Decision Trees and Deep Networks for Pattern Classification

Ananya Gupta
University of Manchester
Oxford Road, Manchester
United Kingdom M13 9PL
+44 7784343777
angup143@gmail.com

ABSTRACT
Pattern classification is a complex task that can be useful in a diverse number of fields such as medicine, stock markets, image recognition and autonomous robots. Machine Learning algorithms can be used to recognise complex patterns in big data, however, the performance of the algorithms depends on the type of data being classified. This study compares the performance of decision trees and neural networks on healthcare data. It also uses neural networks and deep learning to classify handwritten digits as a precursor to object recognition in images.

Keywords
Decision Trees; Neural Networks; Deep Learning; Pattern Classification

1. INTRODUCTION
Patterns in big data are interesting since they can provide useful insights about future unseen instances. For instance, identifying complex patterns in the medical histories of people with cancer can help with early stage cancer detection and prognosis.[1][2].

Prediction of trends in the stock markets is another interesting area where pattern recognition is extremely useful in maximising benefit on trading while reducing the risk of loss[3].

Machine Learning as a field is extremely useful and can provide valuable insights, however, there is no one algorithm that can be said to perform well at all classification tasks, rather there are a number of different algorithms and techniques which have their own strengths and weaknesses.

This study explores the use of decision trees and neural networks in pattern recognition and also looks at some developments in deep learning.

2. METHODS AND RESULTS
2.1 Decision Trees
Decision trees are tree like classification models where each node in the tree depicts an input feature and each branch of the tree corresponds to a value of the attribute.

ID3 and C4.5 decision trees[4][5] were used to classify medical datasets from the UCI benchmark repositories (with 5-fold cross validation) and returned varying accuracies on different datasets. The best performance for these classifiers was on the Breast Cancer (Wisconsin) dataset with 94.7% accuracy. However, decision trees could not be used effectively to classify images due to the way they treat input features.

2.2 Neural Networks
Artificial Neural Networks (ANN) are made up of multiple layers of nodes, with interconnections between layers made up of sets of weights. The weights are updated recursively depending on the output error using backpropagation with gradient descent.

An ANN with a single hidden layer was tested on benchmark datasets from the UCI repository and its performance was not as good as that of the decision trees, with an accuracy of 91% for the Breast Cancer (Wisconsin) dataset.

Neural networks are able to overcome the limitations of decision trees in image classification due to their ability to treat input pixels as features[6]. An ANN was tested with the MNIST handwritten digits benchmark database and it provided an accuracy of 92.2% with a single hidden layer and 96.2% with two hidden layers.

2.3 Deep Learning
Deep Convolutional Neural Networks (CNN) are made of up multiple convolutional layers followed by subsampling layers to reduce the data dimensionality.

A Deep Network with two convolutional layers, with 20 and 12 5*5 feature maps respectively followed by mean pooling gave an accuracy of 99.1% on the MNIST benchmark dataset.

The features learned by the network have been identified and give an insight into how the network functions. The results also show that the digits the network misclassified would probably be similarly incorrectly classified by humans as well.

3. CONCLUSIONS AND FUTURE WORK
The DT algorithms performed better than ANNs on the healthcare based datasets, but cannot be extended for classifying patterns in images.

The CNNs had higher accuracy than ANNs when classifying images of handwritten digits since they could detect features in images and also deal with spatially shifted and mirrored images.

Possible future work involves using Gaussian filters to eliminate noise in the feature sets for CNNs which can then be further adapted for object recognition in images.
4. REFERENCES


