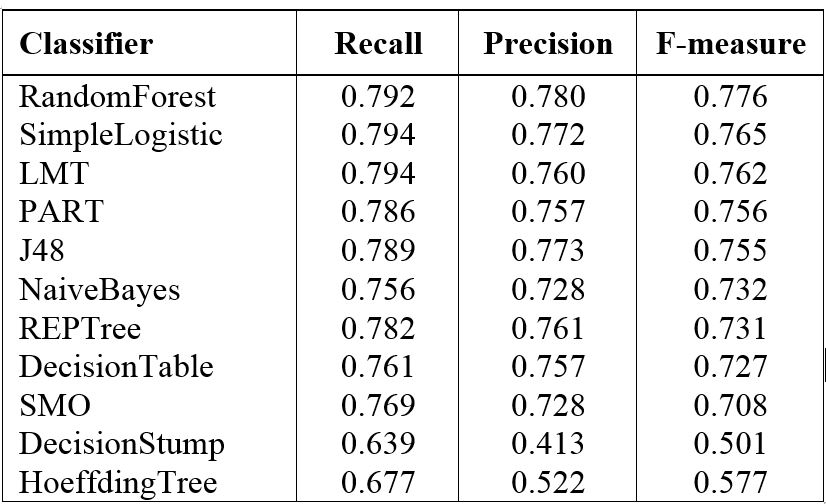
**Classifier Selection**

Dialogue Act Classification means assigning DA types to the individual utterances based on linguistic, perceptual and contextual information. DA classification is a special case of text classification where the text to be classified is the user utterance. So, we need a good classifier to do this classification task. There are different types of classifiers exists such as statistical and probabilistic, decision tree, artificial neural network, support vector machine and rule based classifiers. Each of these classifiers follows different mechanism and has their own pros and cons.

In order to select a suitable classifier to perform the classification, we need to get some measurements about how different classifiers perform on a set of utterances. A set of classifiers have to be trained and tested on a set of utterances. But again the test results affected by the feature set selection, number of training instances and classifier parameters. So it is convenient to use a fixed set of utterances as training set with selected fixed feature sets and without any classifier parameters. There are three widely used measurements call precision, recall and F-measure which can be used to compare the performance of different classifiers.

We have used Weka data mining tool to implement this classification process. For classification task we have used 8000 utterances as training set and 4000 utterances as testing set. As the first step we have tested the classification accuracy by just using the features used for dialogue act recognition in English. From the best performing features Punctuation marks, Trigrams/Bigrams and Frequent words for each tag are the three features used in the related work. The other three features are specific for Sinhala. Using those three features used for English we were able to gain an accuracy of 71.14% in classification using the J48 classifier. Then we have used all six features and classified using the same classifier and we were able to improve the accuracy to 78.68%.

As the next step we have used the same feature set and classified the same data set using different classifiers to model the performance of different classifiers on Sinhala.



We can observe that Simple Logistic and LMT classifiers give the highest recall value. That means they have identified more correctly tagged utterances compared to other classifiers. Also, we noticed that decision tree based classifiers performs well for Sinhala Classification as shown in the table. We obtained several other parameters as well other than presented in the above table. The following diagram shows classifier output on a randomly selected feature set.

