CHAPTER 6

ENSEMBLE METHOD

6.1 INTRODUCTION

An ensemble are a set of classifiers whose individual predictions were combined in some way to classify new examples. The classifiers that make up the ensemble is called base models and the learning systems that produced these models the base learners. An accurate classifier are one that predicts more than 50% of the new examples correctly. Two classifiers were diverse if they make independent errors on new data.

6.2 LITERATURE SURVEY

Bagging predictors are a method for generating multiple versions of a predictor and using these to get an aggregated predictor [114]. The aggregation averages over the versions when predicting a numerical outcome and does a plurality vote when predicting a class. The multiple versions were formed by making bootstrap replicates of the learning set and using these as new learning sets. Tests on real and simulated data sets using classification and regression trees and subset selection in linear regression show that bagging can give substantial gains in accuracy. The vital element are the instability of the prediction method. If perturbing the learning set can cause
significant changes in the predictor constructed, then bagging can improve accuracy.

The technology for building knowledge-based systems by inductive inference from examples have been demonstrated successfully in several practical applications. In [115] J.R Quinlan summarizes an approach to synthesizing decision trees that have been used in a variety of systems, and it describes one such system, ID3, in detail. Results from recent studies show ways in which the methodology can be modified to deal with information that is noisy and/or incomplete. A reported shortcoming of the basic algorithm are discussed and two means of overcoming it are compared. The work concludes with illustrations of current research directions.

Statistics are a subject of many uses and surprisingly few effective practitioners. The traditional road to statistical knowledge was blocked, for most, by a formidable wall of mathematics [117]. The approach in an Introduction to the Bootstrap avoids that wall. It arms scientists and engineers, as well as statisticians, with the computational techniques they need to analyze and understand complicated data sets.

Bagging and boosting are the methods that generate a diverse ensemble of classifiers by manipulating the training data given to a “base” learning algorithm. In [118] Breiman have pointed out that they
rely for their effectiveness on the instability of the base learning algorithm. An alternative approach to generating an ensemble are to randomize the internal decisions made by the base algorithm. This paper compares the effectiveness of randomization, bagging, and boosting for improving the performance of the decision tree algorithm C4.5. The experiments show that in situations with little or no classification noise, randomization are competitive with (and perhaps slightly superior to) bagging but not as accurate as boosting. In situations with substantial classification noise, bagging is much better than boosting, and sometimes better than randomization.

Methods for voting classification algorithms, such as Bagging and AdaBoost, have been shown to be very successful in improving the accuracy of certain classifiers for artificial and real-world datasets. We review these algorithms and describe a large empirical study comparing several variants in conjunction with a decision tree inducer (three variants) and a Naive-Bayes inducer [119]. The purpose of the study are to improve our understanding of why and when these algorithms, which use perturbation, reweighting, and combination techniques, affect classification error. We provide a bias and variance decomposition of the error to show how different methods and variants influence these two terms. This allowed us to determine that Bagging reduced variance of unstable methods, while boosting methods reduced both the bias and variance of unstable methods but increased
the variance for Naive-Bayes, which is very stable. We observed that Arc-x4 behaves differently than AdaBoost if reweighting are used instead of resampling, indicating a fundamental difference. Voting variants, some of which is introduced in this paper, include: pruning versus no pruning, use of probabilistic estimates, weight perturbations (Wagging), and backfitting of data. We found that Bagging improves when probabilistic estimates in conjunction with no-pruning is used, as well as when the data was backfit. We measure tree sizes and show an interesting positive correlation between the increase in the average tree size in AdaBoost trials and its success in reducing the error. We compare the mean-squared error of voting methods to non-voting methods and show that the voting methods lead to large and significant reductions in the mean-squared errors. Practical problems that arise in implementing boosting algorithms were explored, including numerical instabilities and underflows. We use scatterplots that graphically show how AdaBoost reweights instances, emphasizing not only “hard” areas but also outliers and noise.

6.3 ENSEMBLE METHOD

The J48 (C4.8) are the powerful decision tree method that performs on the breast cancer dataset. This work investigate whether we can improve upon the result of the J48 algorithm using ensemble method. We are going to try three popular ensemble methods: Boosting, Bagging and Blending.
BOOSTING

Boosting are the ensemble method that starts out with a base classifier that are prepared on the training data. A second classifier are then created behind it to focus on the instances in the training data that the first classifier got wrong. The process continues to add classifiers until a limit are reached in the number of models or accuracy.

ADABOOST M1 are class for boosting a nominal class classifier using the ADABOOST M1[114] method. Only nominal class problems can be tackled. Often dramatically improves performance, but sometimes overfits. ADABOOST M1 are adaptive in the sense that subsequent weak learners are tweaked in favor of those instances misclassified by previous classifiers.

BAGGING

Bagging (Bootstrap Aggregating) are the ensemble method that creates separate samples of the training dataset and creates a classifier for each sample. The results of these multiple classifiers were then combined (such as averaged or majority voting). The tricks are that each sample of the training dataset are different, giving each classifier that is trained, a subtly different focus and perspective on the problem.
Bagging [115] are the class for bagging a classifier to reduce variance. It can do classification and regression depending on the base learner. We can choose here the bag size that is saying a bag size of 100%, which is going to sample the training set to get another set the same size, but it's going to sample "with replacement". That means we're going to get different sets of the same size every time we sample, but each set might contain repeats of the original training.

**BLENDING**

Blending are the ensemble method where multiple different algorithm is prepared on the training data and a meta classifier are prepared that learns how to take the predictions of each classifier and make accurate predictions on unseen data.

Stacking [116] combines several classifiers using the stacking method. It can do classification or regression. You can choose different meta-classifiers here, and the number of stacking folds. We can choose different classifiers; different level-0 classifiers, and a different meta-classifier. In order to create multiple level-0 models, we can specify a meta-classifier as the level-0 model.

**6.4 EXPERIMENTAL RESULTS**

The dataset is loaded in the WEKA experimenter so as to classify it into class. The main goal of this experiment is to increase the efficiency using ensemble methods. The ADABOOST M1 is used as a
boosting ensemble, BAGGING is used as a bagging ensemble. This work will add Stacking with two classifiers (J48 and IBk) and use Logistic Regression as the meta classifier. The J48 and IBk were very different algorithms. Logistic Regression are a good reliable and simple method to learn how to combine the predictions from these two methods and are well suited to this binary classification problem as it produces binary outputs itself.

6.4.1 Performance Measures

In order to evaluate the performance of the detector, and to allow comparisons, several ratios have been taken into account.

**True Positive (TP):** The detector found normal data when normal data are present.

**True Negative (TN):** The detector found abnormal data when abnormal data are present.

**False Positive (FP):** The detector found abnormal data when a normal sample is present.

**False Negative (FN):** The detector found normal data when abnormal data are present.

Sensitivity is the percentage of abnormal data classified as abnormal by the procedure. Specificity are the percentage of normal
data classified as normal by the procedure. The higher the sensitivity and specificity values, the better the procedure.

From these quantities, Sensitivity, Specificity is chosen as measurement of accuracy and are calculated using the following equation.

\[
\text{Sensitivity (SE)} = 100 \cdot \frac{TP}{TP + FN}
\]

\[
\text{Specificity (SP)} = 100 \cdot \frac{TN}{TN + FP}
\]

\[
\text{Accuracy (A)} = 100 \cdot \frac{TN + TP}{TN + TP + FN + FP}
\]

In order to examine the data mining classification techniques more closely the sensitivity, specificity and positive predictive value have been calculated using the confusion matrix shown in Table 6.1.

<table>
<thead>
<tr>
<th>Actual</th>
<th>Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Positive</td>
<td>TP</td>
</tr>
<tr>
<td>Negative</td>
<td>FP</td>
</tr>
</tbody>
</table>

**Table 6.1 Confusion Matrix**

Table 6.2 shows that the performance of ensemble method using breast cancer dataset.
<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees.J48</td>
<td>74.28%</td>
</tr>
<tr>
<td>Meta.AdaBoosting</td>
<td>70.89%</td>
</tr>
<tr>
<td>Meta.Bagging</td>
<td>72.85%</td>
</tr>
<tr>
<td>Meta.stacking</td>
<td>76.82%</td>
</tr>
</tbody>
</table>

Table 6.2 performance of Ensemble Method

6.4.2 Performance Graph

From the experimental result, it is observed that the accuracy of the ensemble method is 76.82%. Fig. 6.1 show the performance of ensemble method.

Fig. 6.1 Performance of ensemble Method
6.5 CONCLUSION

This work investigated the technique of combining the predictions of multiple classifiers to produce a single classifier. The resulting classifiers is more accurate than any individual classifier. This work try to increase the efficiency of algorithm on breast cancer dataset using ensemble method. From the experimental results it is observed that the accuracy of the ensemble method is 76.82%.