

Oct 29, 13 13:01

573:proj2:Naveen Kumar Lekkalapudi

Page 1/1

```

+Title: Proj 2:CS573
+=====
+I would like to calculate the performance of an k-naive bayes learner.
+Where k are the nearest neighbours on which the naive bayes classifier is appli
ed.
5 +
+Prologue
+-----
+It would seem a bit ridiculous to combine a supervised and unsupervised learnin
g methods.
+But I suppose it would clear the disadvantages of each learner and provide us w
ith an efficient learner with all the advantages.
10 +
+The disadvantage of the naive bayes classifier is that it looses its accuracy a
round the boundaries of clusters.
+We can improve the accuracy of naive bayes by determining the boundaries of eac
h cluster.
+
+We can use k-nn method to our advantage and consider only the values closest to
the given test case to apply naive bayes.
15 +This helps us to zoom into the boundaries of clusters to clear the misprediction
s and improve accuracy.
+
+Approximate Steps
+-----
+1a. For a given test record, Calculate the distance using knn to get the attrib
ute "distance" for each record in training data.
20 +1b. While calculating the distance ignore the goal column.
+2. By using the "distance"s calculated obtain a cluster that is closest to the
given test case.
+3. This cluster consists of the nearest records of given test record.
+4. To get the closest cluster we need to perform cuts on the entire training da
ta.
+5. The tests given in scott-knott procedure i.e. Cohens, Hedges, A12 can be use
d to determine the best cut on "distance".
25 +6. The cluster obtained from such cut is closest to the given test case.
+7. We could perform naive bayes on the values of that cluster (by including the
goal column now) to predict the goal.
+8. As it involves only the values closer to the test case, I assume we are zoom
ing into the part where clusters of goal meet.
+9. With more clarity, I assume the accuracy to be improving.
+
30 +Conclusion
+-----
+I assume that by considering only the values closer to given case, we could imp
rove the accuracy of naive bayes.
+Which I would like to test in the project.

```