

## CLASSIFICATION MODEL AND ANALYSIS ON STUDENTS' PERFORMANCE

H. Nawang, M. Makhtar\* and S. N. W. Shamsudin

Faculty of Informatics and Computing, Universiti Sultan ZainalAbidin, Kuala Terengganu,  
Terengganu, Malaysia

Published online: 10 November 2017

### ABSTRACT

The purpose of this paper is to propose a classification model for classifying students' performance in *SijilPelajaran Malaysiain* order to help teachers plan suitable teaching activities for their students based on the students' performance. Five classifier algorithms have been used during the process which are Naïve Bayes, Random Tree, Multi Class Classifier, Conjunctive Rule and Nearest Neighbour. Data was collected from *MaktabRendahSains MARA* Kuala Berang, Terengganu, Malaysia starting May 2011 until December 2014. The students' performance was evaluated based on the category of students according to their SPM Results. Parameters that contribute to students' performance such as stream, state, gender and hometown are also investigated along with the examination data. This research shows that first semester results can be used to identify students' performance.

**Keywords:** educational data mining; classification model; feature selection.

doi: <http://dx.doi.org/10.4314/jfas.v9i6s.65>

### 1. INTRODUCTION

Data mining, also known as knowledge discovery is the process of analyzing data from different perspectives and summarizing it into useful information. Data mining techniques involve a large amount of data with various parameters to be explored. The advantage of this

technique is its capability to discover underlying relationships and patterns that exist within the data [1]. It combines machine learning and statistical and visualization techniques to discover and extract knowledge in such a way that humans can easily interpret [2].

Currently, there is increasing interest in educational data mining. Educational data mining (EDM) is an emerging interdisciplinary research area that deals with the development of methods to explore data originating in the educational context. EDM uses computational approaches to analyze educational data in order to study educational questions [8]. Educational Data Mining is concerned with developing methods to explore data that come from educational systems and using those methods to understand students' behaviours and performance [3]. Overall, the goal of data mining of educational data is to discover suitable methods and approaches in ensuring the excellence of students.

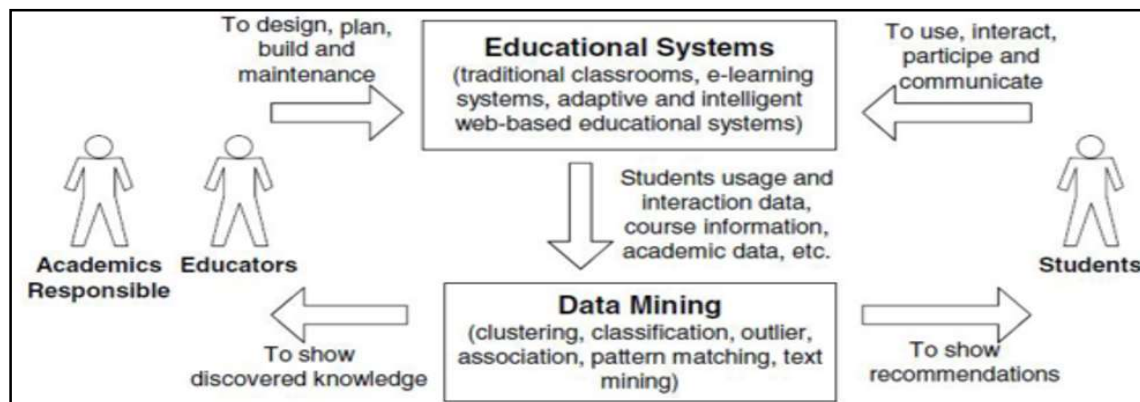
The emerging of educational data mining have been increasing as well as the increasing of data that come from many sources, including standardized tests, quizzes [21], assignments [22] and semester examination [23]. Predicting students' performance becomes more challenging due to the large volume of data in educational databases. In the last two decades, researchers have realized that this data can be useful to teachers and students. With the increase of data storage in educational sector such as online web-based learning system and students' learning profile, the important of data mining in educational field need to be addressed. Currently in Malaysia, the lack of existing system to analyze and monitor the student progress and performance is not being emphasized. As a result, important decisions are not made based on the information retrieved from a database but on intuition of a decision maker who usually is the leader in educational sector.

Examination will play a vital role in any student's life. The marks obtained by a student in the examination decides his/her future. In *Maktab Rendah Sains MARA* (MRSM), both formative and summative evaluation are employed in assessing the effectiveness of the education system. Every semester, the students are evaluated through standardized tests, quizzes, assignments and semester examinations. Those information if properly extracted using data mining techniques will enhance data sharing in order to analyze hidden knowledge such as student relationship management; predict student performances and the effectiveness of academic program [24].

In line with the vision and mission of MRSM Kuala Berang to produce *Bumiputera* students with high potential in the field of science and technology, an appropriate approach has to be applied to educational data to discover hidden patterns and information. Among the challenges faced by educators at the MRSM is to ensure that the excellent students are always

consistent, especially in the science and mathematics subjects. Subject teachers must analyze tests and examination results to identify students who have problems, and therefore formulate an appropriate approach in dealing with lower achievers. To extract the hidden patterns underlying the data in the Student Information System (SIS), a method of data mining is an appropriate approach to be used in helping teachers to analyze and discover useful information to deal with their students. The information gained from the data mining techniques will help teachers classify their students' Student Performance Analysis (SPA) for *Sijil Pelajaran Malaysia* (SPM).

There have been a lot of studies in educational data mining lately. Researchers in educational data mining have been trying to enhance the understanding of the learning process, analyze learning behaviour to predict results and identify factors that correlate with students' performance using different patterns of data. Fig. 1 shows how the applied cycle of data mining in the educational system is utilized to meet the needs of students, educators and the administration as well.



**Fig.1.** The cycle of applying data mining in the educational system [6]

Data mining can be used in the educational field to enhance our understanding of the learning process by focusing on identifying, extracting and evaluating variables related to the learning process of students as described by [4]. In [5] had done a research using the classification task on a student database to predict the students' division on the basis of a previous database. As there are many approaches used for data classification, they chose the decision tree method. Information like attendance, class test, seminar and assignment marks were collected from the previous database to be used to predict the performance of the same students at the end of the semester. Their study will help students and teachers improve the division of students and also work to identify students who need special attention in order to reduce the failure ratio and

take appropriate actions for the next semester examinations.

In another study conducted by [6] during the summer semester among first year students from the Faculty of Economics at the University of Tuzla, they had used three data mining algorithms which are Naïve Bayes, Decision tree and Neural Networks. The algorithms were used to predict success in a course (either pass or fail) and the performance of the learning methods were evaluated based on their predictive accuracy, ease of learning and user friendly characteristics. They found that the results indicate that the Naïve Bayes classifier outperforms the prediction decision tree and neural network [30-33] methods.

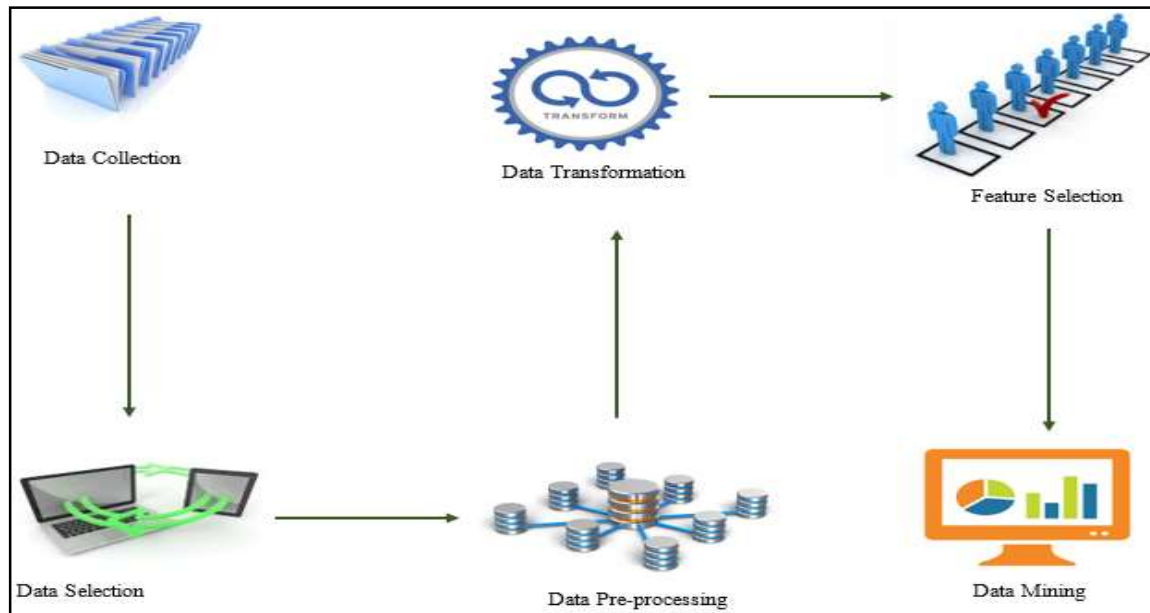
In this research, data mining techniques are used to classify students' of MRSM Kuala Berang performance based on their Individual Grade Point Average (GPI). The main objective of this paper is to use data mining techniques to classify the students' performance in SPM based on semester results. Data was collected from the first semester until the fourth semester. Students were divided into three streams namely Science Arabic, Pure Science and Science Accounts. A major number of students were in the Pure Science and Arabic streams, which involve 11 subjects. The GPI system was first used by the Malaysian Examination Board starting in 2009, where each grade is represented by the numbers 0 to 9. Table 1 shows the number that represents each grade.

**Table 1.** Numbers representing each grade in semester examinations

<b>Grade</b>	A+	A	A-	B+	B	C+	C	D	E	G
<b>Number Represent Grade</b>	0	1	2	3	4	5	6	7	8	9

## 2. METHODOLOGY

During the research, there were six stages involved in the SPA process which are data collection, data selection, data pre-processing, data transformation, feature selection and finally data mining as shown in the SPA proposed framework. Fig. 2 shows the proposed framework for the SPA architecture. Brief descriptions for every stage as shown in Fig. 2 are explained as below.



**Fig.2.**The proposed framework for SPA architecture

**2.1. Data Collection**

The implementation for data mining in this research is composed of three batches starting from year 2011 until 2014. Data were collected from upper secondary students of MRSM Kuala Berang, Terengganu, Malaysia beginning from semester one until semester 4. The data collected from the SIS of MRSM Kuala Berang involves 491 students’ records that were separated into four semesters in order to identify the percentage of data mining results based on semester. The SIS contains details about student data, standardized test and examination data and also the co-curriculum data. While examination result for *Sijil Pelajaran Malaysia* (SPM) comes in the form of hard copy from the *Pejabat Pendidikan Daerah* (PPD) Kuala Berang, Terengganu. Table 2 shows the number of students who represent each year.

**Table 2.**Number of students for each batch

Year	Number of Students
2011-2012	169
2012-2013	167
2013-2014	159

**2.2. Data Selection**

Data selection is a process of selecting a target data set or focusing on a subset of variables on which the finding is to be performed. The SIS for MRSM Kuala Berang contains data about students’ personal information, semester examination, test results and co-curriculum activities.

However, some files were not kept in the same system such as SPM Results which was kept in Microsoft Office Excel 2010 format.

Before integrating those files, the first step in data selection is to identify the relevant attributes in order to make the integration process easier when the target sources is already known. In this research, the attributes selected are the subjects in semester examinations such as BM (Malay Language), BI (English), PI (Islamic Education), SJ (History), MM (Mathematics), MT(Additional Mathematics), AC (Accounting Principles), SN (Art), FZ (Physics), KM (Chemistry), BO (Biology), BA (Arabic Language), QS (Al-Quran and Sunnah), SI (SyariahIslamiah) along with the data on students' gender, homeroom, stream, state, individual grade point average (GPI) for semester, category for semester results and category for SPM results.

### **2.3. Data Pre-Processing**

The next phase in this research is data pre-processing. During the data pre-processing, there are two processes involve which are data cleaning and data integration.

#### **2.3.1. Data Cleaning**

Data cleaning is also known as data scrubbing is the process of identifying errors, inaccurate or incomplete records from a table and then replace the data with the correct format of data [25]. The importance of data cleaning process is to maintain the quality of the data by identifying the missing values or remove outliers and resolve inconsistencies [26].

During the data cleaning process, instances are checked to ensure that the record is complete before the mining process is done in order to provide quality data mining results. Table 3 shows the incomplete data of students' results for the whole semester, where the grades contain TS values. This may happen when a student transfers to another school, but the school administrator did not remove his/her record before processing the semester's examination results. The data has to be removed in order to ensure that the data mining results are not distracted due to unambiguous values. After the cleaning process, there were 488 records left since three records were removed due to missing values for the selected parameters.

**Table 3.**The missing values

GENDER	HOMETOWN	1BM	1BI	1PI	1SJ	1MM	1BA	1SN	1MT	1AC	1FZ
P	PAHANG	B	A	B	A	C			C		E
P	TERENGGANU	B+	B	B+	A	E			G	G	C
P	TERENGGANU	B+	A-	B+	B+	C			G		A-
L	PAHANG	C	C+	B	A-	E			G		C
L	TERENGGANU	TS	TS	TS	TS	TS			TS		TS
P	WILAYAH PERSEKUTUAN	TS	TS	TS	TS	TS			TS		TS
P	KELANTAN	TS	TS	TS	TS	TS			TS		TS

### 2.3.2. Data Integration

Data integration is the process of combining data located in different location and allow the unified view called global schema. The global schema is the interface where users submit their queries to a data integration system. The quality of the data sources can dramatically change as data may be incomplete, inaccurate or out of date. In fact, the quality of the result depends mainly on two factors: the quality of the data at the data sources and the manipulation process that builds the resulting data from the data sources [27].

Data for this research involved students from three different years, so the data consist of three databases that needed to be merged together. A total of 488 students' data were chosen as representative of the sample. Data contained in the SIS can be accessed by different users such as teachers, school management and parents. In this research, the situation is to integrate data from students' information system based on user queries that were kept in 3 different sources. To integrate these files, these three schemes need to be match to find the similarity between the data. In order to ensure the consistency of the data, the primary key of each table needs to be identified. One of the minimum standard upon designing data, development and execution of data integration process is assuming all dimension tables have surrogate keys that function as the primary key[28]. The primary and the foreign key cannot contain the null value. An unhandled null value can destroy the integration process.

Queries were used in data integration to formulate the information request by user. Query over a table is normal, but mostly queries require more than a table to display. Join queries usually used when user set the limitation on what to display based on how the tables are connected to each other. There are two common structured Query Language that has been used in this research that are Select and Inner Join. The SQL Select statement returns a result set of

records from one or more tables. The second SQL language is Inner Join that used to combine all the records from two or more tables whenever there are matching values in a common field. Inner joins will produce query results by combining rows between two tables or more to match up the tables based on specific criteria given by the user before the queries are mapped to a global schema.

#### 2.4. Data Transformation

During data transformation, some data need to go through a normalization process. This happens because some of the attributes that contain numerical values such as Stream need to be converted into categorical classes. Stream was grouped into three groups; science Arabic, pure science and science accounts. While the hometown parameter was categorized into four groups which are east, west, south and north based on state data. The category parameter was grouped into five groups which are excellent, good, average, poor and very poor. Table 4 shows the category for grade point per individual. Each category depends on the result of GPI achieved by students.

**Table 4.**Category for grade point per individual

GPI	Category
0-2	Excellent
2.1-3	Good
3.1-5	Average
5.1-7	Poor
7.1-9	Very Poor

Details of the attributes involved in this research are as shown in Table 5. It represents data for the first semester. However, the names of parameters in the other three semesters are still the same.

**Table 5.**The final attributes for DM process

Parameter	Category
Stream	Science Arabic,Pure Science, Science Account
Gender	L, P
State	East, West, South, North
BM	A+, A, A-,B+,B,C+,C, D,E,G



---

BI	A+, A, A-,B+,B,C+,C, D,E,G
PI	A+, A, A-,B+,B,C+,C, D,E,G
SJ	A+, A, A-,B+,B,C+,C, D,E,G
MM	A+, A, A-,B+,B,C+,C, D,E,G
SN	A+, A, A-,B+,B,C+,C, D,E,G
MT	A+, A, A-,B+,B,C+,C, D,E,G
AC	A+, A, A-,B+,B,C+,C, D,E,G
FZ	A+, A, A-,B+,B,C+,C, D,E,G
KM	A+, A, A-,B+,B,C+,C, D,E,G
BO	A+, A, A-,B+,B,C+,C, D,E,G
BA	A+, A, A-,B+,B,C+,C, D,E,G
QS	A+, A, A-,B+,B,C+,C, D,E,G
SI	A+, A, A-,B+,B,C+,C, D,E,G
GPP	Numerical Value
Category_1	Excellent
	Good
	Average
	Poor
	Very Poor
Category	Excellent
	Good
	Average
	Poor
	Very Poor

---

## 2.5. Feature Selection

The next process is the attribute evaluation that also known as feature selection has been used to the dataset. Feature selection is a pre-processing stage used to reduce dimensionality and delete irrelevant data to increase learning accuracy and improve result comprehensibility [15].

The purpose of attribute evaluation is to increase classification accuracy by eliminating noise

features [29]. Feature selection helps in knowledge discovery by selecting relevant features to a particular application. The advantage of attribute selection is that no information about the importance of single features is lost. During the feature selection process of this research, some instances were removed to increase the accuracy of the mining results. Using cfs Subset Eval as the attribute evaluator together with the Best first methods has shown increase of data mining result percentage on some classifiers.

The next process before data mining was splitting the data into two parts that are testing and training set. Training set and testing set are the most common approaches applied in supervised learning. In this research, the 10 cross fold validation have been used on the dataset. Cross-Validation is a statistical method of evaluating and comparing learning algorithms by dividing data into two segments: one used to learn or train a model and the other used to validate the model. The basic form of cross-validation is k-fold cross-validation that will split the original dataset into k equal parts. In k-fold cross-validation, the data is first partitioned into k equally (or nearly equally) sized segments or folds.

## **2.6. Data Mining**

The data mining phase is the important part in KDD process. The data mining process will apply intelligent methods to extract useful patterns from the data sets. During this phase, the classification will play the role. For the data mining stage, WEKA tools that were developed in University of Waikato in New Zealand were used to determine the classification of students' performance. The selected Classifiers used during the data mining process are Naïve Bayes, Random Tree, Multi Class Classifier, Conjunctive Rule and Lazy IB1. During the classification process, a model will be build using a training dataset and verified by using testing dataset.

The first applied algorithm is Random Tree (RT). Random Trees are formed by a stochastic process [10]. Random binary trees are binary trees with a given number of nodes, formed by inserting the nodes in a random order or by selecting all possible trees uniformly at random. Random trees can also be formed using spanning methods. The second algorithm applied in this research is the Nearest Neighbour Classifier (IB1). This algorithm uses normalised Euclidean distances to find the training instances closest to the given test instances and predict the same class as these training instances.

The third applied algorithm is the Naïve Bayes Classifier (NBC). The Naïve Bayes Classifier uses the Bayes Theorem which assumes that the effect of an attribute value of a given class is

independent on the values of the other attributes. It represents a descriptive and predictive approach to predict the class membership for a target tuple [14]. The fourth classifier used in this research is the Multi Class Classifier (MCC). The MCC is a meta classifier in handling multiclass datasets with two class classifiers. The Conjunctive Rule (CR) classifier was used in this research as the fifth test on the data. The CR is a class that implements a single CR learner that can predict numeric and nominal class labels.

The Bayes Theorem provides the calculation of the posterior probability,  $P(c|x)$ , from  $P(c)$ ,  $P(x)$  and  $P(x|c)$ . The Naive Bayes classifier assumes that the effect of the value of a predictor ( $x$ ) on a given class ( $c$ ) is independent of the values of other predictors. This assumption is called class conditional independence [9]. The following is the Naïve Bayes formula:

$$P(c|x) = \frac{P(x|c) P(c)}{P(x)}$$

where  $P(c)$  is the prior probability of class that reflects background knowledge due to the chance of  $c$  to be correct. In this research, the  $P(c)$  can be representative as  $P(\text{Good})$ .  $P(x)$  is the probability of  $x$  to be observed. In this research, the  $P(x)$  can be representative of  $P(\text{Pure Science})$ .  $P(x|c)$  is the probability of observing  $x$  given a world in  $c$  holds. In this research, the  $P(x|c)$  can be representative as  $P(\text{Pure Science}|\text{Good})$ .  $P(c|x)$  is the posterior probability of class (target) given predictor (attribute).  $P(x|c)$  can be representative as  $P(\text{Good}|\text{Pure Science})$ . To specify the target class, the prior probability and conditions generated from a ten-fold cross validation were used to classify the data by using this formula:

$$P(c|x) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

### 3. RESULTS AND DISCUSSION

During the experiment, data were divided into a training set to train the model and a test set to evaluate the data using Cross-Fold Validation. As for the test options, the 10 cross fold validation applied in the experiment which means the dataset was divided into 10 parts or folds. Each fold is used once for testing and nine for training. As there are 488 total data, the separation were 49 data used for testing and 439 data used for training.

As explained in section 2, five classifiers were used in these experiment, which are Random Tree, Nearest Neighbour (IB1) and Naïve Bayes, Multi Class Classifier and Conjunctive Rule. The accuracy results during the experiment were recorded and presented in Table 6. For cross-fold validation, the data were split into 10 subsets for each classifier. Results were also

divided into 4 semesters starting from semester 1 to semester 4 with and without Feature Selection.

**Table 6.**Data mining results before feature selection

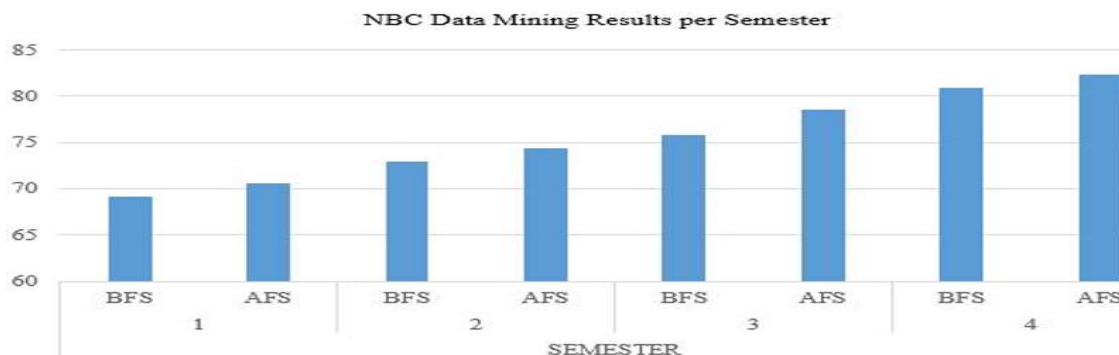
<b>Before Feature Selection</b>					
<b>Semester</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	
RT	63.40	67.42	63.52	72.54	
NBC	69.12	72.95	75.82	80.94	
Classifier	IB1	64.82	65.78	68.24	78.28
	MCC	61.35	60.66	64.14	67.21
	CR	66.67	66.80	67.83	73.57

From the results shown in Table 6, it could be observe that every classifier shows the ascending in result percentage starting from semester 1 to semester 4. However, the 10 cross-fold validations in the Naïve Bayes Classifier for the fourth semester show the highest percentage compared to the other classifiers with the result 80.94%. This may happen due to the factors that come from students and teachers themselves.

**Table 7.**Data mining results after feature selection

<b>After Feature Selection</b>					
<b>Semester</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	
RT	63.40	71.31	70.90	75.61	
NBC	70.55	74.39	78.48	82.38	
Classifier	IB1	63.40	69.68	75.95	75.20
	MCC	67.28	74.59	77.25	80.12
	CR	66.67	66.81	67.83	73.57

Meanwhile, when the feature selection method applied to the data mining process, the results in Table 7 shows the increase accuracy of the percentage for every classifier in each semester except for the Conjunctive Rule. The percentage increase about 1.4% to 13% for the certain classifier. However, it could also be observed that Naïve Bayes classifier still holding the highest accuracy result with the 82.38%.



**Fig.3.** Data mining results for Naïve Bayes Classifier

Fig. 3 shows the data mining results for Naïve Bayesian Classifier from semester 1 until semester 4 before (BFS) and after (AFS) feature selection method applied. The graph shows the increasing of the classification accuracy for every semester.

As shown in Fig. 4, the result indicates that the model has correctly classified 402 instances among 488 students. The results obtained from the feature selection process are the selected parameters which are subjects BI, PI, SJ, MM, MT, FZ, KM and also Grade Point per Individual. The selected parameters shows that the certain subjects have influenced in determine the students' results for SPM.

Therefore, it can be concluded that the students' mastery of Core subject such as English, Islamic Education and History is as essential as having the mastery in Science and Mathematics Subjects. Having full grasp in both the core subject and the elective subject will somehow influence the students' overall academic performance. Proving that both of the subjects are of equally important in determining the students' excellent performance.

```

==== Stratified cross-validation ====
==== Summary ====
Correctly Classified Instances      402      82.377 %
Incorrectly Classified Instances    86       17.623 %
Kappa statistic                    0.715
Mean absolute error                 0.0939
Root mean squared error             0.2742
Relative absolute error             31.0096 %
Root relative squared error         70.537 %
Total Number of Instances          488

==== Detailed Accuracy By Class ====
      TP Rate   FP Rate   Precision   Recall   F-Measure   ROC Area   Class
      0.837     0.039     0.96       0.837   0.894     0.969     EXCELLENT
      0.84      0.18     0.661     0.84    0.74     0.895     GOOD
      0.767     0.037     0.815     0.767   0.79     0.965     AVERAGE
      0         0         0         0       0        0.848     POOR
Weighted Avg.   0.824     0.08     0.844     0.824   0.828     0.946

==== Confusion Matrix ====
a  b  c  d  <-- classified as
215 42  0  0 | a = EXCELLENT
 9 121 14  0 | b = GOOD
 0 20 66  0 | c = AVERAGE
 0  0  1  0 | d = POOR
    
```

**Fig.4.**The 10 Cross fold validation for Naïve Bayes Classifier after feature selection

#### 4. CONCLUSION

In this paper, we study the various methods in classifying the model analysis on students' performance in SPM for MRSM Kuala Berang. From the experiments, the results show that:

1. The accuracy of classification increase from semester 1 until semester 4 for all classifiers used in this research.
2. Feature selection method improves the accuracy for Random Tree, Naïve Bayes, Nearest Neighbour (IB1) and Multi Class Classifier.
3. Naïve Bayes Classifier has shown the highest accuracy with or without feature selection.

The aim of this paper is to classify students' performance based on data collected from the SIS. Through the pre-processing stage, the data were cleaned and the variables were identified and selected to be categorized according to students' performance for SPM. Through this research, the school management and teachers can see the pattern of students' performance increasing from semester 1 until semester 4. To our knowledge, this paper has two major contributions:

1. Significant parameters can be used as a guideline for teachers and school management to plan suitable teaching and learning activities for students based on their performance.
2. The selected parameters and Naïve Bayes Classifier that effect the accuracy of classification for MRSM Kuala Berang such as discussed in section 3.

This paper should be further enhanced in the future with more data from other schools with more parameters and more classification techniques applied, in order to get quality data mining results.

#### 5. ACKNOWLEDGEMENTS

This work is partially supported by UniSZA and KPT (Grant No.FRGS/2/2013/ICT07/UniSZA/02/2).

#### 6. REFERENCES

- [1] Kumar V, Chadha A. An empirical study of the applications of data mining techniques in higher education. *International Journal of Advanced Computer Science and Applications*, 2011, 2(3):80-84
- [2] Zhang Y, Oussena S, Clark T, Kim H. Using data mining to improve student retention in higher education-A case study. *In 12th International Conference on Enterprise Information Systems*, 2010, pp. 1-8

- [3] Baker R S, Yacef K. The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 2009, 1(1):3-17
- [4] Romero C., Ventura S. *Data mining in e-learning*. Southampton: Wit Press, 2006
- [5] BaradwajBnK, Pal S. Mining educational data to analyze students' performance. *International Journal of Advanced Computer Science and Applications*, 2011, 2:63-69
- [6] Osmanbegović E, Suljić M. Data mining approach for predicting student performance. *Economic Review*, 2012, 10(1):3-12
- [7] Wang X, Huang L P, Zhang Y, Xu X H, Chen J Q. A solution of data inconsistencies in data integration-designed for pervasive computing environment. *Journal of Computer Science and Technology*, 2010, 25(3):499-508
- [8] Romero C, Ventura S. Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 2010, 40(6):601-618
- [9] Aziz A A, Ismail N H, Ahmad F. Mining students' academic performance. *Journal of Theoretical and Applied Information Technology*, 2013, 53(3):485-495
- [10] Srimani P K, Balaji K. A comparative study of different classifiers on search engine based educational data. *International Journal of Conceptions on computing and Information Technology*, 2014, 2:6-11
- [11] Priyama A, Abhijeeta R G, Ratheeb A, Srivastavab S. Comparative analysis of decision tree classification algorithms. *International Journal of Current Engineering and Technology*, 2013, 3(2):334-337
- [12] Kumar V, Chadha A. Mining association rules in student's assessment data. *International Journal of Computer Science Issues*, 2012, 9(5):211-216
- [13] Leung C. S., Lee M., Chan J. H. *Neural information processing: 16th International Conference, ICONIP 2009, Bangkok, Thailand, December 1-5, 2009, Proceedings*. Berlin: Springer Verlag, 2009
- [14] Anoopkumar M, Rahman A M. A comprehensive survey on educational data mining and use of data mining techniques for improving teaching and predicting student performance. *Advances in Innovative Engineering and Technologies*, 2015
- [15] Kabir M M, Islam M M, Murase K. A new wrapper feature selection approach using neural network. *Neurocomputing*, 2010, 73(16):3273-3283

- [16] Kamarudin N S, Makhtar M, Fadzli S A, Mohamad M, Mohamad F S, Kadir M F. Comparison of image classification techniques using Caltech 101 dataset. *Journal of Theoretical and Applied Information Technology*, 2015, 71(1):79-86
- [17] Muhammad S Y, Makhtar M, Rozaimie A, Aziz A A, Jamal A A. Classification model for water quality using machine learning techniques. *International Journal of Software Engineering and Its Applications*, 2015, 9(6):45-52
- [18] Ismail M R, Awang M K, Rahman M N, Makhtar M. A multi-layer perceptron approach for customer churn prediction. *International Journal of Multimedia and Ubiquitous Engineering*, 2015, 10(7):213-222
- [19] Nafis S, Makhtar M, Awang M K, Rahman M N A, Deris M M. Feature selection and classification model for customer churn. *Journal of Theoretical and Applied Information Technology*, 2015, 75:356-365
- [20] Kavyashree K R, Lakshmi D. A review on mining students' data for performance prediction. *International Journal of Advanced Research in Computer and Communication Engineering*, 2016, 5:1104-1106
- [21] Al-Musharraf A, Alkhatabi M. An educational data mining approach to explore the effect of using interactive supporting features in an LMS for overall performance within an online learning environment. *International Journal of Computer Science and Network Security*, 2016, 16(3):1-13
- [22] Tair M M, El-Halees A M. Mining educational data to improve students' performance: A case study. *International Journal of Information*, 2012, 2(2):140-146
- [23] Superby J F, Vandamme J P, Meskens N. Determination of factors influencing the achievement of the first-year university students using data mining methods. In *Workshop on Educational Data Mining 2006*, pp. 1-8
- [24] Ranjan J, Malik K. *Effective educational process: A data-mining approach*. Vine, 2007, 37(4):502-515
- [25] Kochar B, Chhillar RS. An effective data warehousing system for RFID using novel data cleaning, data transformation and loading techniques. *International Arab Journal of Information Technology*, 2012, 9(3):208-216
- [26] Elgamal A F, Mosa N A, Amasha N A. Application of framework for data cleaning to handle noisy data in data warehouse. *International Journal of Soft Computing and Engineering*, 2014, 3:226-231



- [27] Abdel-Moneim M S, El-Bastawissy A H, Kholief M H. Data quality based data integration approach. *World of Computer Science and Information Technology Journal*, 2015, 5(10):155-164
- [28] Han J., Pei J., Kamber M. *Data mining: Concepts and techniques*. Amsterdam: Elsevier, 2011
- [29] Romero C., Ventura S., Pechenizkiy M., Baker R. S. *Handbook of educational data mining*. Florida: CRC Press, 2010
- [30] Yassin I M, Jailani R, Ali M, Baharom R, Hassan A, Rizman Z I. Comparison between cascade forward and multi-layer perceptron neural networks for NARX functional electrical stimulation (FES)-based muscle model. *International Journal on Advanced Science, Engineering and Information Technology*, 2017, 7(1):215-221
- [31] Zabidi A, Yassin I M, Hassan H A, Ismail N, Hamzah M M, Rizman Z I, Abidin H Z. DETECTION of asphyxia in infants using deep learning convolutional neural network (CNN) trained on Mel frequency cepstrum coefficient (MFCC) features extracted from cry sounds. *Journal of Fundamental and Applied Sciences*, 2017, 9(3S):768-778
- [32] Hashim F R, Daud N N, Ahmad K A, Adnan J, Rizman Z I. Prediction of rainfall based on weather parameter using artificial neural network. *Journal of Fundamental and Applied Sciences*, 2017, 9(3S):493-502
- [33] Hashim F R, Adnan J, Ibrahim M M, Ishak M T, Din M F, Daud N G, Rizman Z I. Heart abnormality detection by using artificial neural network. *Journal of Fundamental and Applied Sciences*, 2017, 9(3S):1-10

**How to cite this article:**

Nawang H, Makhtar M, Shamsudin S N W. Classification model and analysis on students' performance. *J. Fundam. Appl. Sci.*, 2017, 9(6S), 869-885.