Journal, Advances in Mathematical & Computational Sciences
Vol. 8 No. 1, March 2020
www.mathematics-computationaljournal.info

Article Citation Format
Ukoba, J.O, Ochei, L.C. & Okengwu, U.A. (2020): Detecting Students' Assessment Score Anomalies Using KNN Classification Algorithm. Journal of Advances in Mathematical & Computational Sc. Vol. 8, No. 1. Pp 55-66

Article Progression Time Stamps

Article Type: Research Article
Manuscript Received 10th February, 2019
Final Acceptance: 17th March, 2020
Article DOI: dx.doi.org/10.22624/aims/mathscompv8n1p5

Detecting Students' Assessment Score Anomalies Using KNN Classification Algorithm

[1]Ukoba, J.O, [2]Ochei, L.C. & [3]Okengwu, U.A.
Department of Computer Science
University of Port Harcourt
Port Harcourt, Nigeria
***E-mail**: ukobajosephorode@gmail.com

ABSTRACT

The challenge of detecting and classifying anomalies in student's assessment scores has been a key issue in analyzing student results, especially in institutions of higher learning. There is little research that applies machine learning to detect and classify anomalies in student assessment scores. In this study, we applied the K-Nearest Neighbour algorithm to improve the detection and classification of anomalies in student's assessment scores. Two categories of students' assessment dataset were considered: (a) CGPA-based dataset, which represents an anomaly where there were too many high CGPAs, and (b) course-based dataset which represents anomalies related to - (i) inconsistent CA vs Exam score, (ii) too many high scores, and (iii) borderline failure. Experimental result shows at least 95% accuracy in detection rate, precision, F1 score, and recall (sensitivity), especially in the disproportionate CA vs Exam Score anomaly and borderline anomaly.

Keywords: Anomaly Detection, Classification, Student Assessment dataset, K-Nearest Neighbour (K-NN), Algorithms, Normal distribution

1. INTRODUCTION

Assessing the performance of people, whether staff, students or a company's inputs and outputs are a crucial and important activity that takes place in every sector or institution (Ukoba *et al*, 2020). The performance of staff in companies, industries, and other institutions are been examined either quarterly or annually so as to give the necessary feedback to the management for effective decision making. This is done by the superintendents or supervisors examining each staffs' performance, using some criteria for evaluation. This appraisal is then sent to the management for an effective business decision which may include promotions or queries. Assessing the companies' input resources and corresponding output are a necessity in evaluating the company's performance and take necessary decision.

In education, assessments are by administering continuous assessment and examinations. An examination is considered as the peak of every academic process (Salami et al, 2016). Students would have to take a series of internal and sometimes external examinations before they could advance to another level/stage. Assessing students' performance is key, as it evaluates the students' rate of assimilation and gives feedback for improving students learning. After teaching, examinations are set for the students. A student can only progress to the next level when the pass mark obtained is above the cut off mark in an examination. This cut of mark is agreed upon by the management based on the general performance.

Once examinations are administered and written, it is the duty of course lecturers to allocate marks to the students based on answers provided. The total score that determines the grade of the student is computed by adding the CA score to the examination score. This is what determines the student's general performance in any course. This, in turn, is used to compute the CGPA of the students which determines the class the student graduates with (Salami et al, 2016; Ukoba et al, 2020). An accurate result computation, compilation, and approval bring the semester or term to end. Evaluating and approving students' results manually is faced with some setbacks such as fatigue, waste of energy and time and the inability of the board/senate members to detect certain anomalies because of the large information they are to assimilate (Salami et al, 2016).

Analyzing these anomalies is of great interest to many researchers in diverse fields especially in data-science (Xiaodan Xu et al, 2019). Behaviors or patterns in students' result which deviate from what is normal are termed 'anomalies'. Anomalies are seen as deviations, outliers or peculiarities caused by some malevolent activities or intrusions (Chandola et al, 2009; Agrawal, 2015). Results of a student or set of students that go against the rules of an existing standard or regular model or rule are of great interest to academicians with regards to its detection (Dokas et al, 2002). Anomalies are surprising or unforeseen patterns that are notably different from others in the given dataset, and this is what the study of anomaly is aimed at finding or identifying (Xiaodan Xu et al, 2019). The problem of discovering these patterns which do not obey the existing standard or pattern defines anomaly detection (Chandola et al, 2009). It is identifying these patterns that go against that which is normal or standard (Shahreza et al 2011).

Anomaly detection is applied in diverse fields (Shahreza et al 2011). In the medical sciences, it is applied in the health monitoring systems for spotting malignant tumours in an MRI scan and for intrusion detection in weird patterns which could signal that a traffic network has been hacked. In the banking institution, anomaly detection discovers credit card frauds in transactions and in operating environments to detect faults in the sensor of spacecraft (Pramit; 2017; Chandola et al, 2009; Kumar, 2005). Anomalies can be identified using a data mining approach or machine learning approach. In data mining, patterns that are strange are extracted in vast and large database evaluating what is discovered on the unknown strange patterns from those known previously (Raval, 2012; Buczak and Guven, 2016).

Machine learning allows the computer to learn by extracting, without programming explicitly, information automatically using computational methods (Buczak and Guven, 2016; Svensson and Söderberg, 2008). Parmar and Patel (2017) classified machine learning approaches into three: learning which involves or requires identifying beforehand normal instances and anomalous once called supervised learning. Here the dataset must be first trained (Shahreza et al 2011); learning by constructing a model to mimic the standard or normal pattern from an existing normal dataset already trained, with the possibility that the learned model will generate testing instances, called the semi-supervised learning (Pamar and Patel, 2017) and the unsupervised learning where the dataset is unlabelled, thus dataset training is not required (Pamar and Patel, 2017). This paper seeks to improve the detection of anomalies in the student assessment dataset using the K-Nearest Neighbour algorithm.

The rest of this paper is organized as follows: Section II gives a review of related work on students' assessment dataset and the use of K-NN classification algorithm in anomaly detection. An x-ray of anomaly detection in students' assessment dataset was carried out in Section III. The KNN classification algorithm was examined in Section IV. The experimental results and the discussion of the results are presented in Section V and VI, respectively. Section VII concludes the paper with future work.

2. REVIEW OF RELATED WORK

Many research work have been carried out on anomalies and their detection in several fields like identification of malevolent webpage, intrusion detection, detection of faults in space aircraft, etc. In this section, we review related work on the use of the KNN classification algorithm and that of detecting anomalies in students' assessment dataset.

A. Review of Related Work on K-Nearest Neighbour Algorithm

Chandola et al (2009) did a survey on anomaly detection. The survey helped to provide a structural and broad overview of anomaly detection. They discussed anomalies in six categories, viz., classification-based, clustering-based, Nearest Neighbour based, information-theoretic and spectral techniques. Viswanath et al (2014) used a framework of PCA and the KNN to distinguish potentially bad redundant behavior in social networks from normal ones. The PCA was to detect high-dimensional data patterns. The KNN then identifies the class of anomaly. The framework gave 66% accuracy with a false positive of less than 0.3%. Djenouri et al (2019) considered the traffic flow distributions in detecting surprising patterns or distributions in Spatio-temporal traffic flow in specified time intervals using the Flow Distribution Probability (FDP) and the K Nearest Neighbour algorithm. Testing results with a real dataset from Odense traffic flow at ten locations reveals efficient real distribution flow outliers by the proposed framework. Testing the model on Beijing data showed a performance above the baseline algorithm for high urban traffic flow.

Liao and Vemuri (2002) investigated the K-NN classifier performance in detecting program intrusion and classifying system call frequencies using text categorization techniques. The text categorization techniques were adopted to convert each process to a vector, evaluating similarities between two program activities. K-NN then detects and classifies normal and intrusive attacks. The K-NN classifier with a text categorization technique detected and classified intrusion attacks effectively, achieving a low false positive rate. However, the model was unable to detect anomalies in the frequency system calls if the incursion doesn't reveal any attack in operation in as much as the process does not exist.

Oladeji and Adeleke (2017) proposed an improved K-NN classifier for detecting and classifying intrusions. The K-mean clustering algorithm was for clustering verification and the generic algorithm for optimization. The improved K-NN classifier had 99.6% efficiency for instances classified correctly and 0.3222% for incorrectly classified. Chaurasia and Anurag (2013) proposed a combining classification of the Neural Network model and K-NN method for intrusion detection. Implementation was done in 2 phases. In phase 1, Neural Network was used for better results and improves the K-NN classifier and then ensemble (bagging) them and used in phase 2 for classification.

B. Review of Related Work on Students' Assessment Dataset

Cortez and Silva (2008), examined the performance of decision trees, Random forest, Neural Network and Support Vector Machines on secondary school grades using binary classification, regression and five-level classification for evaluation. They explained that social, demographic and school related variables also affect students' performance. Salami et al (2016) carried out a study on the detection anomalies in students' results using the decision trees. The decision tree model was able to detect efficiently anomalies in student results in most cases. However, it was unable to detect or identify anomalies in a situation where the training dataset has few anomalous instances. Thomas and Jayagopi (2017) measured the students' engagement using a machine learning algorithm based on students' facial expressions, head poses, and eye gazes. The experimental result showed that the machine learning algorithm performed well in predicting student engagement in class. Hamid et al (2018) measured students' engagement using a machine learning approach and concluded that the SVM and K-NN classifiers are appropriate for predicting students' engagement. Ukoba et al (2020) carried out a review on the detection and classification of anomalies in students' assessment dataset. The review pointed out that very few works have been done with regard to detecting anomalies in the educational sector/domain. Salami and Yahaya (2018) described how the Extreme Learning Machines (ELM) can be used to automatically detect anomalies in students' result. However, it was unable to detect anomalous instances in some of the dataset especially where the anomalous instances were few. This paper seeks to use the K-Nearest Neighbor to improve the detection of anomalies in students' assessment dataset which follows the normal distribution curve.

3. ANOMALIES IN STUDENTS' ASSESSMENT DATASET

Student result anomalies are salient or unusual observations that need further elucidation. A typical result of students in a department or a particular course ought to follow the normal (Gaussian) distribution curve, where few students should have A's and F's and bulk of the students score C as shown in figure 1 below (Ukoba et al, 2020). Any deviation from this is considered an anomaly. The anomaly on its own is not necessarily a bad thing. For instance, when we have more first-class and second class upper than second class lower in a graduating set isn't actually bad. However, this result or grade would cause a deviation from the normal distribution curve., thus an anomaly. This paper seeks to detect anomalies in course-based and CGPA based assessment data

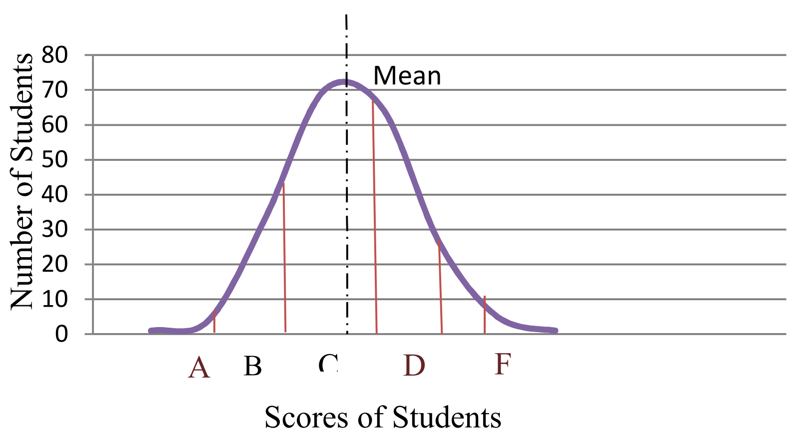

Figure 1: Normal Distribution for Scores/Grade of Students (Source: Ukoba et al, 2020)

4. K-Nearest Neighbor Classification Algorithm

K-nearest neighbour (K-NN) is a supervised, classification machine learning algorithm. It performs quite complex classification tasks but simple to implemented (Scott, 2018). KNN is a modest and conventional non-parametric technique for classifying samples. It doesn't assume anything about the underlying data. This is exceptionally one feature the K-NN has because most data in the real world doesn't actually follow any theoretical assumptions like the uniform distribution (Manocha and Girolami, 2007; Scott, 2018). The K-NN algorithm works by evaluating the estimated distances between various points on data inputted, thereafter assigns the unlabeled point a class of its K-nearest neighbours. The distance could be the Euclidean distance or the Manhattan distance etc. It thereafter selects the K-nearest data points where K could be an integer. It finally assigns to the class that the K data points belong to that data point.

The "k" is a useful parameter and various (k) values can cause various performances. If 'k' is big, the neighbours used for classification would affect the accuracy rate and prediction that can consume a lot of classification time. The K Nearest Neighbour algorithm doesn't have a specialized training phase; hence the KNN is termed 'lazy algorithm'. The KNN algorithm trains all the data as it classifies new data instances during testing (Scott, 2018).

Strengths

i. They are very intuitive and independent of the distribution of the data and capable of detecting isolated objects.
ii. One of the best classification algorithms with regards intuitive approach with high predictive power
iii. Very robust to noisy training data and outliers on the predictors.
iv. It is extremely easy to implement
v. Requires no training before making predictions. This makes K-NN much faster than other algorithms that require training like the SVM, linear regression, Decision trees, etc.
vi. The K-NN algorithm allows for seamless addition of new data, since it requires no training before making predictions.
vii. They are very simple to interpret, understand and are very powerful

Weaknesses

i. The K-NN algorithm does not work well with high dimensional data because with large number of dimensions, it becomes difficult for the algorithm to calculate distances in each dimension.
ii. Has a high prediction cost for large datasets. This is because in large datasets the cost of calculating the distance between new point and each existing point becomes higher.
iii. It does not work well with categorical features.

5. EXPERIMENT

This paper is aimed at improving the detection of course-based and CGPA based anomalies using the K- Nearest Neighbor algorithm and was implemented using Python programming language (Spyder [Anaconda] IDE).

A. Dataset

The dataset used in this paper was automatically generated randomly to follow the Guassian distribution (normal distribution), using Microsoft excel. This was done suing the the NORMINV function (i.e NORMINV(Rand(),mean, standard deviation). The procedure for generating the dataset is summarized below:

Procedure for Generating the Dataset:

Step 1: Start
Step 2: Set the mean and SD you which to consider
Step 3: Use the Norminv () function that returns random numbers which follows the normal distribution around the mean and SD by tying Normiv(Rand(), mean's Cell address, Cell address of the SD) and press enter.
Step 4: Use the formula Normiv(Rand(),$Mean column letter$Mean row number, $SDcolumn letter$SD row number) and drag to produce random numbers.
Step 5: To make the values to be static and not change copy and paste in a different cell.
Step 6: End

The dataset was classified broadly into two categories: the course-based assessment dataset and the CGPA based assessment dataset. In each categories of dataset acquired and described (course-based assessment scores and CGPA based assessment score), there are sets of scores generated randomly which follows the normal distribution and are labelled *baseline dataset*. These sets of scores follows the Gaussian distribution, hence assumed not to contain neither error nor anomalies.

In each category, there are sub-categories of datasets to which anomalies were inserted with respect to the type of anomaly considered in this research.

1. ***Course-based assessment dataset*:** In this dataset, each line represents scores of different students that took a course. Each row represents the score of different students in the course. The first column represents the candidate's numbers, the second column represents CA score of students and third column represents the examination scores of the students offering that course. We presume that one hundred and sixty (160) students offered the course of which each student partook in the test and examination. This paper considers three kinds of course-based anomalies. These are: Inconsistent CA vs Exam score anomalies, too many good high score anomalies, borderline failure anomalies.

2. ***CGPA-based assessment dataset:*** In this dataset, each line represents the CGPA of different students of department. Each row represents the CGPA of different students in a department. The first column represents the candidate's numbers; the second column represents the CGPA of each student of department. This paper considers one kind of CGPA based anomalies: "Too many high CGPA anomalies". We presume that a department had one hundred and fifty (150) students that have completed their program, passing all their courses, and thus CGPA computed.

B. Evaluation Metrics'

In validating detection and classification performance, several evaluation metrics' have been used. Some of which are F-Ratio, accuracy, sensitivity/recall, specificity, rank power etc. However, in this work, we use accuracy, sensitivity/recall, specificity and F-Ratio to validate the K-Nearest Neighbour algorithm performance.

Table 1 presents the relationship between the actual class and the predicted class with regards the four metric evaluation parameters.

i. **Accuracy**: This metric is for evaluating observations correctly classified. This is the ratio of observations predicted correctly (True positive) to the total observations.

$$\text{Accuracy} = \frac{TP+}{TP+FP+TN+F} \quad \text{x } 100\% \qquad (1)$$

In a class imbalanced dataset, accuracy alone doesn't tell if a model or system is doing excellently well

ii. **Sensitivity/Recall:** This measure evaluates observations which are positive, correctly classified. It shows the ratio of positive targets to all targets in the main class. It gives the extent to which instances that are anomalous are correctly identified. Equation 3.2 gives the computation of sensitivity/recall.

$$\text{Sensitivity}= \frac{TP}{TP+FN} \qquad (2)$$

iii. **Specificity:** This measure evaluates observations which are negative, correctly classified. It shows the ratio of negative targets to all targets in the main class. It gives the extent to which instances that are anomalous are correctly identified. The higher the percentage, the better. Equation 3.2 gives the computation of specificity.

iv.

$$\text{Specificity} = \frac{TN}{TN+FP} \qquad (3)$$

v. **F1 score:** This is the weighted average of precision and recall. Therefore, the score takes both false positives and false negatives into account. F1 score is usually more useful than accuracy especially if we have an imbalance class distribution.

$$\text{F1 Score} = \frac{\mathbf{2^*(Recall^*Precision)}}{\mathbf{Recall+Precision}} \qquad (4)$$

vi. **Precision:** Precision is the percentage of positive accurately predicted observations to all positive predicted observations.

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (5)$$

Where TP = Number of true positive targets (anomalous instances) correctly classified.
FP = Number of positive targets [anomalous instances] wrongly classified.
TN = Number of true negative targets (normal instances) correctly classified.

Table 1: Relationship between the Actual and Predicted Class with regards the Metric Parameters

		Predicted Class	
		Class = Positive	Class = Negative
Actual Class	Class = Positive	TP	FN
	Class = Negative	FP	TN

C.

C. Experimental Results

Table 2 shows the K-Nearest Neighbour system for the Inconsistent CA vs Exam score, too many high score and the borderline failure anomalies. Out of the 160 instances in the CA vs Exam score anomaly dataset, the dataset were randomly splitted into 120 instances (75%) for training and 40 instances (25%) for testing. Table 3 shows the K-Nearest Neighbour system for the too many high CGPA anomalies. Out of the 150 instances, the dataset were randomly spitted using the split-train-test into 112 for training instances (75%) and 38 instances (25%) for testing.

Table 2: Evaluation of K- NN Models for Course Based Anomalies

K-NN Anomaly		**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 Score (%)**	**Specificity (%)**
Inconsistent CA vs Exam Scores	Training	99	96	100	98	100
	Testing	100	100	100	100	100
Too many high Score	Training	93	86	86	86	86
	Testing	90	80	92	86	92
Borderline Failure	Training	99	75	100	86	100
	Testing	100	100	100	100	100

Table 3 Evaluation of K- NN Models for Too Many High CGPA Anomalies

K-NN Anomaly		**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 Score (%)**	**Specificity (%)**
Too many High CGPA	Training	96	85	96	90	96
	Testing	100	100	100	100	100

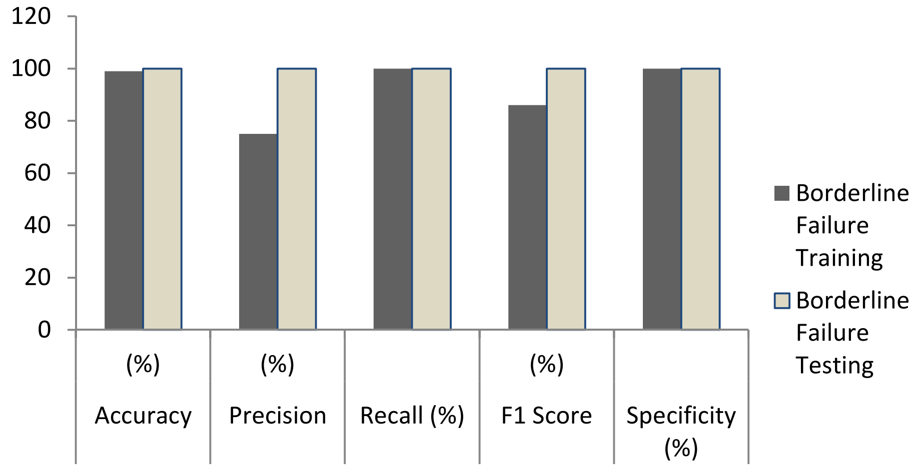

Figure 2: Chart showing the training and testing results of Inconsistent CA vs Exam Score Anomaly

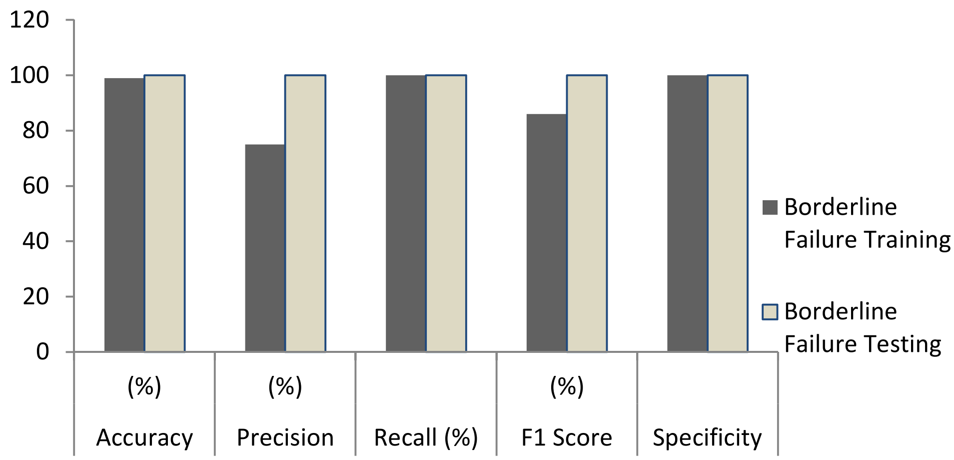

Figure 3: Chart showing the training and testing results of Borderline Failure Anomaly

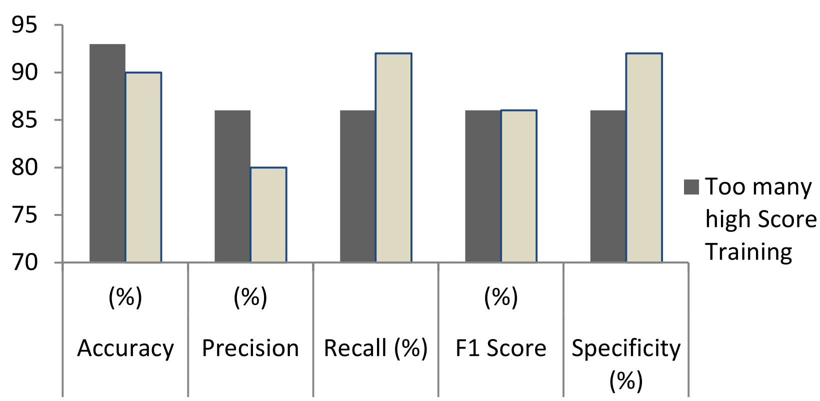

Figure 4: Chart showing the training and testing results of Too many High Score Anomaly

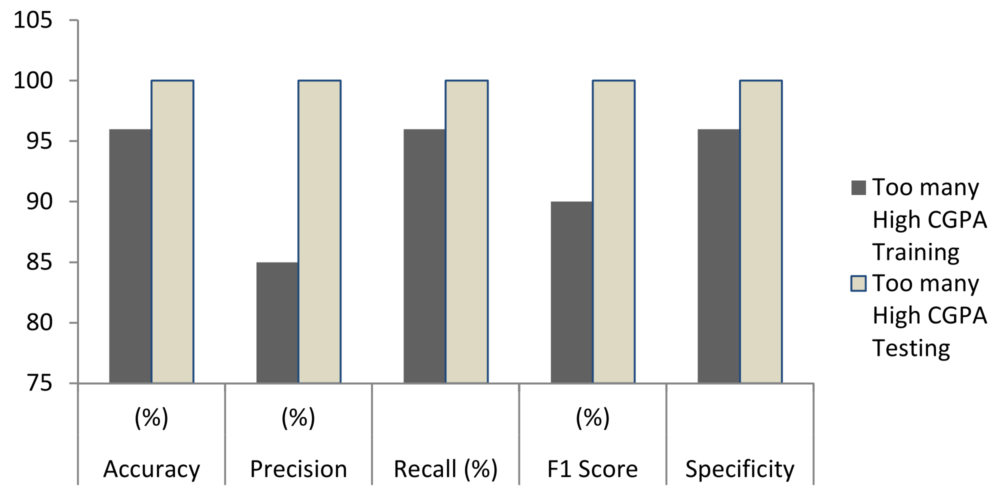

Figure 5: Chart showing the training and testing results of Too many High CGPA Anomaly

6. DISCUSSION OF RESULTS

Results from the research are discussed next.

1. **Inconsistent CA vs Exam score Anomaly**
   Results from Table 2 which presents K-Nearest Neighbour performance for the Inconsistent CA vs Exam anomaly shows that all 26 anomalous instances in the training dataset were correctly detected by the K-NN model and 93 normal instances out of 94 instances in the training dataset were correctly classified, giving a specificity of 100%, accuracy of 99%, recall of 100%, precision of 96% and F1 Score of 98%. Testing results showed that all 6 anomalous instances and 34 normal instances were classified correctly, giving a specificity of 100%, accuracy of 100%, recall of 100%, precision of 100% and F1 Score of 100%.
2. **Too Many High Score Anomaly**
   Results from Table 1, which presents K-Nearest Neighbour performance for Too many high score anomaly, showed that 26 anomalous instances out of the 27 anomalous instances in the training dataset were correctly classified by the K-NN model and 88 normal instances out of 93 instances in the training dataset were correctly classified giving a specificity of 86%, 93% accuracy, recall of 86%, precision of 86% and F1 Score of 86%. Testing results showed that out of the 14 anomalous instances, 13 were classified correctly and 23 out of the 26 normal instances were correctly classified, giving a specificity of 92%, accuracy of 90%, recall of 92%, precision of 80% and F1 Score of 86%.
3. **Borderline Failure Anomaly**
   From Table 2, the performance of the K-Nearest Neighbour model built for the borderline failure anomaly, showed that no anomalous instances out of the 2 anomalous instances in the training dataset were classified by the K-NN model and all 118 normal instances in the training dataset were classified correctly, giving 99% accuracy, recall of 100%, specificity of !00%, recall of 86% and precision of 75%. Testing result showed that all 3 anomalous instances and 37 normal instances were correctly classified, giving a specificity of 100%, accuracy of 100%, recall of 100%, precision of 100% and F1 Score of 100%.

4. Too Many High CGPA Anomaly

Table 2 presents K-Nearest Neighbour performance for Too many high CGPA anomaly. The result shows that 22 anomalous instances out of the 23 anomalous instances in the training dataset were classified correctly by the K-NN model and 85 normal instances out of 89 instances in the training dataset were correctly classified, giving a specificity of 96%, accuracy of 96%, recall of 96%, precision of 85 % and F1 Score of 90%. Testing results showed that all 7 anomalous instances and 31 normal instances were correctly classified, giving a specificity of 100%, accuracy of 100%, recall of 100%, precision of 100% and F1 Score of 100%.

7. CONCLUSION AND RECOMMENDATIONS

In this paper, we designed and applied KNN, a supervised learning machine learning algorithm to improve the detection and classification of anomalies in course-based and CGPA based student assessment scores. Testing results showed over 95% in the accuracy detection rate, precision, F1 score, recall (sensitivity) and specificity in detecting and classifying anomalies in inconsistent CA vs Exam Score anomaly and the borderline anomaly which had few instances of anomalies. In future, we will plan to use a hybrid model that combines KNN and with an unsupervised machine learning technique (e.g., Expectation–Maximization (EM) Clustering using Gaussian Mixture Models (GMM)) to improve the detection and classification of anomalies in assessment scores. GMMs assume that the data points are Gaussian distributed, and this means that the shape of the clusters can be described using multiple parameters including the mean and standard.

REFERENCES

1. Agrawal S. and J. Agrawal (2015). Survey on Anomaly Detection Using Data Mining Techniques, *Procedia Computer Science: 19th International Conference on Knowledge Based and Intelligent Information and Engineering Systems*, 6, 708 - 713
2. Buczak A. L and E. Guven (2016). A survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection, *IEE Communication Surveys*, 24, 1153 - 1176
3. Chandola V., A. Banerjee and V. Kumar (2009). Anomaly Detection: A survey, *ACM Computing,* 59, 1-57.
4. Chaurasin S. and J. Anurag (2013). Review: Ensemble neural network and KNN classifiers for intrusion detection, *International Journal of Scientific & Engineering Research*, 5, 213-217.
5. Cortez P. and A. Silva (2008). Using data mining to predict secondary school students performance. *EUROSIS*.
6. Djenouri Y., B. Asma, C. L. Jerry. and C. Alberto. (2019). Adapted K-Nearest Neighbours for detecting anomalies on spatio-temporal Traffic flow, *IEEE Access*, 13, 10015-10027.
7. Dokas P., L. Ertoz, V. Kumar, A. Lazarevic, J. Srivastava, P.N Tan (2002).Data mining for network intrusion detection, *In Proceedings of NSF Workshop on Next Generation Data Mining*; 10, 21-30
8. Hamid S.S.A. , N. Admodisastro, N. Manshir, A. Kamaniddin and A.A.A. Chani (2018). Dyslexia adaptive learning model: Student engagement prediction using machine learning approach in advances to intelligent systems and computing, Springer Berlin Germany, 13, 372-384.
9. Manocha, S. and M. Girolami (2007). An empirical analysis of the probabilistic K-nearest Neighbor Classifier*,* *Pattern Recognition Letters*, 7, 1818–1824
10. Oladeji P. A. and I. A Adeleke (2017). An improved KNN classifier for anomaly detection system using cluster optimization, *International Journal of Computer Science and Telecommunications,* 5, 34-38.

11. Parmar J. D. and J. T, Patel (2017). Anomaly Detection in Data Mining: A Review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 9, 32- 40
12. Pramit C. (2017). Introduction to Anomaly Detection, *Data Science.com.* Retrieved on 4th April, 2019 from http/datascience.com/blog/python-anomaly-detection.html.
13. Raval K. M. (2012). Data Mining Techniques, International *Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, 4, 439-442
14. Salami H.O. .S. I. Ruquyyah, M. O. Yhaya (2016). Detecting Anomalies in Students' Results Using Decision Trees, *International Journal of Modern Education and Computer Science,* 7, 31-40.
15. Salami H.O., M. O. Yahaya (2018). On the use of extreme learning machines for detecting anomalies in students' result. 2nd *International Conference on Information and Communication Technology and its Applications (ICTA), Federal University of Technology, Minna, Nigeria,* 7, 402- 408.
16. Shahreza L. M. , D. Moazzami, B. Moshiri, M.R. Delavar (2011). Anomaly Detection using a Self Organizing Map and Particle Swarm Production, *Scientia Iranica,* 18(6), 1460 - 1468
17. Thomas C. and D. B. Jayagopi (2017). Predicting student engagement in classrooms using facial behaviour cues, Proceedings of 1st ACM SIGCHI, *International Workshop on Multimodal Interaction for Education,* 8, 33-40
18. Vemur R.V. and Y. Liao (2002). Use of K-Nearest Neighbour Classifier for intrusion detection, *Journal of Computing and Security,* 21, 439-448.
19. Viswanath B., M. A. Bashir, M. Crovella, S. Guha, K. P. Gummadi, B. Krishnamurthy and A. Mislove (2014). Towards detecting anomalous user behaviour in online social network, *Proceedings of 23rd USENIX Security Symposium.* San Diego, CA: USENIX, USENIX Security Symposium (USENIX Security 14), 16, 223-238.
20. Xiaodan X., H. Liu and M. Yao (2019). Recent Progress of Anomaliy detection. *Complexity*, 11, 2686378