=== 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 3 -min 3 -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: 54

spicy korean chicken

the wine list

should have been

in las vegas

at the front

bottle of wine

to try the

was really good

glass of wine

wife had the

the royal palms

a phone call

have a lot

feel like i

taken care of

we would have

i took it

next to the

i look forward

an oil change

customer service i

cant go wrong

the place and

what the hell

for an oil

i told them

of wine and

a bottle of

la grande orange

of the day

votes.cool

look forward to

at least the

this is not

on the same

hung up on

food is delicious

i called to

great food and

for a place

dont have to

chicken noodle soup

right next to

to have it

as if they

for my taste

at the time

a few more

never come back

i asked her

last time we

have gone to

a gift card

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: 94.45 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 64.7 seconds

=== Summary ===

Correctly Classified Instances 17651 63.3925 %

Incorrectly Classified Instances 10193 36.6075 %

Kappa statistic 0.0361

Mean absolute error 0.3661

Root mean squared error 0.605

Relative absolute error 78.239 %

Root relative squared error 125.0915 %

Coverage of cases (0.95 level) 63.3925 %

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

0.985 0.956 0.634 0.985 0.771 0.090 0.515 0.634 High Rating

0.044 0.015 0.641 0.044 0.082 0.090 0.515 0.385 Low Rating

Weighted Avg. 0.634 0.605 0.636 0.634 0.514 0.090 0.515 0.541

=== Confusion Matrix ===

a b <-- classified as

17194 256 | a = High Rating

9937 457 | b = Low Rating