# الآيـة

**قال تعالى:** (أَمَّنْ هُوَ قَانِتٌ آنَاء اللَّيْلِ سَاجِدًا وَقَائِمًا يَحْذَرُ الآخِرَةَ وَيَرْجُو رَحْمَةَ رَبِّهِ قُلْ هَلْ يَسْتَوِي الَّذِينَ يَعْلَمُونَ وَالَّذِينَ لا يَعْلَمُونَ إِنَّمَا يَتَذَكَّرُ أُوْلُوا الأَلْبَابِ)

**صدق الله العظيم**

**سورة الزمر الآية (9)**

# Dedication

Dedicated to

My parents

My sister, my brothers

Colleagues and all my friends

To everyone who tried to guide me to a better life

With my love

# ACKNOWLEDGEMENT

Praise is to Allah, I would like to express my gratitude to many people who helped me in different ways with the development of this research. Without their continuous support and guidance, the completion of my research would be impossible. I wish to express my most sincere thanks to my supervisor **Dr. Shaza Merghani**, who had a major role in ending this research in this way and gave me a lot of help and assistance, it is my luck to study and work under her guidance. I wish to thank also my parents, my sister, my brothers, and special deep thanks are extended to Dr. EzzEldin Mohammed, Hisham abdallha, Ustaz Mojtaba FadlAlmola, Mosab Hassan altom and mosheera kabashi for their kind help, assistance and support. Finally, my deepest thanks are extended to all the staff of Faculty of Economics and Business Studies, University of Kordofan, and all the other individuals who helped me during the study period.

# ABSTRACT

Data mining is the automatic search of huge data to discover patterns and trends that go beyond simple analysis. The high rate of student's failure is one of the major problems and represents a worry for many universities because no rule for distribution specialization, it's difficult to determine which specialization is better for the student. This study proposed a model for Predicting the best specialization for the student. Weka data mining tool used to evaluate performance of student's. The data consists of academic information contain 1402 records from University of Kordofan, Faculty of Economics. The experiment conducted using three algorithms Naive Bayes classifier, j48 and Random Forest to predict best student's specialization. Then, Apriori algorithm was also applied to find close correlation between courses and specialization. Results showed that the best technique is J48 classifier was achieved 97.6% of accuracy which is better than Naïve Bayes and Random. The Apriori algorithm used for generates strong rules that helped to identify if the student's academic qualify to study the Specialization and relationship between specialization and courses. The experiment conducted generated strong rules their number are fifteen with (Support 11% and Confidence 100%).

# المستخلص

تنقيب البيانات هو البحث التلقائي عن البيانات الضخمة لاكتشاف الأنماط والاتجاهات التي تتجاوز التحليل البسيط. يعد ارتفاع معدل فشل الطالب إحدى المشكلات الرئيسية ويمثل مصدر قلق للعديد من الجامعات لأنه لا توجد قاعدة محدده لتوزيع الطلاب للتخصصات المختلفة بالكلية، ومن الصعب تحديد التخصص الأفضل للطالب.

اقترحت هذه الدراسة نموذجا لتوقع أفضل تخصص للطالب. تستخدم أداة Weka لتنقيب البيانات ولتقييم أداء الطالب. تتكون البيانات من معلومات أكاديمية تحتوي على 1402 سجل من جامعة كردفان، كلية الاقتصاد والأعمال. أجريت التجربة باستخدام ثلاث خوارزميات Naive Bayes مصنف، J48 وRandom Forest للتنبؤ بالتخصص الأفضل للطالب. كما تم تطبيق خوارزمية Apriori أيضا لإيجاد معامل الارتباط الوثيق بين المقرارت الدراسية والتخصص.

أظهرت النتائج أن أفضل تقنية هي المصنف J48 الذي تم تحقيقه بنسبة 97.6٪ من الدقة وهو أفضل من Naïve Bayes وRandom. تستخدم خوارزمية Apriori في إنشاء قواعد قوية ساعدت في تحديد ما إذا كان الطالب الأكاديمي مؤهلاً لدراسة التخصص والعلاقة بين التخصص والمقررات الدراسيه. التجربة التي تم إجراؤها ولدت قواعد قوية وعددهم خمسة عشر (دعم 11٪ والثقة 100٪).

## List of contents

## List of Tables

## List of Figures

## List of Abbreviations

| | |
|---|---|
| DM | Data Mining |
| EDM | Educational Data Mining |
| GPA | Grade Point Average |
| WEKA | Waikato Environment for Knowledge Analysis |
| CSV | Comma-separated values (Comma-delimited) |
| ARFF | Attribute Relation File Format |