استخدام تقنيات تنقيب البيانات لتحليل استبيان الطلاب لتقييم العملية التدريسية على ضوء معايير الجودة الشاملة

Using data mining techniques to analyze students Survey to evaluate the teaching process according to total quality standards

رانية عبد العال عثمان سليمان
أستاذ محاضر
Nucsit016@gmil.com
كلية علوم الحاسوب و تقانة المعلومات - جامعة النيلين
الخرطوم ، السودان

1. المستخلص :
في الآونة الأخيرة حرصت أغلب المؤسسات الأكاديمية على إيجاد العوامل التي تزيد من نتائج التعليم وأصبحت جودة التعليم هدف مهم تسعى جميع المؤسسات التعليمية اليه خاصة في ظل توفر قدر ضخم من البيانات التي يمكن الاستفادة منها للحصول على نتائج ممتازة. و كثر استخدام خوارزميات تنقيب البيانات لتحليل هذه العوامل. في هذه الدراسة تم بناء نموذج لتحديد أكثر العوامل المؤثرة في تقييم الطلاب للأستاذ الجامعي وذلك عبر تصنيف اجابات الطلاب المتحصل عليها بواسطة استبيان مكون من 36 فقرة، بالاضافة لبعض البيانات الأخرى الخاصة بالأستاذ مثل معدله التراكمي الذي تخرج به، الدرجة العلمية ، وعدد سنوات الخبرة. كان حجم مجموعة البيانات 709 صف، وتم استخدام ثلاث خوارزميات تصنيف، هي: Random Forest, Decision Tree و Naïve Bayes، كما تم استخدام خوارزمية Sequential Feature Selection(SFS) وذلك لسببين؛ الأول هو تقليل عدد السمات لزيادة دقة النتائج والثاني هو الحصول على السمات الأكثر صلة والتي تؤثر على النتيجة النهائية. تمت المقارنة بين اداء الخوارزميات وبعد الاختبار والتحقق كانت الـ Random Forest هي الأكثر دقة .
يمكن تقسيم عملية تصميم النموذج النهائي الى أربع مراحل، المرحلة الأولى هي المعالجة المسبقة (preprocessing)، و الثانية هي تطبيق SFS للعثور على السمات المرتبطة ببعضها، في المرحلة الثالثة تم تطبيق خوارزميات التصنيف الثلاث وأخيرا المرحلة الرابعة تم التحقق من النموذج عن طريق تقنية التحقق المتقاطع الطبقي Stratified Cross Validation، وذلك لأن البيانات غير متوازنة imbalanced.

2. الكلمات المفتاحية : تقييم الأداء، الجودة الشاملة في التعليم ، تنقيب البيانات التعليمية، بيانات غير متوازنة، خوارزميات التصنيف.

3. Abstract:
Recently, most academic institutions have been keen to find factors that increase the results of education, and the quality of education has become an important goal that all educational institutions seek, especially in light of the availability of a huge amount of data that can be used to obtain excellent results and the frequent use of data mining algorithms to analyze these factors. In this study, a model was built to determine the most influencing factors in students' evaluation of a university professor, by classifying students' answers obtained through a survey consisting of 36 items, in addition to some other data related to the professor, such as his cumulative grade point average, academic degree, and the number of years of experience. The size of the data set was 709 rows, and three classification algorithms were used: Random Forest, Decision Tree, and Naïve Bayes. The Sequential Feature Selection (SFS) algorithm was used for two reasons, the first is to reduce the

number of features to increase the accuracy of the results and the second is to get the most relevant features that affect the final result. The comparison has been made between the performance of the algorithms and after testing and verification, the Random Forest was the most accurate. The process of designing the final model can be divided into four stages, the first stage is preprocessing, the second is the applying of SFS to find the features associated with each other, the third stage is applied three classification algorithms and finally the fourth stage is validation of the model by stratified cross-validation technique, because the data is imbalanced.

4. Keywords: performance evaluation, total quality in education, educational data mining, imbalance data, classification algorithms

5. المقدمة :

يمثل التعليم الجامعي في الدول المتقدمة والنامية مصدر إشعاع علمي، وثقافي، وحضاري؛ لأنَّه المسئول عن إعداد الكفاءات المتخصصة اللازمة للنهوض بأعباء التنمية في مختلف المجالات.

وما من شك أنَّه لن يستطيع أن يقوم بتلك الوظيفة إلاَّ إذا توافرت له الإمكانات التي تعينه، وفي مقدمتها أستاذ الجامعة الذي يستطيع بإمكاناته العلمية والخُلقية والنفسية أن يساهم مساهمةً فعالةً في تحقيق الأهداف المنشودة من الجامعة . [1]

ونظرًا لأهمية أستاذ الجامعة في تحقيق أهداف الجامعة، وللتباين في القيمة الكمية والنوعية في تحقيق هذه الأهداف بين أساتذة الجامعة؛ فمن الضروري أن يخضع عملهم للتقويم؛ ذلك أنَّ التقويم الموضوعي يساعد متخذي القرارات والمنفِّذين على معرفة مدى تحقيق الأهداف المنشودة . كما يعمل على تحسين نوعية التعليم الجامعي وتطويره من خلال زيادة فاعلية أداء أستاذ الجامعة. ومن هنا جاءت الدراسة الحالية لتقييم الأداء الأكاديمي لأستاذ الجامعة وما ينبغي أن يكون عليه في جامعة النيلين – كلية علوم الحاسوب و تقانة المعلومات- من وجهة نظر الطلاب؛ ممَّا يتيح الفرصة لأستاذ الجامعة وإدارتها للتعرف على نواحي القوة وتشجيعها، والوقوف على نواحي الضعف لعلاجها.

وجاءت فكرة البحث بسبب قلة أو انعدام الدراسات التي أجريت في مجال تقييم الأداء الأكاديمي لأستاذ الجامعة وما ينبغي أن يكون عليه في السودان عامة، وجامعة النيلين بشكل خاص.

6. الدراسات السابقة :

1. في هذه الدراسة [2] تم استخدام نموذج التصنيف classification model للتنبؤ بأداء الأساتذة وذلك بالاعتماد على بيانات من الطلاب و الزملاء ، كما تم تعريف العوامل التي تؤثر على أداء الأستاذ. استخدم الباحث خوارزمية : J48 decision tree(random tree) and naïve bayes . بإستخدام أداة ويكا weka . وتم التوصل الى أن رأي الطالب وحده غير كافي لتقييم الأستاذ لذلك لا بد من استخدام بيانات أخرى مثل بيانات الزملاء ورؤساء الأقسام .
2. تهدف هذه الدراسة [3] للتنبؤ بجودة أداء الاستاذ الجامعي و تحليل العوامل التي تؤثر على إستفادة الطلاب لتحسين جودة النظام التعليمي . تم العمل علي بيانات طلاب في جامعة تركية واستخدم الباحث خوارزميات التصنيف : J48 Decision tree, Multilayer Perception, Naïve Bayes, and Sequential Minimal Optimization . وبعد مقارنة الخوارزميات الأربعة اظهرت الدراسة مدى أهمية استخدام تنقيب البيانات في مجال التعليم ، ايضا توصلت الدراسة الى أن استخدام بيانات الطالب لتقييم المقررات مفيد في التنبؤ بالعوامل التي تؤثر على تحصيله وعلى التنبؤ بجودة أداء الأستاذ .
3. في هذه الدراسة [4] تم انشاء نظام تصنيف يعتمد على طبقتين: الشبكة العصبية Artificial Neural Network (ANN) و شجرة القرار Decision Tree . تم اختبار النظام بنجاح باستخدام بيانات دراسة حالة من الجامعة النيجيرية . البيانات هي: المؤهلات الاكاديمية لأعضاء هيئة التدريس وكذلك الخبرة درجات الطلاب في المواد التي يدرسونها .اظهرت النتائج انه من بين الست سمات المستخدمة تم تصنيف الخبرة و المؤهل كأفضل سمتين ساهمتا في اداء أعضاء هيئة التدريس . ايضا بالاخذ في الاعتبار الزمن المستغرق في بناء النماذج و مستوى دقة اداء شجرة القرار C4.5 تفوقت على الخوارزميتين MLP و ID3 .
4. تبحث هذه الدراسة [5] في العوامل المرتبطة بتقييم أداء الأساتذة باستخدام تقنيتين : الانحدار التدريجي stepwise regression و أشجار القرار decision trees . تم جمع البيانات من تقييمات الطلاب بالإضافة إلى ذلك تم تضمين بعض المتغيرات الأخرى . أشارت النتائج إلى أن حالة توظيف الأستاذ ، و عبء العمل ، وحضور الطلاب و مستوى الطلاب الأكاديمي هي أبعاد مهمة في تقييم الأستاذ .
5. حاولت هذه الورقة [6] الإجابة على أسئلة مثل : هل لدى الطلاب النضج و المعرفة الكافية لتقديم ملاحظات مفيدة يمكن الاعتماد عليها لتحسين القدرات التدريسية للأساتذة ؟ ، هل صعوبة المقرر لديه علاقة قوية مع التصنيف الذي يعطيه الطالب للأستاذ؟ . تم استخدام بعض تقنيات تنقيب البيانات الإحصائية الحديثة مثل :

Support vector machines, classification and regression trees, boosting, random forest, factor analysis, k-Means clustering. hierarchical clustering.

تم استخدام جوانب مختلفة من البيانات من منظورين مختلفين ، التعلم الخاضع للإشراف و الغير خاضع للإشراف supervised and unsupervised learning. البيانات التي تم تحليلها في هذه الورقة تم جمعها من جامعة تركية ، تم اكتشاف بعض الانماط مثل الارتباط القوي بين جدية الطلاب (تقاس بالحضور) و بين نوع الدرجات التي يمنحونها لاساتذتهم.

6. في هذه الورقة [7] تم استخدام 4 خوارزميات تصنيف : decision tree algorithms, support vector machines, artificial neural networks, and discriminant analysis.
7. تم مقارنة اداءها بالتطبيق على بيانات تم الحصول عليها بواسطة استبيان الطلاب لتقييم المقرر الدراسي . أظهرت خوارزمية C5.0 نتائج أكثر دقة مقارنة بالخوارزميات الأخرى , و بتحليل متغيرات كل نموذج تصنيفي تم التوصل الى انه العديد من اسئلة الاستبيان ليست ذات صلة بمشكلة الدراسة ، علاوة على ذلك اظهرت النتائج ان نجاح الاساتذة يعتمد على اهتمام الطلاب باالمقرر وعليه يمكن استخدام النتائج لتحسين ادوات القياس.

باستقراء الدراسات السابقة يلاحَظ أنَّ العديد من الدراسات تناولت موضوع أستاذ الجامعة من جوانب متعددة ومختلفة؛ لأهمية أستاذ الجامعة، ودوره في خدمة الجامعة والمجتمع. وقد اختلفت هذه الدراسات وفقًا لطبيعتها، وأهدافها، والعينة المختارة، والنظام التعليمي، والمجتمعات التي أجريت فيها. والدراسة الحالية تتفق مع الدراسات السابقة من حيث موضوعها -أستاذ الجامعة -، ومع ذلك فإنَّ الدراسة الحالية تكتسب وضعًا مختلفًا في معالجتها لتقويم الأداء الأكاديمي لأستاذ الجامعة وفقا لعدد من المتغيرات المتمثلة في فقرات الاستبيان الذي يمثل إحدى ادوات جمع بيانات الدراسة ، كما أن الدراسة الحالية تميزت بقدرتها على ايجاد حلول للتعامل مع البيانات الغير متوازنة وكان ذلك واضحا في تحسن أداء النموذج .

7. تنقيب البيانات التعليمية Educational Data Mining: [9]

مع إنشاء المؤتمر الدولي السنوي لاستخراج البيانات التعليمية ومجلة تعدين البيانات التعليمية في عام 2008، برزت EDM كمجال بحثي موثوق به (Baker et al.، 2010). تقدم الجمعية الدولية لتعدين البيانات التعليمية، التي تستضيف المؤتمر الدولي لتعدين البيانات التعليمية وتنشر مجلة تعدين البيانات التعليمية ، هذا التعريف لـ EDM:

"التنقيب في البيانات التعليمية هو تخصص ناشئ، يهتم بتطوير طرق لاستكشاف البيانات الفريدة والواسعة النطاق المتزايدة التي يتم الحصول عليها من الإعدادات التعليمية، ويستخدم هذه الأساليب لفهم الطلاب بشكل أفضل والإعدادات التي يتعلمون فيها"

وفقًا لجمعية التنقيب عن البيانات التعليمية الدولية (2011) ، غالبًا ما تتكون المعلومات في أي سياق تعليمي من مستويات هرمية متعددة ، والتي لا يمكن تحديدها مسبقًا ولكن يجب التحقق منها من خلال الخصائص الموجودة في البيانات. تعتبر العوامل ، مثل الوقت والتسلسل والسياق مهمة أيضًا في الاعتبار عند دراسة البيانات التعليمية. على سبيل المثال ، يمكن تحليل السلوكيات التعليمية للطلاب (مشاركة الطلاب، وتكرار تسجيل الدخول، وعدد رسائل الدردشة، ونوع الأسئلة المرسلة إلى المعلم) جنبًا إلى جنب مع درجاتهم النهائية.

8. البيانات الغير متوازنة Imbalanced Data :

أحد التحديات المهمة في التنقيب عن البيانات هو التعامل مع البيانات غير المتوازنة في التصنيف. ، يظهر عدم التوازن عندما يتم توزيع البيانات بشكل غير متساوٍ إلى فصائل ؛ قد تحتوي بعض الفصائل على كمية كبيرة من البيانات تسمى فصائل الأغلبية Majority class وبعضها قد يحتوي على حالات قليلة فقط من البيانات تسمى فئات الأقليات Minority class . يتسبب هذا التوزيع غير المتكافئ في أداء متحيز للمصنفات التقليدية لأنها تأخذ في الاعتبار معدل الخطأ وليس توزيع البيانات ، وبسبب قلة عدد حالات البيانات ، يتم تجاهل فئات الأقليات في نتيجة التصنيف الإجمالية. تظهر هذه المشكلة في العديد من تطبيقات العالم الحقيقي ، مثل قطاع الرعاية الصحية ، و اكتشاف الانسكاب النفطي ، و اكتشاف الاحتيال في استخدام بطاقات الائتمان ، و نمذجة للثقافات ، و اكتشاف التسلل في الشبكات ، وتصنيف النصوص ، وما إلى ذلك. [17]

9. الجودة الشاملة في التعليم : [10]

تعرف الجودة الشاملة في التعليم بأنها فلسفة شاملة للحياة والعمل في المؤسسات التعليمية، تحدد أسلوباً في الممارسة الإدارية بهدف الوصول إلى التحسين المستمر لعمليات التعليم والتعلم وتطوير مخرجات التعليم على أساس العمل الجماعي، بما يضمن رضا الأساتذة والطلبة وأولياء الأمور وسوق العمل

10. إدارة الجودة الشاملة في التعليم العالي: [10]

تزايد الاهتمام العالمي في العقدين الأخيرين من القرن العشرين بجودة التعليم، ومن المتوقع أن يزداد هذا الاهتمام في المستقبل نظرا للشكوى العالمية من انخفاض مستويات الجودة في التعليم الأساسي أو التعليم العالي على حد سواء، وقد بدأ الاهتمام المتزايد بإدارة الجودة الشاملة في التعليم في منتصف السبعينيات من القرن الماضي، وذلك للعديد من العوامل التي كان لها الأثر في هذا الاهتمام، والتي من أبرزها: التغيرات الاقتصادية المصاحبة للانفجار العلمي والتكنولوجي والتوسع في التعليم، وزيادة الإقبال عليه في جميع المراحل التعليمية بما فيها التعليم الجامعي والعالي.

أما عن أهمية تطبيق إدارة الجودة الشاملة في مؤسسات التعليم العالي فيرى البعض أن الجودة الشاملة هي مجموعة الخصائص / السمات التي تعبر بدقة وشمولية عن جوهرها وحالتها بما في ذلك كل أبعادها : مدخلات وعمليات ومخرجات وتغذية راجعة، وأن تطبيق إدارة الجودة الشاملة يتطلب وجود أرضية متينة في كافة بنياتها التنظيمية والإدارية والاجتماعية داخل المنظمة بحيث تعمل على توفير المناخ المناسب والثقافة الملائمة المؤمنة بإمكانية تطبيقها. لذا فلابد من توافر القناعة التامة لدى الإدارة العليا والمجالس

الأكاديمية فيها بأهمية مفهوم إدارة الجودة الشاملة وبضرورة تطبيقه واتخاذه في مقدمة استراتيجياتها والعمل على نشر قناعاتها لجميع العاملين فيها.

11. معايير التعليم العالي في السودان لتقييم أعضاء هيئة التدريس : [16]

المعيار العام هو هل للمؤسسة العدد الكافي و المؤهل من أعضاء هيئة التدريس في مختلف التخصصات العلمية ، وتعمل على تقويم أداءهم ، وبناء قدراتهم لأداء وظائفهم بكفاءة و فاعلية .

المؤشرات :

1. هل يتم تعيين أعضاء هيئة التدريس وفقا لحاجات الأقسام و التخصصات المختلفة ؟ و هل تحدد أعباءهم و مهامهم المختلفة وفق الوصف الوظيفي ؟
2. كم عدد أعضاء هيئة التدريس المتفرغين على رأس العمل و المتعاونين في الثلاث سنوات الأخيرة موزعين حسب الجنسية و التخصص و الدرجة العلمية و الرتبة العلمية ؟
3. هل توزع أعباء أعضاء هيئة التدريس بصورة متوازنة في المجالات المختلفة ؟
4. الى أي مدى يتلاءم التخصص العلمي لعضو هيئة التدريس مع المقررات التي يقوم بتدريسها ؟
5. الى أي مدى يتناسب عدد أعضاء هيئة التدريس في كل تخصص مع أعداد الطلاب في التخصص المعني ؟
6. هل تتبع المؤسسة سياسة التفرغ العلمي لأعضاء هيئة التدريس؟ و كم عدد المستفيدين في الثلاث سنوات الأخيرة ؟
7. كم عدد أعضاء هيئة التدريس المعارين و ممن هم في إجازات بدون مرتب و المتسربين في الثلاث سنوات الأخيرة ؟
8. هل يتم تقويم أعضاء هيئة التدريس؟ و ما الجهات التي تشارك في تقويم أداءهم ؟

12. خوارزمية Random Forest Classifier : [13]

كيف تعمل خوارزمية الغابة العشوائية ؟

فيما يلي الخطوات الأساسية المتضمنة في تنفيذ خوارزمية الغابة العشوائية:

1. اختر N سجلات عشوائية من مجموعة البيانات.
2. قم ببناء شجرة قرار بناءً على هذه السجلات N.
3. اختر عدد الأشجار التي تريدها في الخوارزمية وكرر الخطوتين 1 و 2.
4. في حالة وجود مشكلة انحدار ، بالنسبة لسجل جديد ، تتنبأ كل شجرة في الغابة بقيمة Y (الإخراج). يمكن حساب القيمة النهائية بأخذ متوسط جميع القيم التي تنبأت بها جميع الأشجار في الغابة. أو ، في حالة وجود مشكلة تصنيف ، تتنبأ كل شجرة في الغابة بالفئة التي ينتمي إليها السجل الجديد

13. خوارزمية Decision Tree [14]

شجرة القرار هي تقنية تعليمية خاضعة للإشراف يمكن استخدامها لكل من مشاكل التصنيف والانحدار ، ولكنها في الغالب مفضلة لحل مشاكل التصنيف.

تمثل العقد الداخلية سمات مجموعة البيانات ، وتمثل الفروع قواعد القرار وتمثل كل عقدة ورقية النتيجة.

في شجرة القرار ، توجد عقدتان ، وهما عقدة القرار والعقدة الورقية. تُستخدم عقد القرار لاتخاذ أي قرار ولها فروع متعددة ، في حين أن العقد الورقية هي نتاج تلك القرارات ولا تحتوي على أي فروع أخرى.

يتم تنفيذ القرارات أو الاختبار على أساس ميزات مجموعة البيانات المحددة.

لماذا نستخدم أشجار القرار؟

فيما يلي سببان لاستخدام شجرة القرار:

- عادةً ما تحاكي أشجار القرار قدرة التفكير البشري أثناء اتخاذ القرار ، لذلك من السهل فهمها.
- يمكن فهم المنطق وراء شجرة القرار بسهولة لأنها تظهر بنية شبيهة بالشجرة.

مصطلحات شجرة القرار :

عقدة الجذر Root Node: عقدة الجذر هي من حيث تبدأ شجرة القرار. يمثل مجموعة البيانات بأكملها ، والتي يتم تقسيمها إلى مجموعتين أو أكثر من المجموعات المتجانسة.

العقدة الورقية Leaf Node: العقد الورقية هي عقدة الإخراج النهائية ، ولا يمكن فصل الشجرة أكثر بعد الحصول على عقدة ورقية.

التقسيم Splitting: التقسيم هو عملية تقسيم عقدة القرار / العقدة الجذرية إلى عقد فرعية وفقًا للشروط المحددة.

الفرع / الشجرة الفرعية Branch/Sub Tree: الشجرة التي تكونت بتقسيم الشجرة.

التقليم Pruning: التقليم هو عملية إزالة الفروع غير المرغوب فيها من الشجرة.

العقدة الأصلية / الفرعية Parent/Child node: تسمى العقدة الجذرية للشجرة العقدة الأصلية ، وتسمى العقد الأخرى بالعقد الفرعية.

كيف تعمل خوارزمية شجرة القرار؟

في شجرة القرار ، للتنبؤ بفئة مجموعة البيانات المحددة ، تبدأ الخوارزمية من العقدة الجذرية للشجرة. تقارن هذه الخوارزمية قيم سمة الجذر بسمة السجل (مجموعة البيانات الحقيقية) ، وبناءً على المقارنة ، تتبع الفرع وتنتقل إلى العقدة التالية.

بالنسبة للعقدة التالية ، تقارن الخوارزمية مرة أخرى قيمة السمة مع العقد الفرعية الأخرى وتتحرك أكثر. تستمر العملية حتى تصل إلى العقدة الورقية للشجرة. يمكن فهم العملية الكاملة بشكل أفضل باستخدام الخوارزمية أدناه:
الخطوة 1: ابدأ الشجرة بالعقدة الجذرية ، لنرمز لها بالرمز S، والتي تحتوي على مجموعة البيانات الكاملة.
الخطوة 2: ابحث عن أفضل سمة في مجموعة البيانات باستخدام مقياس تحديد السمة Attribute Selection Measures(ASM)
الخطوة 3: قسّم S إلى مجموعات فرعية تحتوي على القيم الممكنة لأفضل السمات.
الخطوة 4: إنشاء عقدة شجرة القرار ، والتي تحتوي على أفضل سمة.
الخطوة 5: إنشاء أشجار قرارات جديدة بشكل متكرر باستخدام مجموعات فرعية من مجموعة البيانات التي تم إنشاؤها في الخطوة 3. استمر في هذه العملية حتى يتم الوصول إلى مرحلة حيث لا يمكنك تصنيف العقد بشكل أكبر وتسمى العقدة النهائية كعقدة طرفية.

14. خوارزمية Naive Bayes

Naive Bayes هي تقنية تصنيف إحصائي تعتمد على نظرية بايز. إنها واحدة من أبسط خوارزميات التعلم الخاضع للإشراف. مصنف Naive Bayes هو خوارزمية سريعة ودقيقة وموثوقة. تتميز مصنفات Naive Bayes بدقة عالية وسرعة في مجموعات البيانات الكبيرة. [15]

يعمل مصنف Naive Bayes على الاحتمال الشرطي. الاحتمال الشرطي هو احتمال حدوث شيء ما ، بالنظر إلى أن شيئًا آخر

قد حدث بالفعل. باستخدام الاحتمال الشرطي ، يمكننا حساب احتمال حدث باستخدام معرفته السابقة. [18]

يوجد أدناه معادلة حساب الاحتمال الشرطي.

$$P(H \mid E) = \frac{P(E \mid H) * P(H)}{P(E)}$$

معادلة حساب الاحتمال الشرطي لخوارزمية Naive Bayes [18]

حيث أن :

P (H) هو احتمال أن تكون الفرضية H صحيحة. يُعرف هذا بالاحتمال السابق.

P (E) هو احتمال الدليل (بغض النظر عن الفرضية).

P (E | H) هو احتمال وجود الدليل بالنظر إلى صحة الفرضية.

P (H | E) هو احتمال الفرضية بالنظر إلى وجود الدليل.

15. مصفوفة الارتباك Confusion Matrix [19]

مصفوفة الارتباك هي مصفوفة N x N تُستخدم لتقييم أداء نموذج التصنيف ، حيث N هو عدد الفئات المستهدفة. تقارن المصفوفة القيم المستهدفة الفعلية بتلك التي تنبأ بها نموذج التعلم الآلي. يمنحنا هذا نظرة شاملة عن مدى جودة أداء نموذج التصنيف الخاص بنا وأنواع الأخطاء التي يرتكبها.
يتم حساب دقة النموذج بناء على المعادلات الآتية : [20]

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{\text{Correct Predictions}}{\text{Total Predictions}}$$

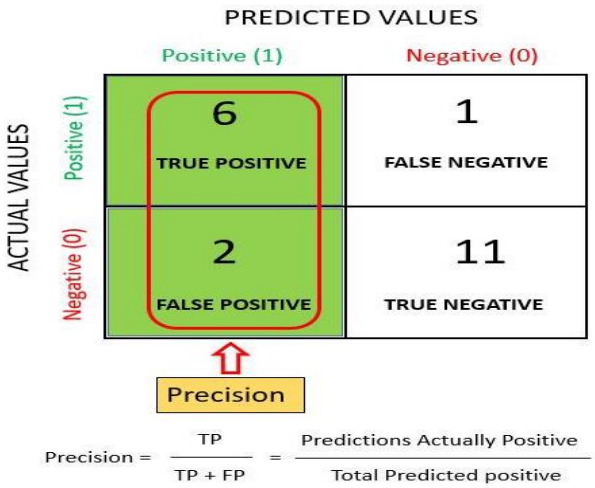
معادلة حساب دقة النموذج [20]

ما هو TP و FP و FN و TN هنا؟ هذا هو الجزء الأساسي من مصفوفة الارتباك [21]
إيجابي حقيقي (TP)
القيمة المتوقعة تطابق القيمة الفعلية
كانت القيمة الفعلية موجبة وتنبأ النموذج بقيمة موجبة
سلبي حقيقي (TN)
القيمة المتوقعة تطابق القيمة الفعلية
كانت القيمة الفعلية سالبة وتنبأ النموذج بقيمة سالبة
موجب كاذب (FP) - خطأ من النوع الأول

تم التنبؤ بالقيمة المتوقعة بشكل خاطئ كانت القيمة الفعلية سالبة لكن النموذج تنبأ بقيمة موجبة
سلبي كاذب (FN) - خطأ من النوع 2
تم التنبؤ بالقيمة المتوقعة بشكل خاطئ
كانت القيمة الفعلية موجبة لكن النموذج تنبأ بقيمة سالبة

تخبرنا precision عن عدد الحالات الايجابية التي تم التنبؤ بها بشكل صحيح . أي أنها النسبة بين الإيجابيات الحقيقية التي تم التنبؤ بها و بين كل الإيجابيات التي تم التنبؤ بها. [22] أي أنها تجاوب عن السؤال التالي : كم من عدد الأساتذة الذين تم تصنيفهم من قبل الطلاب على أن مستوى أداءهم جيد أو وسط أو منخفض هم بالفعل كذلك ؟ ، "الدقة هي مقياس مفيد في الحالات التي يكون فيها الإيجابي الكاذب مصدر قلق أكبر من السلبيات الكاذبة" . [23]

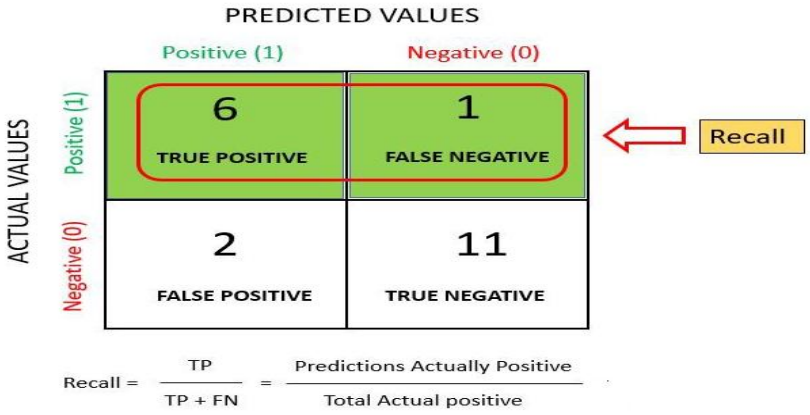
والشكل (1) يوضح طريقة حساب الـ Precision :



الشكل (1) : طريقة حساب الـ Precision [23]

تخبرنا Recall عن عدد الحالات الإيجابية الفعلية التي تمكن النموذج من التنبؤ بها بشكل صحيح. أي أنه هو مقياس النموذج الذي يحدد الإيجابيات الحقيقية بشكل صحيح ، و نلجأ اليه عندما يكون أهمية السلبية الكاذبة عالية أي أن يخبرنا النموذج أن أداء الأستاذ منخفض و هو ليس كذلك .

الشكل (2) يوضح طريقة حساب Recall :



الشكل (2) : طريقة حساب Recall [23]

F1 score :

F1-Score هي الوسيلة التوافقية للدقة والاستدعاء ، ولذا فهي تعطي فكرة مجمعة عن هذين المقياسين. وتعطي نفس الوزن لكل من الـ Recall و precision .

$$F1\ score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2 * (Precision * Recall)}{(Precision + Recall)}$$

معادلة حساب f1 score [23]

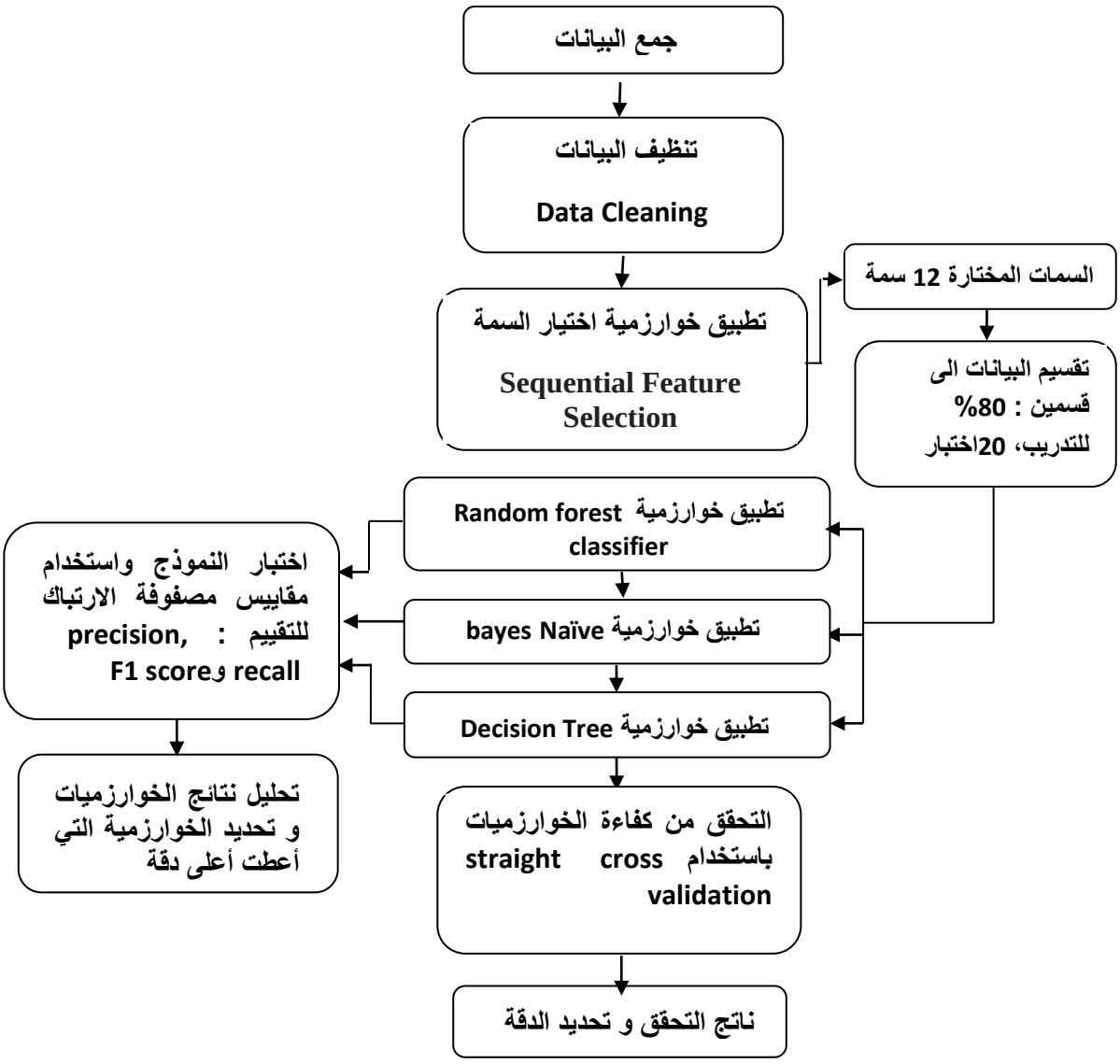
Macro average
هو متوسط precision / recall / f1 score

$$\text{Macro avg Precision} = \frac{P1 + P2}{2}$$

معادلة حساب متوسط precision / recall / f1 score [23]

16. النموذج :

كما هو ملاحظ في الشكل (3) فان خطوات النموذج موضحة بداية بجمع البيانات وانتهاء بنتيجة التقييم



الشكل (3) : خطوات بناء النموذج

1-16 . مرحلة جمع البيانات:

تضمنت هذه المرحلة جمع البيانات بواسطة الاستبيان وبعض البيانات الأخرى الخاصة بالاستاذ مثل المؤهل العلمي، سنوات الخبرة و المعدل التراكمي.

2-16 . مرحلة تنظيف البيانات : Data Cleaning

بعد أن قام الباحث بدراسة مجموعة البيانات و اجراء بعض المقارنات الاحصائية بين السمات قرر استبعاد السمات : الجامعة ، الكلية ، القسم ، الرقم الجامعي ، اسم الطالب ، اسم المقرر ، اسم أستاذ المقرر و المعدل التراكمي وذلك للاسباب التالية :

1. الجامعة، الكلية، القسم واسم المقرر واسم الاستاذ سمات غير مؤثرة في النتائج.
2. المعدل التراكمي انحصرت أغلب القيم بين 72 و 74 أي أن المدى بسيط نوعا ما .
3. الرقم الجامعي واسم الطالب تحتوي على قيم شاغرة كثيرة.

3-16 . تحديد السمة المتسلسل Sequential Feature Selection [13]:

تبدأ هذه الخوارزمية بمجموعة فارغة ثم تضيف سمة واحدة إلى الخطوة الأولى التي تصور أعلى قيمة لوظيفة الهدف ومن الخطوة الثانية فصاعدًا ، تتم إضافة السمات المتبقية بشكل فردي إلى المجموعة الفرعية الحالية وبالتالي يتم تقييم المجموعة الفرعية الجديدة . تتكرر هذه العملية حتى يتم إضافة العدد المطلوب من السمات .

رأى الباحث أن عدد السمات كثير نوعا ما وعليه قرر تطبيق طريقة اختيار السمات feature selection لتقليل أبعاد البيانات أثناء إجراء التحليل التنبؤي وزيادة كفاءة النموذج .

تم اختيار طريقة التغليف Wrapper Method وطبقت خوارزمية تحديد السمة المتسلسل Sequential Feature Selection(SFS) و قامت الخوارزمية بتقليل السمات من 29 إلى 12 سمة باعتبارها السمات الأكثر تنبؤا والتي تعطي معدل دقة جيد ، و كانت السمات المختارة هي :

Batch, show_study_plan ,clear_presentation , Prefer_technology ,Lecturer_respect_std , IS_lecturer_good_model ,Control_class ,recommend_references ,Want_same_lect ,Std_mark ,Lect_drgree ,Lect_experience

أي أن السمات على التوالي : فرقة الطالب، يعرض الأستاذ الخطة الدراسية للمقرر في بداية الفصل ، يقدم الأستاذ شرحا واضحا ومفهوما للمادة العلمية أثناء المحاضرات، تفضل استخدام الوسائل التدريسية أكثر من الطريقة التقليدية للشرح، يتعامل الأستاذ مع الطلاب بتقدير واحترام، يلتزم الأستاذ بالسلوك الحسن قولا وعملا (القدوة الحسنة)، يضبط الأستاذ النظام في المحاضرة، ينبه الأستاذ طلابه إلى الاستعانة بمصادر خارجية (المكتبة / الانترنت)، ترغب في تلقي دروس أخرى بواسطة ذات الأستاذ، التقدير الذي حصلت عليه في الامتحان، الدرجة العلمية وعدد سنوات الخبرة هي السمات التي تحسن أداء النموذج وتقلل من الوقت الحسابي وكذلك المساحة المطلوبة لتشغيل الخوارزمية.

4-16 . طريقة تقسيم البيانات :

حجم البيانات هو 12 سمة و 709 صف ، تم تقسيمها الي قسمين : للتدريب و الاختبار بنسبة 80% و 20% على التوالي فاصبح الجزء المخصص للتدريب هو عبارة عن 567 صف و الباقي للاختبار و هو 142 صف .

تم اختيار السمة Std_satisfaction والتي تعبر عن رضاء الطالب العام عن أداء الاستاذ، تم اختيارها كمتغير هدف في عملية التصنيف، عدد الفصائل/خيارات هذه السمة هي 0،1،2 حيث أن 0 تعني أن درجة الرضاء منخفضة و 1 تعني أنها جيدة و 2 تعني أنها متوسطة .

الشكل (4) يوضح توزيع البيانات على الفصائل و كما نلاحظ أن انحياز أغلب الاجابات الى الفصيل (1) و ذلك بنسبة تفوق الـ 400 من أصل 709 ، يليها الاداء الوسط (الفصيل (2)) ثم اخيرا الاداء الضعيف (الفصيل (0)) . هذا الانحياز يجعل البيانات غير متوازنة imbalanced مما سيؤثر لاحقا في نتائج النموذج . لذلك ولمزيد من التأكد من أداء النموذج و لمقارنة وتحليل دقة خوارزميات التصنيف الثلاث، استخدم في مرحلة التحقق Validation تقنية Stratified K-fold Cross-Validation وذلك لأنها تتحكم في نسبة توزيع البيانات في الفصائل بنسب متساوية تقريبا .

```
y3.value_counts()
```

```
1    411
2    200
0     98
Name: Std_satisfaction, dtype: int64
```

الشكل (4) : توزيع البيانات على الفصائل

5-16 . تطبيق خوارزميات التصنيف الثلاث :

1-5-16. خوارزمية الغابة العشوائية Random Forest Classifier:

تم استخدام خوارزمية الغابة العشوائية لتدريب النموذج على بيانات التدريب و كانت دقة التعلم تساوي 0.97 .

و عند اختبار النموذج و من ثم استخدام مصفوفة الارتباك لتقييم النموذج كانت الدقة تساوي 0.77 ، كما موضح في الشكل (5)

```
print (metrics.classification_report(y_test,y_pred_rd))
```

	precision	recall	f1-score	support
0	0.69	0.65	0.67	17
1	0.87	0.88	0.87	89
2	0.56	0.56	0.56	36
accuracy			0.77	142
macro avg	0.70	0.69	0.70	142
weighted avg	0.77	0.77	0.77	142

الشكل (5) : تقييم خوارزمية Random Forest بعد اختيار السمات

عند مرحلة التحقق واستخدام Stratified K-fold Cross-Validation كان متوسط الدقة للنموذج هو % 77.17

16-5-2.خوارزمية Decision Tree:

تعلم هذه النموذج بدقة تساوي 0.1 و عند اختبار النموذج و من ثم استخدام مصفوفة الارتباك لتقييم النموذج كانت الدقة تساوي 0.73 ، كما موضح في الشكل (6)

```
print (metrics.classification_report(y_test,y_pred))
```

	precision	recall	f1-score	support
0	0.45	0.59	0.51	17
1	0.89	0.80	0.84	89
2	0.55	0.61	0.58	36
accuracy			0.73	142
macro avg	0.63	0.67	0.64	142
weighted avg	0.75	0.73	0.73	142

الشكل (6) : تقييم خوارزمية Decision Tree بعد اختيار السمات

عند مرحلة التحقق واستخدام Stratified K-fold Cross-Validation كان متوسط الدقة للنموذج هو 69.13%

16-5-3.خوارزمية Naive Bayes :

تعلم هذه النموذج بدقة تساوي 0.78 و عند اختبار النموذج و من ثم استخدام مصفوفة الارتباك لتقييم النموذج كانت الدقة تساوي 0.80 ، كما موضح في الشكل (7)

```
print (metrics.classification_report(y_test,y_pred))
```

	precision	recall	f1-score	support
0	0.65	0.88	0.75	17
1	0.89	0.89	0.89	89
2	0.70	0.58	0.64	36
accuracy			0.81	142
macro avg	0.75	0.78	0.76	142
weighted avg	0.81	0.81	0.81	142

الشكل (7) : تقييم خوارزمية Naive Bayes بعد اختيار السمات

عند مرحلة التحقق واستخدام Stratified K-fold Cross-Validation كان متوسط الدقة للنموذج هو% 77.87

17.تقييم النتائج :

كانت نتائج الخوارزميات الثلاث كما موضح في الجدول (1):

الجدول (1) : ملخص نتائج الخوارزميات

Algorithm	classes	Micro avg/ accuracy	precision	recall	F1-score	Accuracy after validation
Random Forest	0	0.77	0.69	0.65	0.67	77 %
	1		0.87	0.88	0.87	
	2		0.56	0.56	0.56	
		Macro avg	0.70	0.69	0.70	
Decision Tree	0	0.73	0.45	0.59	0.51	69 %
	1		0.89	0.80	0.84	
	2		0.55	0.61	0.58	
		Macro avg	0.63	0.67	0.64	
Naïve Bayes	0	0.81	0.67	0.82	0.74	79 %
	1		0.89	0.89	0.89	
	2		0.66	0.58	0.62	
		Macro avg	0.74	0.76	0.75	

نلاحظ من الجدول (1) أن خوارزمية Naïve Bayes أعطت معدل دقة أعلى من خوارزميتي الغابة العشوائية وشجرة القرار

النتائج و التوصيات :

بعد أن تم تصميم النموذج قيد الدراسة و ذلك باستخدام ثلاث من خوارزميات التصنيف وهي : Random Forest , Decision Tree and Naïve Bayes تم إجراء مقارنة دقة جميع الخوارزميات خلال عملية التنبؤ بأكثر السمات تأثيرا على التنبؤ بأداء الاستاذ. ووجد أن استخدام خوارزمية Naïve Bayes اعطت دقة عالية وهي 0.81 و عند استخدام طريقة Stratified K-fold Cross-Validation كانت الدقة تساوي % 79 .

تعتبر دقة الخوارزميات جيدة نوعا ما مقارنة بحجم العينة و عدم توازن بياناتها، ولا يوجد اختلاف كبير في الدقة بين مرحلة التنبؤ و مرحلة التحقق وهذا يشير الى أن السمات المختارة بواسطة خوارزمية SFS كانت جيدة في عملية التنبؤ وعليه يمكن أن يعول على هذه السمات في اعتمادها كإطار عام وقابل للتطبيق و الاستخدام في جامعات أخرى وبيانات مختلفة. لكن من ناحية أخرى نلاحظ أن النموذج يعطي نتائج منخفضة في recall و precision مع الفصائل 0 و 2 بينما النتائج مرتفعة مع الفصيل 1 وذلك بسبب انحياز النموذج لفصيل الأغلبية .

يوصي الباحث بالآتي :

1. المتغير الهدف و الذي يتمثل في الرضاء العام عن أداء الاستاذ يفضل أن يحمل قيمتين فقط اما جيد أو منخفض بدلا عن ثلاثة قيم أي الغاء القيمة "متوسط" لأنه من تحليل نتائج الخوارزميات وجد الباحث أنه و بالرغم من أن عدد البيانات في الفصيل 2 أكثر من الفصيل 0 الا أن دقة مصفوفة الارتباك تعطي نتائج سيئة وهذا يعني أن نسبة التنبؤ في هذا الفصيل ضعيفة. و

كتحليل فلسفي توقع الباحث هذه النتيجة لأن أغلب الطلاب لا يميلون لملء مثل هذه الاستبيانات لاعتبارات عدة واذا كان هنالك خيار يحمل القيمة "وسط" أو "الى حد ما" سيقوم الطالب باختياره بدلا عن الخيار "منخفض".

2. اذا كان حجم العينة كبير و وجدت بيانات متراكمة لعدد من السنين يمكن تقييم أداء كل أستاذ على حدى و مطابقة النتائج مع نتائج الدراسة الحالية .
3. عمل توليد للبيانات في فصيل الأقلية Minority classes وذلك لموازنة البيانات ومن ثم تحسين أداء الخوارزميات . يقترح الباحث استخدام Synthetic Minority Oversampling Technique (SMOTE)
4. يوصى الباحث باستخدام مجموعة بيانات أخرى من جامعة أخرى و اختبارها على النموذج قيد الدراسة ومقارنة النتائج للتأكد من دقة النموذج قيد الدراسة .

المراجع :

[1] كمال بـــورزق . عبد الحميد شـــطة . محمد رحماني (2018) ، " نظرة الطلبة نحو أداء أستاذ البدنية والرياضية بين الواقع والمأمول" ، دراسة ميدانية في معهد علوم وتقنيات الأنشطة البدنية والرياضية ، https://www.asjp.cerist.dz/en/article/64575

[2] Debalkew Geremew & Rupali Gangrade, " Data Mining Techniques to Evaluate and Predict Teachers's Performance in Higher Education", International Journal of Innovations & Avancement in computer scince , IJIACS , (october 2017)

[3] Ahmed Mohamed Ahmeda. AhmetRizanerc. Ali HakanUlusoyc, " Using data mining to predict instructor performance" , International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS, (29-30 August 2016)

[4] Sona MARD˙IKYAN, Bertan BADUR, "Analyzing Teaching Performance of Instructors Using Data Mining Techniques" ,Department of Management Information Systems, Bogazici University ,Istanbul, Turkey, (July 2011)

[5] Asanbe M.O. Osofisan A.O. William W.F, "Teachers' Performance Evaluation in Higher Educational Institution using Data Mining Technique" , University of Ibadan Ibadan, Nigeria , International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868 Foundation of Computer Science FCS, New York, USA Volume 10 – No.7 , (March 2016) , www.ijais.org

[6] Sona MARD˙IKYAN, Bertan BADUR , " Analyzing Teaching Performance of Instructors Using Data Mining Techniques" , Department of Management Information Systems, Bogazici University Istanbul, Turkey , (July 2011)

[7] Necla G¨und¨uza_ and Ernest Fokou´eb , " Pattern Discovery in Students' Evaluations of Professors A Statistical Data Mining Approach" : https://arxiv.org/abs/1501.02263 , 11/2/2020 . 3:00 PM , (JAN 2015)

[8] MUSTAFA AGAOGLU , " Predicting Instructor Performance Using Data Mining Techniques in Higher Education" , IEEE, Received May 6, 2016, accepted May 10, 2016, date of publication May 13, 2016, date of current version June 3, 2016. Digital Object Identifier 10.1109/ACCESS.2016.2568756 (June 3, 2016)

[9] "Why is Educational Data Mining important in the research?" [online],

Available: https://towardsdatascience.com/why-is-educational-data-mining-important-in-the-research-e78ed1a17908 , Last Updated : 19 Jan, 2019 , Accessed: 18/3/2021 . 3:50 PM

[10] "الجودة الشاملة في التعليم"[online], Available: https://mawdoo3.com/ الجودة الشاملة في التعليم , Last Updated : 31 Jul, 2016 , Accessed:1/2/2021 . 12:26 AM

[11] Costing, H. (1994). “Reading in Total Quality Management”, Copyright, by Har Court Brace & Company, Sandigo, New York , Global Journal of Management and Business Research: A Administration and Management , (2019)

[12] “What Are Feature Selection Techniques In Machine Learning?” [online] , Available : https://analyticsindiamag.com/what-are-feature-selection-techniques-in-machine-learning/ , Last Updated : 1 Apr, 2019, Accessed: 2021/3/5 . 12:35AM

[13] “Random Forest Algorithm with Python and Scikit-Learn” [online] , Available: https://stackabuse.com/random-forest-algorithm-with-python-and-scikit-learn/, Last Updated : 21 Mar, 2021 , Accessed: 7/3/2021. 11:53 PM

[14] “Decision Tree Classification Algorithm” [online] ,

Available: https://www.javatpoint.com/machine-larning-decision-tree-classification-algorithm , Last Updated : 2018 , Accessed: 18/3/2021 . 3:45 PM

[15] “Naive Bayes Classification using Scikit-learn” [online], Available:https://www.datacamp.com/community/tutorials/naive-bayes-scikit-learn , Last Updated : 4 Des, 2018 , Accessed: 18/3/2021. 4:45PM

[16] وزارة التعليم العالي و البحث العلمي ، الهيئة العليا للتقويم و الاعتماد ، "دليل المعايير الوطنية لضمان جودة التعليم العالي في السودان" ، (2012)

[17] Harshita Patel. Dharmendra Singh Rajput. G Thippa Reddy1. Celestine Iwendi. Ali Kashif Bashir. Ohyun Jo ,” A review on classification of imbalanced data for wireless sensor networks” ,International Journal of Distributed Sensor Networks , Received: January 06, 2020; Accepted: March 04, 2020 .

[18] https://dataaspirant.com/naive-bayes-classifier-machine-learning/, “HOW THE NAIVE BAYES CLASSIFIER WORKS IN MACHINE LEARNING” , February 6, 2017 , accessed : 28/4/2021 1:26 PM

[19] “Everything you Should Know about Confusion Matrix for Machine Learning” [online],

Available: https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/, Last Updated : 17 Apr 2020 , Accessed: 4/3/2021 . 11:32 PM

[20] “Classification: Accuracy” [online] ,

Available: https://developers.google.com/machine-learning/crash-course/classification/accuracy , Accessed: 10/4/2021 . 2:10 AM

[21] “ Understanding True Positive, True Negative, False Positive and False Negative in a Confusion Matrix “ [online] ,

Available: https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machin e-learning/ , Accessed: 10/4/2021 . 2:10 AM

[22] “Precision vs. Recall – An Intuitive Guide for Every Machine Learning Person” [online], Available : https://www.analyticsvidhya.com/blog/2020/09/precision-recall-machine-learning/ , Accessed: 10/4/2021 . 2:23 AM

[23] “ What is a confusion matrix?” [online], Available : https://medium.com/analytics-vidhya/what-is-a-confusion-matrix-d1c0f8feda5, Accessed: 10/4/2021 . 3:04 AM