



**University of Science and Technology**

**College for Graduate Studies & Academic Advancement**

**COMPARATIVE STUDY BETWEEN  
DECISSION TREE AND NEURAL NETWORK**

**Thesis submitted in Partial Fulfillment of the Requirements for the  
Degree of Master of Computer Science**

**By**

**Rawia Ahmed Mohammed Ibrahim**

**Supervisor**

**Dr. Altyeb Altaher**

**May 2013**

## **. Abstract**

Classification of data objects based on a predefined knowledge of the objects is a data mining and knowledge management technique used in grouping similar data objects together. Classification algorithms have a wide range of applications like churn prediction, fraud detection, artificial intelligence, and credit card rating etc .

In this research present a comparison between neural network and decision tree on application of data classification , we used standard dataset to evaluate the performance of the neural network and decision tree algorithms in terms of accuracy and classification time. The dataset wear preprocessed using Information Gain algorithm to select the best feature for the classification.

The experimental results show that the neural network algorithm (radial basis function ) achieved the highest classification accuracy , while the decision tree algorithm (j48) achieved the minimum classification time.

## المستخلص

تصنيف كائنات البيانات بناءا علي المعرفة المسبقة للكائنات يعتبر تنقيب للبيانات و إدارة للمعرفة ،ويستخدم لتجميع كائنات البيانات المتشابهة مع بعضها. خوارزميات التصنيف ذات مدي واسع من التطبيقات مثل :

churn prediction , fraud detection , artificial intelligence , credit card rating. هذا البحث يقدم مقارنة بين خوارزميات شجرة القرار و خوارزميات الشبكات العصبية في تطبيق تصنيف البيانات.

تم استخدم بيانات قياسية لتقييم آدا شجرة القرار والشبكات العصبية من حيث الدقة وزمن التصنيف ، تمت معالجة البيانات باستخدام خوارزمية مكسب المعلومات (Information Gain) لاختيار أفضل الخصائص للتصنيف . نتائج التجارب ان خوارزمية الشبكات العصبية (radial basis function) حققت اعلي دقة تصنيف ، بينما خوارزمية شجرة القرار (j48) حققت اقل زمن للتصنيف .

## **Introduction1 .1**

Classification can be described as a supervised learning algorithm in the machine learning process. It assigns class labels to data objects based on prior knowledge of class which the data records belong.

In classification, a given set of data records is divided into training and test data sets. The training data set is used in building the classification model, while the test data record is used in validating the model. The model is then used to classify and predict new set of data records that is different from both the training and test data sets [2].

Supervised learning algorithm (like classification) is preferred to unsupervised learning algorithm (like clustering) because its prior knowledge of the class labels of data records makes feature/attribute selection easy and this leads to good prediction/classification accuracy.

Some of the common classification algorithms used in data mining and decision support systems are: neural networks , logistic regression , Decision trees

## 1.2 Research Problem

Classification of data objects based on a predefined knowledge of the objects is a data mining and knowledge management technique used in grouping similar data objects together. There are many classification methods , including the decision tree and neural network .

The problem of this research is to compare between the decision tree algorithms and neural network algorithms , to find the algorithm which achieve the highest classification accuracy and minimum classification time .

## 1.3 Research Objective

The objectives of this thesis include the following:

- Comparing the classification accuracy of the decision tree and neural networks algorithms .
- Comparing the classification time of decision tree and neural networks algorithms .
- Comparing the Root mean squared error (RMSE) of decision tree and neural networks algorithms .

## **.4 Thesis Structure 1**

Chapter one is an introduction .

Chapter two demonstrates basic information about machine Learning and classification algorithm ( Decision Trees ,Neural Networks). and detailed description of algorithms decision tree (j48, simple cart) and neural network (RBF network, MLP) , also explains the advantages and disadvantages and comparison of Decision Trees and Neural Networks algorithm.

Chapter three

Chapter four presents the implementation and the result of the comparison between the Decision Trees and Neural Networks.

Chapter five shows the conclusion and Future work .