



**UNIVERSITY OF SCIENCE AND TECHNOLOGY**  
**COLLEGE OF POSTGRADUATE STUDIES AND**  
**ACADEMIC ADVANCEMENT**

Faculty of Computer Science and Information Technology

**A Comparative Study in WEKA between the Classifier algorithms:  
AD, J48 and LADTree**

By

Marwa Awad Sharfi Abdalla

A Thesis Submitted to the College of Graduate Studies and Academic  
Advancements in Partial Fulfillment of the Requirement for the Degree of Master of  
Science in Information Systems

Supervisor

Dr. Hassan FadlAllah

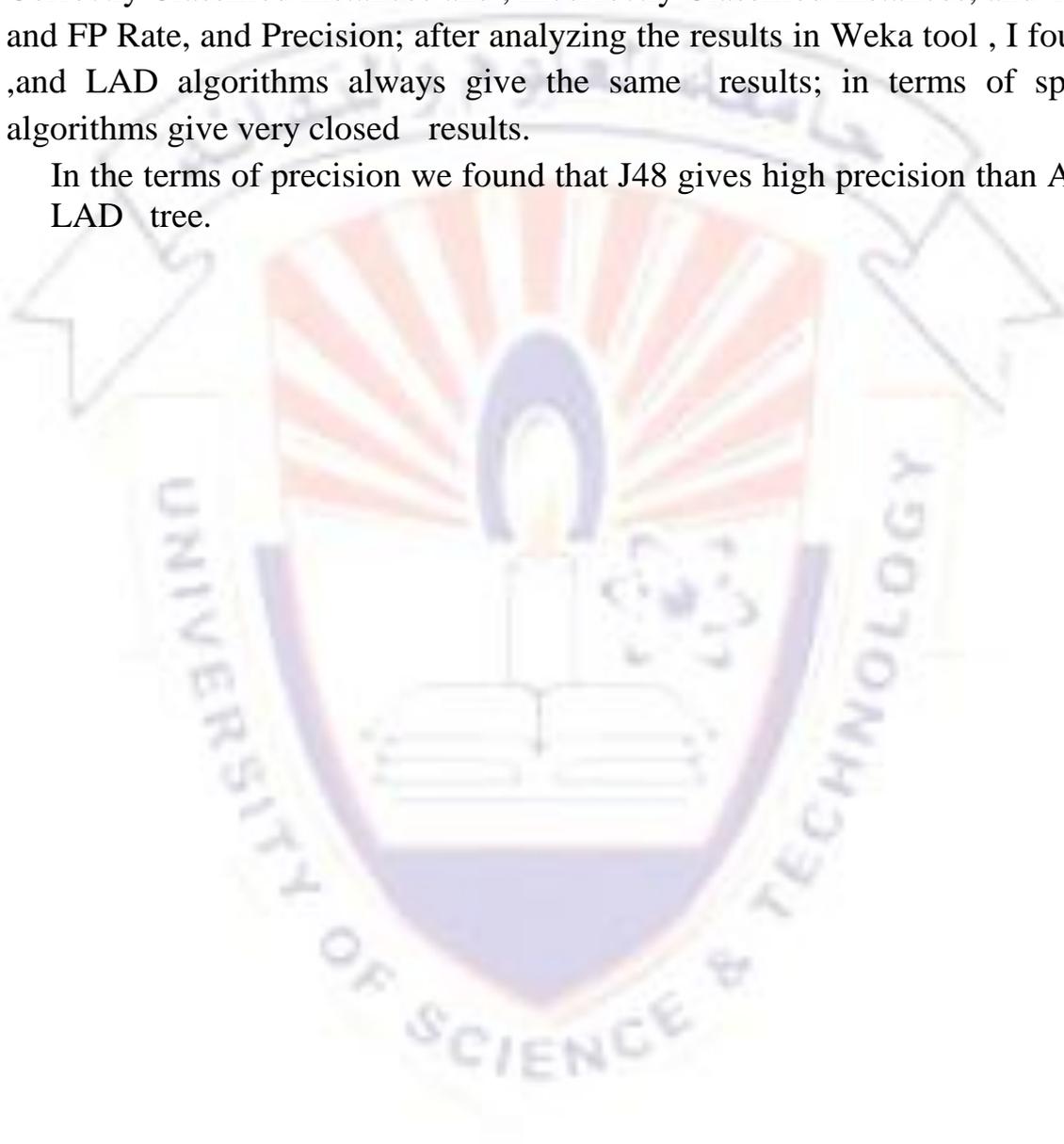
September 2015

**Abstract:**

This study measure the performance of AD , J48 and LADTree algorithm in WEKA tool, as applied to parts of the results of students some of Kibieda international schools for years 2012-2013 , 2013-2014 , and 2014-2015 .

The model used to measure the performance of the algorithms based on Correctly Classified Instances and , Incorrectly Classified Instances, and TP Rate, and FP Rate, and Precision; after analyzing the results in Weka tool , I found AD ,and LAD algorithms always give the same results; in terms of speed all algorithms give very closed results.

In the terms of precision we found that J48 gives high precision than AD and LAD tree.



## المستخلص:

هذه الدراسة تعمل علي قياس أداء خوارزمية ،48لو LADTree، و AD في برنامج الويكا (WEKA)، المطبقة علي بعض نتائج مدارس كبيدة العالمية لسنة 2012-2013 ، و 2013-2014، و 2014-2015 ، وكان قياس اداء الخوارزمية مبني علي صحة عمل الخوارزميات في استخراج النتائج ومدى دقتها؛ وبعد تحليل النتائج بواسطة برنامج الويكا (WEKA) وجدنا ان نتائج الدارسة تتفق دائما لدى خوارزميتي LAD، و AD ؛ أما من ناحية السرعة تتفق جميع الخوارزميات في سرعه اعطاء النتائج.

كما وجدنا أيضا ان خوارزمية J48 تعطي نتائج اكثر دقة من خوارزميتي LAD ، و AD.



## 1.1 INTRODUCTION

We live in a world where vast amounts of data are collected daily. In the areas and different departments, so it is necessary to analyze this data as needed here and this is through data mining.

Data mining is an interdisciplinary field, the confluence of a set of disciplines, including database systems, statistics, machine learning, visualization, and information science depending on the data mining approach used, techniques from other disciplines may be applied, such as neural networks, fuzzy and/or rough set theory, knowledge representation, inductive logic programming, or high performance computing.

Depending on the kinds of data to be mined or on the given data mining application, the data mining system may also integrate techniques from spatial data analysis, information retrieval, pattern recognition, image analysis, signal processing, computer graphics, Web technology, economics, business, bioinformatics, or psychology.[1]

Data Mining is a step in the Knowledge Discovery in Databases (KDD) process consisting of applying data analysis and discovery algorithms that, under acceptable computational efficiency limitations, produce a particular enumeration of patterns over the data [2], as shown in Figure 1.1.



Figure(1.3): Data mining steps

## **1.2 Research Problem:**

In this research I compare algorithms AD Tree, J48 and LAD Tree, to measure the performance of each one.

## **1.3 Research Question:**

Is there a difference in the performance of these algorithms (AD Tree, J48 and LAD Tree).

## **1.4 Research Objectives:**

1. Comparison of the results of each algorithm.
2. Determine the best algorithm.
3. To investigate algorithms (AD Tree, J48 and LAD Tree)

## **1.5 Data Set:**

The data set that used in this research have been collected from kibeida International Schools.

## **1.6 Methodology:**

This is an Analytic experimental method firstly I collected data from kibeida International Schools and then we prepared the data set in the suitable format before starting the experiments. Then we used the selected algorithms (AD Tree, J48 and LAD Tree) with the selected data set. we applied the management model which based on Correctly Classified Instances, Incorrectly Classified Instances, Precision, time, True Positive (TP) rate, and False Positive FP rate. Finally we discussed the results of experiments.

## **1.7 Research Tools:**

In this research, I used WEKA version 3-6-11.

## **1.8 Research Organization:**

In this research there are five chapters. First chapter deals with the proposed, subject of research, objective of research, method used in this research and research field. Second chapter include some concepts which deal with Data mining generally and some studies in Data mining. Third chapter deal with weka tool, classification and the algorithms used in detail. Fourth chapter include practical part in classification, analysis data and conclude the result. Finally the conclusion and recommendation for future work.