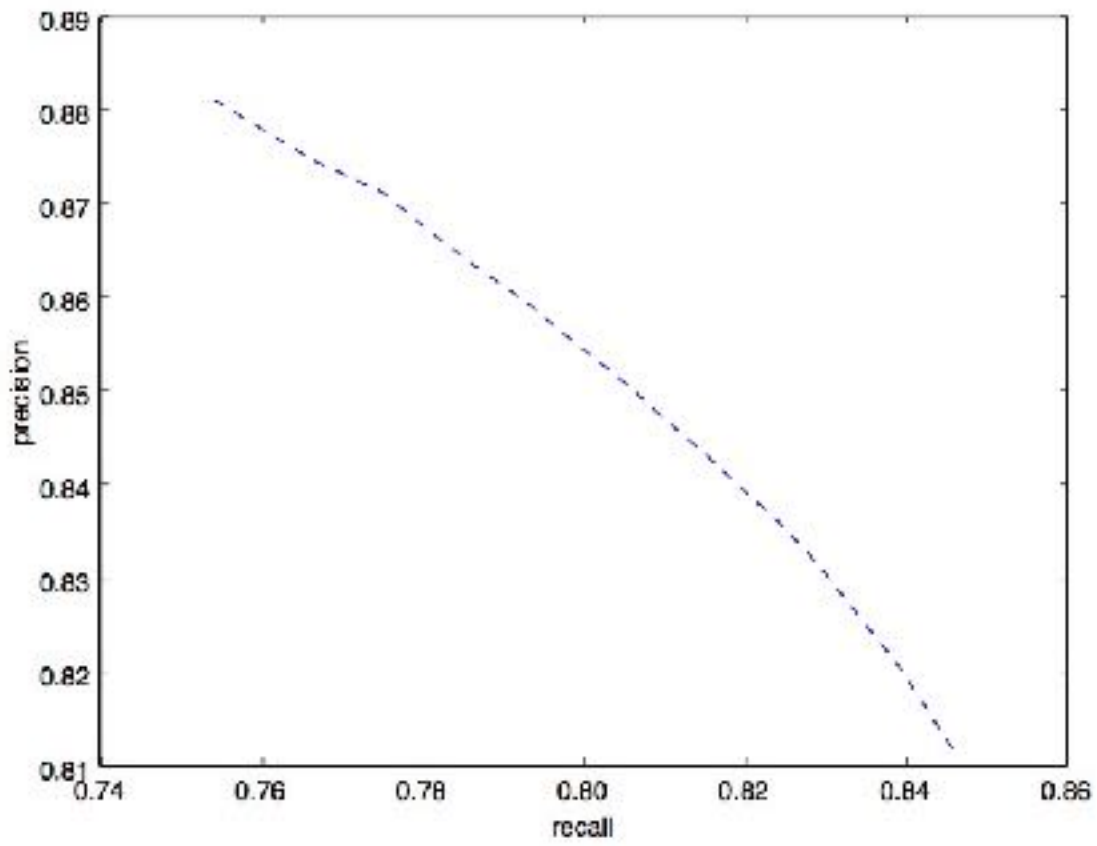Precision of negative class = 0.85056
Recall of negative class = 0.80592

Precision and Recall Curves(thresholds are 0.10,0.15,0.20,0.25,0.30,0.35,0.40,0.45,0.50,0.55,0.60,0.65,0.70,0.75,0.80,0.85,0.90)

Positive Class

Negative Class

Relationship: When recall increases, precision decreases.

Features:
20 most positive words:
(word, weight): (the,-0.468),(and,-1.112),(a,-1.171),(of,-1.246),(to,-1.391),(is,-1.557),
(in,-1.699),(i,-2.078),(it,-2.088),(that,-2.091),(this,-2.100),(as,-2.346),(/><br,-2.375),
(with,-2.460),(for,-2.501),(was,-2.514),(but,-2.611),(his,-2.743),(on,-2.816),(film,-2.895)

20 most negative words:
(word, weight): (the,-0.526),(a,-1.223),(and,-1.308),(of,-1.355),(to,-1.358),(is,-1.699),
(in,-1.838),(i,-1.959),(this,-1.960),(that,-2.064),(it,-2.091),(/><br,-2.308),(was,-2.336),
(for,-2.527),(but,-2.567),(with,-2.578),(as,-2.595),(movie,-2.724),(on,-2.820),(not,-2.859)