=== Run information ===

Scheme: weka.classifiers.functions.LibSVM -S 0 -K 2 -D 3 -G 0.0 -R 0.0 -N 0.5 -M 40.0 -C 1.0 -E 0.001 -P 0.1 -model "C:\\Program Files\\Weka-3-7" -seed 1

Relation: yelp\_task2\_trainingset\_binaryclassification-weka.filters.unsupervised.attribute.NominalToString-Cfirst-weka.filters.unsupervised.attribute.StringToWordVector-R1-W2000-prune-rate-1.0-C-T-I-N1-L-S-stemmerweka.core.stemmers.NullStemmer-M1-tokenizerweka.core.tokenizers.NGramTokenizer -max 2 -min 2 -delimiters " \r\n\t.,;:\'\"()?!"-weka.filters.unsupervised.attribute.Reorder-R2-last,1-weka.filters.unsupervised.attribute.Reorder-R2-last,1-weka.filters.unsupervised.attribute.Reorder-R2-last,1-weka.filters.unsupervised.attribute.Reorder-R2-last,1-weka.filters.supervised.attribute.AttributeSelection-Eweka.attributeSelection.InfoGainAttributeEval-Sweka.attributeSelection.Ranker -T 0.0 -N -1

Instances: 27844

Attributes: 43

in vegas

the bruschetta

of wine

the wine

the airport

the best

customer service

wine and

bottle of

is great

oil change

las vegas

tried to

to enjoy

rude and

money and

i love

do not

this office

in henderson

the casino

the wings

called to

the food

the office

i called

the other

the manager

me to

hot dog

a bottle

chicken and

the guy

lack of

delicious i

place and

the perfect

told them

the wrong

but nothing

votes.cool

glass of

review.rating

Test mode: evaluate on training data

=== Classifier model (full training set) ===

LibSVM wrapper, original code by Yasser EL-Manzalawy (= WLSVM)

Time taken to build model: 182.83 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 91.01 seconds

=== Summary ===

Correctly Classified Instances 17672 63.4679 %

Incorrectly Classified Instances 10172 36.5321 %

Kappa statistic 0.04

Mean absolute error 0.3653

Root mean squared error 0.6044

Relative absolute error 78.0778 %

Root relative squared error 124.9626 %

Coverage of cases (0.95 level) 63.4679 %

Mean rel. region size (0.95 level) 50 %

Total Number of Instances 27844

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

0.983 0.951 0.635 0.983 0.771 0.094 0.516 0.634 High Rating

0.049 0.017 0.639 0.049 0.091 0.094 0.516 0.386 Low Rating

Weighted Avg. 0.635 0.602 0.636 0.635 0.517 0.094 0.516 0.542

=== Confusion Matrix ===

a b <-- classified as

17162 288 | a = High Rating

9884 510 | b = Low Rating