اسم الطالب: أمل مهدي علي ال حسان.

اسم المشرف: بسام ظفر، احمد معين.

عنوان الرسالة: استخدام تقنيات تنقيب البيانات للتنبؤ بالأداء الأكاديمي للطلاب.

تعد القدرة على التنبؤ بالأداء الأكاديمي للطلاب أمرًا بالغ الأهمية لكل مؤسسة تعليمية تهدف إلى تحسين أداء طلابها ومثابرتهم. على الرغم من انه قد تم دراسة التنبؤ بأداء الطلاب على نطاق واسع، إلا أنها لا يزال يمثل تحديًا وقضية معقدة لأن أداء الطلاب يتأثر بالعديد من السمات. درست العديد من الأعمال السابقة تأثير السمات المختلفة على الأداء الأكاديمي للطلاب، لكن عدد قليل من الأعمال ركزت على تأثير درجات تقييم الطلاب ونشاطهم عبر الإنترنت على نظام إدارة التعلم (LMS) معًا.

لذلك كان الهدف من هذا البحث هو التحقق من تأثير درجات التقييم وبيانات النشاط عبر الإنترنت في شكل مشترك ومنفصل لتحديد تأثيرها على أداء الطلاب. من خلال استخدام مجموعات البيانات الفرعية لإنشاء نماذج التنبؤ بأداء الطلاب. استنادًا إلى خمسة خوارزميات تصنيف هي شجرة القرار والغابة العشوائية ومتسلسل الحد الأدنى الأمثل والإدراك متعدد الطبقات والانحدار اللوجستي. بالإضافة إلى تطبيق خوارزميات اختيار السمات باستخدام طريقتي المرشح (Filter) والمغلف (Wrapper) لتحديد أهم السمات التي تؤثر على أداء الطلاب. تم تنفيذ هذه الدراسة باستخدام البيانات المستخرجة من نظام إدارة التعلم (LMS) لـ 241 طالب جامعي في قسم نظم المعلومات، كلية الحاسبات وتقنية المعلومات في جامعة الملك عبد العزيز.

أظهرت نتائج هذا البحث أن درجات التقييم هي أهم السمات التي تؤثر على أداء الطلاب، خاصة علامة الواجبات والامتحان النهائي .علاوة على ذلك، كان أداء نماذج التنبؤ التي تم إنشاؤها بناءً على درجات التقييم بشكل منفصل أو بالاشتراك مع بيانات النشاط أفضل من النماذج القائمة على بيانات النشاط وحدها. تشير هذه النتائج إلى أن درجات التقييم لها تأثير أكبر على أداء الطلاب، بينما بيانات النشاط لها تأثير أقل. أيضًا، النماذج التي شملت بيانات التقييم والنشاط تعمل معًا بشكل جيد، وهذا يعني أنه من المهم تضمين درجات التقييم مع بيانات النشاط للتنبؤ بأداء الطلاب. كما تفوقت الغابة العشوائية على خوارزميات التصنيف الأخرى من خلال الحصول على أعلى دقة، تليها شجرة القرار.

Student name: Amal Mahdi Ali Alhassan.

Supervisor name: Bassam Zafar and Ahmed Mueen.

Thesis title: Using data mining techniques to predict students' academic performance.

Ability to predict students' academic performance is critical for any educational institution aims to improve their students' performance and perseverance. Although students' performance prediction studied widely, it still represents a challenge and complex issue because student performance influenced by different features. Previous works have studied the impact of different features on students' performance, but few works have focused on the impact of students' assessment grades and online activity on Learning Management System (LMS) together.

This research aims to investigate the impact of assessment grades and online activity data jointly and separately to determine their effect on students' performance. By employing sub-datasets to create prediction models for students performance. Based on five classification algorithms are decision tree, random forest, sequential minimal optimization, multilayer perceptron, and logistic regression. In addition to feature selection task based on filter and wrapper methods for identifying the most important features that affect student performance. This work carried out using data extracted from LMS for 241 undergraduate students in Information Systems department, Faculty of Computing and Information Technology at King Abdulaziz University.

Results showed that assessment grades is the important features that affect students' performance, especially assignments mark and final exam. Moreover, prediction models created based on assessment grades separately or jointly with activity data were perform better than models using activity data alone. That indicate assessment grades had a greater impact on student performance, while activity data had a lower impact. However, models that include assessment and activity data together work well, which means it is important to include assessment grades with activity data for predicting students' performance. Random forest performed better than other algorithms, followed by a decision tree.