1 Hyper parameter optimization

The goal of hyper parameter optimization is to find the best hyperparameters for the given classifier, in order to optimise the loss function and to avoid overfitting [1]. Hyperparameter optimisation is one of the most important steps in machine learning [2]. Some hyper parameter optimisation algorithms not only find the best hyperameters, but identify those that carry the most weight, [1]. We needed an algorithm to find the best way to tune the hyper parameters of our chosen classifiers. As we were recommended grid search, this was the algorithm we decided to look into.

1.1 Grid Search

Grid Search is an algorithm for hyperparameter optimisation. It does this by iterating over all of the variables in a given range, and selecting the best combination to use in the chosen classifier. As grid search was recommended to us, and is the most widely used way to optimize hyper parameters, we decided to use it in our project.

1.2 Grid Search in Skilearn

SkiLearn has a built in function for Grid Search [3], which is defined below:

```
GridSearchCV(estimator, param_grid, scoring=None, fit_params=None,
    n_jobs=1, iid=True, refit=True, cv=None, verbose=0,
    pre_dispatch='2*n_jobs', error_score='raise',
    return_train_score=True)
```

The param _grid is a dictionary of the parameters you want to choose and the range of parameters to try. The grid search algorithm will then run the given classifier over the entire parameter grid for the best possible combination.

Scoring allows you to choose to train for precision or recall. Precision is the fraction of retrieved instances that are relevant, whilst recall is the fraction of relevant instances that are retrieved.

$$Precision = truepositive/truepositive + falsepositive$$

$$Recall = truepositive/truepositive + falsenegative$$

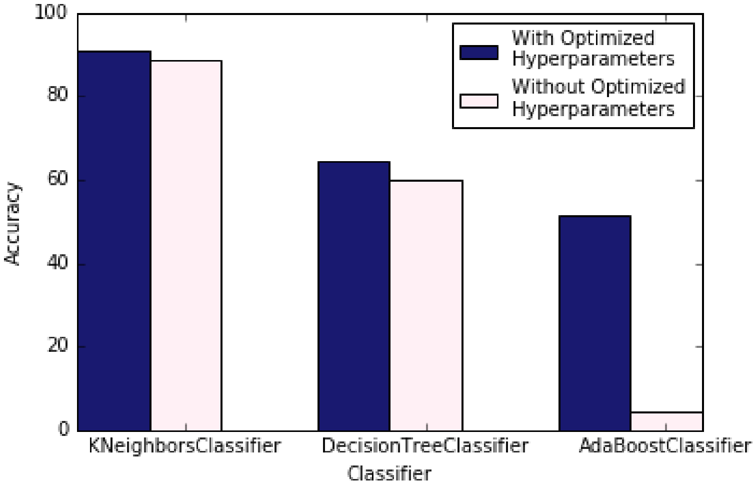
We decided to focuss on precision, as we wanted to tune accurate classifiers.

The output of GridSearch is the a dict, with the names of the hyperparamter as keys, and the optimal result as values.

```
{hyperparamter: optimal value, hyperparameter: optimal value}
```

2 Results

The accuracy of classifiers with optimised hyperparameters were graphed against those without optimised hyperparameters:



In each of the cases, grid search clearly improved the accuracy of the classifier. This highlights the importance of hyperparameter optimisation.

References

[1] James S. Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 2546–2554. Curran Associates, Inc., 2011.

[2] R. Bardenet, M. Brendel, B. Kegl, and M. Sebag. Collaborative hyperparameter tuning. *30th Annual Conference on Machine Learning.*, 28:199–207, 2013.

[3] skilearn. Gridsearchcv. [Online; accessed 15-December-2016].