

# Self-motivation and Academic Performance In Computer Programming Language Using a Hybridised Machine Learning Technique

**Isaac Kofi Nti**

*Department of Computer Science  
Sunyani Technical University  
Sunyani, Ghana*

*ntious1@gmail.com*

**Juanita Ahia Quarcoo**

*Department of Computer Science  
Sunyani Technical University  
Sunyani, Ghana*

*juan83p@yahoo.com*

---

## Abstract

The advancement in artificial intelligence (AI) and Machine learning (ML) have made it easier to foreknown feature happens from current and past trends. Once Self-efficacy and self-confidence are believed to be, an individual trait associated with academic brilliance. Using a hybridised Random Forest and Support Vector Machine (RFSVM) ML model we predicted students' academic performance in computer programming courses, based on their self-confidence, self-efficacy, positive thinking, focus, big goals, a motivating environment and demographic data. Benchmarking our RFSVM model against Decision Tree (DT) and K-Nearest Neighbour (K-NN) model, the RFSVM recorded an accuracy of 98% as against 95.45% for DT and 36.36% for K-NN. The error between actual values and predicted values of the RFSVM model was better (RMSE = 0.326401, MAE = 0.050909) and compared with the K-NN (RMSE = 2.671397, MAE = 1.954545) and DT models (RMSE = 0.426401, MAE = 0.090909). The results further revealed that students with a high level of self-confidence, self-efficacy and positive thinking performed well in computer programming courses.

**Keywords:** Self-Confidence, Self-Efficacy, Individual-Interest, Positive-Thinking, Focus, Personal Goals, Motivating Environment, Academic Performance, Random Forest, Support Vector Machine.

---

## 1. INTRODUCTION

Effectively achievement of a preferred academic long-term goal requires both intelligence and effort. Even though much attention in research is regularly paid to cognitive factors, such as intellect and academic ability, non-cognitive factors significantly contribute to academic performance (Jung, Zhou, & Lee, 2017). Academic self-directive is a crucial contributor to academic achievement (Honicke & Broadbent, 2016; Jung et al., 2017; Lee, Lee, & Bong, 2014), it signifies the orderly and the active application of self-development to achieve academic goals, but students' carry out academic self-directive more efficiently when they are hugely motivated (Lee et al., 2014). Having low motivation towards study-associated tasks may be a common and frequent obstacle in self-regulated learning.

Consequently, students should regulate their motivation to attain their study-associated goals, particularly on choices and actions that affect their academic work (e.g., meeting friends). To control their motivation students can harness from the various psychological feature regulation methods, that is, self-monitoring methods that aim to manage and improve motivation (Grunschel, Schwinger, Steinmayr, & Fries, 2016). When students have strong self-believe about their

competence, high values for their academic goal, they are more likely strategised, observer and reflect their goal accomplishment and fine-tune their regulatory processes accordingly.

In (Dweck, Walton, & Cohen, 2014) the authors argue that psychological factors also known as a motivational or non-cognitive factor are higher predictors students' academic performance as compared cognitive factors in some cases. Again in (Dweck et al., 2014), authors further elaborated that these motivational factors also offer hopeful levers for levitating the achievement of underprivileged children and, eventually, narrowing achievement gaps based on race and income. In other studies, the authors see every person as an island of their own with different inbuilt motivations, ambitions, perception and abilities. Whiles self-efficacy has long been seen as a critical factor of academic performance. A counter-position is that self-efficacy is simply a mirror image of past performance (Talsma, Schüz, Schwarzer, & Norris, 2018).

Despite the numerous contribution in research on the effect of cognitive and non-cognitive factors on academic performance, little is known on how demographic data and non-cognitive factors jointly affect students' academic performance in a specific course such as computer programming language.

The present study proposes an artificial intelligence (AI) model to predict students' academic performance in programming courses such as C++, JAVA and VB.net based on their self-motivation level and demography data, using a hybridised machine learning technique of Random Forest (RF) and Support Vector Machine (SVM) (RFSVM). To achieve our stated aim, this study outline these specific objectives: (i) to proposed a hybridised Random Forest (RF) and Support Vector Machine (SVM) (RFSVM) students' academic performance prediction model. (ii) To identify the weak, learners in programming courses at the earliest stage of their computer science education, so that counsellor and management can offer them the necessary help. (iii) Reduce students' failure in computer programming courses by increasing their academic performance in programming courses. It is hypothesised that (i) student with a higher motivation level would perform better in computer programming courses (ii) Good performance in programming courses has nothing to do with age.

The current study would lead to students' academic performance prediction based on their demographic data and self-motivation level, so that course lectures and an academic counsellor will be able to take the necessary actions to identify weak learners and help them accordingly. Also, based on the outcome, the probable performance of newly admitted students can be the accessed forehead

The remaining of the current paper is organised as follows. Section 2 covers the review of the literature on students' academic performance and cognitive and non-cognitive factors. Section 3 discusses our methodology for analysing and predicting students' academic performance in programming language courses based on the degree of self-motivation. Section 4 presents finding and discusses results. Section 5 concludes this study and Section 6 describes limitations and possibilities for future research.

## 2. SELF-MOTIVATION

Self-motivation is believed to be complicated; it is linked with one's level of creativity in setting thought-provoking goals for him or herself. The following are considered for building a healthy level of self-motivation; the following four factors are needed. (i) Self-confidence and self-efficacy. (ii) Positive-thinking and positive thinking about the future. (iii) Focus and strong goals. (iv) A motivating environment.

**Self-confidence and Self-efficacy:** a crucial part of being self-motivated is having acceptable levels of self-assurance, self-efficacy and self-confidence. Self-efficacy is defined as a belief in one's capability to succeed and one's ability to achieve the goals set for oneself (AL-Baddareen, Ghaith, & Akour, 2015). High self-efficacy to results in an individual's ability to see tough goals as

a dare, whereas individuals with low self-efficacy would likely see same goals as being away from their capabilities, and might not even try to attain them.

**Positive-thinking and Positive-thinking about The Future:** thinking positively is closely associated with self-confidence as a factor in self-motivation. It is vital to look at everything positively, even when things are going opposite to planned. Thinking positively helps one to think about an attractive future.

**Focus and healthy goals:** setting a goal for oneself gives a clear direction. One goal should be clear, measurable, and specificity and behaviour based not outcomes, Challenge, commitment and regularity of feedback.

**Motivating environment:** the last point is surrounding oneself with the resources and people that remind one of his/her goals.

## 2.1 Related Works

The environment of tertiary education comprises several stakeholders, such as learners, parents, staff, faculty, university administrators, and other contributors (Ferguson, 2017). Each to contribute to the academic success of their learners, thus their semester overall grade point average (GPA). The performance of a student is a key and essential factor in the life of every student and the university administration. Tertiary education demands high academic success of every student; low academic success can lead to academic failure, which can cause one to go on probation, loss a scholarship and grant or expelled from the institution for good. Most students go to universities and colleges, progress through the system, and generally graduate successfully without having a deeper understanding of what it takes to be successful academically and what factors will possibly contribute to that success. Hence, everything that contributes either positively or negatively, to the academic performance of students' needs to be examined.

The following researchers (Bernard & Dzandza, 2018; Maya, 2015; Okyeadie Mensah, Nizam, Mensah, & Nizam, 2016; Owusu-Acheaw & Larson, 2015; Rashid & Asghar, 2016; Salomon & Ben-David, 2016; Sudha & Kavitha, 2016; Wentworth & Middleton, 2014) surveyed how students impact students' academic performance use of social media. Their results revealed that a small negative correlation exists between the amounts of time students spent on social media and their academic performance. Other surveys of social media by (Al-rahmi, 2013; Al-rahmi, Zeki, Alias, & Saged, 2017) effect on academic performance, concluded that social media aids the academic experience of students, but there is a need for them to manage and control their time (Alwagait, Shahzad, & Alim, 2015).

In (AL-Baddareen et al., 2015), the authors examined the effects that achievement goal self-efficacy and metacognition bear on the academic motivation of university students in Jordan. The study outcome revealed that the joint effect of metacognition and mastery goals influence the university student's academic motivation; that is, metacognition and mastery goals can predict academic motivation of university students. The effects of 3 prominent five personality qualities (conscientiousness, agreeableness, and openness to experience) on academic achievement were examined by (Mammadov, Cross, & Ward, 2018) on 161 gifted middle and high school learners. The researchers established that a relationship existed between all these three qualities and academic achievement. In other research work, the link between one's trait of emotional intelligence (TEI) and academic performance (AP) was examined based on perceived social support (PSS), engagement coping (EC), and adjustment, in the environment of the university transition. Their outcome showed that TEI is a direct predictor of higher PSS and the greater use of EC strategies (Perera & DiGiacomo, 2015).

The effect of using motivational regulation strategies on academic procrastination, students' academic performance, and well-being were examined (Grunschel et al., 2016). The extent to which students' AP is affected by their engagement (learners' participation in academic activities and commitment to the school's rules, vision and mission), academic self-efficacy (the learners'

intellect of their own abilities), and academic motivation (the learners' wish in intensification of their academic performance) (Dogan, 2015). The effect of self-efficacy and the big five personality traits on AP was carried out by (Stajkovic, Bandura, Locke, Lee, & Sergent, 2018). An assessment of the influence of learner-faculty communications on student academic motivation was carried out in (Trolan, Jach, Hanson, & Pascarella, 2017). Their outcome revealed that learner-faculty communications greatly influence student's self-motivation. In (Holland et al., 2016), the authors investigated the effect of improved extrinsic motivation on fluent (i.e., efficient and accurate) AP in pediatric medulloblastoma survivors.

The associations between self-regulation of learning, self-efficacy and AP was examined in (Agustiani, Cahyad, & Musa, 2016) among students of faculty of Psychology at Universitas Padjadjaran. Their results revealed a positive relationship exists between self-efficacy, self-regulation of learning and AP of students. Also, in (Schwartz, 2017) an assessment of the effect of the weekly student-led discussions on goal-setting, goal accomplishment, effort, achievement, intrinsic motivation, and satisfaction. In (Cetin, 2015), a survey in the Early Childhood Education Department revealed that there is no correlation between academic self-regulated learning, academic-motivation and students' academic performance. Again (Suprayogi & Valcke, 2014) examined the association between goal orientation, academic efficacy and academic achievement and see which of these, is a crucial predictor of academic achievement.

In (Cigan, 2014), the relationship between the strength of learners' inherent and extrinsic learning-motivation and their socio-demographic characteristics were examined. The relation between self-discipline, self-efficacy and AP was examined in (Jung et al., 2017). In the report presented in (Ferguson, 2017), an assessment of the correlation between AP and mindset, academic motivation, academic Self-Efficacy among Undergraduate communication sciences and disorders students at Andrews University. In (Lee et al., 2014), the researchers examined how personal interest, as a sentimental motivational variable, could predict academic performance and self-regulation, beyond what academic self-efficacy can predict. The link between AP and self-efficacy was examined in (Komarraju & Nadler, 2013). In (Ross, Perkins, & Bodey, 2016) the authors examined the interrelationships between information literacy, self-efficacy and the different types of academic motivation. (Wilson & Narayan, 2016) investigated the associations between self-efficacy, self-regulated learning strategy use and AP.

In (Chang et al., 2014), the authors investigated by what means does Internet self-efficacy aids learners to convert motivation into learning achievement, and its effect on learning performance. In (Zimmerman & Kitsantas, 2014), authors compared learners' self-directive and self-discipline measures and their prediction of AP. Research work by (Rimfeld, Kovas, Dale, & Plomin, 2016) examined how personal pure perseverance and passion for long-term goals and genetics predicts academic achievement. The impact of students' motivation and personality traits on academic performance in blended and online learning environs was investigated in (Alkış & Temizel, 2018). An investigation of the association between achievement-motivation, academic self- concept and academic achievement of high school students was conducted by (Affum-osei, Adom, Barnie, & Forkuoh, 2014). An examination of the connections sandwiched between motivational beliefs (task value, self-efficacy and control of learning beliefs) and use of metacognitive learning schemes amid teacher education students in Uganda was carried out by (Muwonge, Schiefele, Ssenyonga, & Kibedi, 2017). According to reviewed literature, the demographic data, which affect a student's performance, are as shown in Table 1.

S/N	Factor	Reference
1.	Age	(Kieti, 2017; Musso, Kyndt, Cascallar, & Dochy, 2013; Oladokun, Adebajo, & Charles-Owaba, 2008)
2.	Gender	(Chen, Hsieh, & Do, 2014; Musso et al., 2013; Oladokun et al., 2008)
3.	Father's occupation	(Bhardwaj & Pal, 2011; Devasia, Vinushree, & Hegde, 2016; Goga, Kuyoro, & Goga, 2015; Musso et al., 2013; Yadav & Pal, 2012; Yusif, Yussof, & Noor, 2011)

4.	Mother's occupation	(Bhardwaj & Pal, 2011; Devasia et al., 2016; Fleischer, 2015; Goga et al., 2015; Musso et al., 2013; Yadav & Pal, 2012; Yusif et al., 2011)
5.	Ethnicity	Ethnicity (Adejo & Connolly, 2018; Attuquayefio, NiiBoi, & Addo, 2014; Schmitt et al., 2009; Tuen et al., 2019)
6.	Parents' marital status	(Cortez & Silva, 2008; Goga et al., 2015)
7.	Family size	(Agrawal & Mavani, 2015; Bhardwaj & Pal, 2011; Devasia et al., 2016; Goga et al., 2015; Khasanah & Harwati, 2017; Osmanbegovic & Suljic, 2012; Osmanbegović, Suljić, & Agić, 2014; Yadav & Pal, 2012; Yeboah, 2014; Yusif et al., 2011)

**TABLE 1:** Demographic data and students' academic performance.

Given the relationship between one's self-motivation (self-confidence, self-efficacy, positive thinking, focus, reasonable goals, and a motivating environment) and demographic data, it is possible that demographic data will play a role in the effects that self-motivation has on academic performance. However, little is known about how demographic data may influence the relationship between self-motivation and academic performance, especially in developing countries. This paper examined the extent to which the combination of students' demographic data and self-motivation (self-confidence, self-efficacy, individual interest, positive thinking, focus, personal goals, and a motivating environment) predict students' academic performance among Computer Science (CPS) and Information and Communication Technology (ICT) students in computer programming language. Using a hybridised machine learning technique of Random Forest (RF) and Support Vector Machine (SVM) (RFSVM).

### 3. MATERIALS AND METHODS

A quantitative and experimental research approach was adopted for this study. Self-motivation factors priority order questionnaire was adopted for the collection of the research primary data. The first phase of the questionnaire concentrated on getting the demographic data (gender, age, educational background and region of origin). The second part, Self-Motivation had four subsections, which tested each participant's (i) Self-confidence and self-efficacy. (ii) Positive-thinking and positive-thinking about the future. (iii) Focus and reasonable goals (iv) A motivating environment.

#### 3.1 Research Model

Figure 1 shows the research model for this study, showing how the independent variable's demographic data (gender, age, educational background and region of origin) and self-motivation (self-confidence, self-efficacy, individual interest, positive thinking, focus, personal goals, and a motivating environment) joint predict student's score in 3 programming language courses.

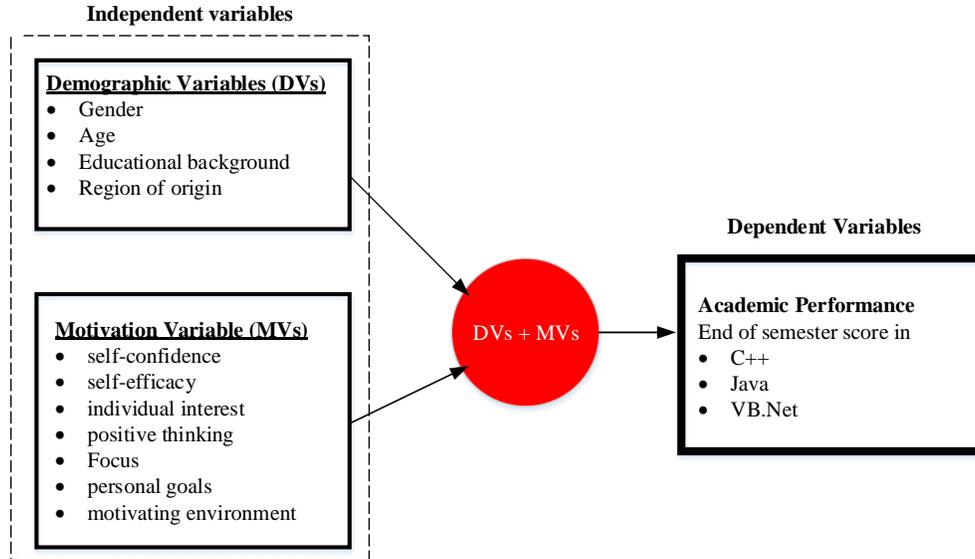


FIGURE 1: Research Model.

### 3.2 Variables

Two variables were considered, Student’s performance (course score) as the dependent variable while demographic data and self-motivation as independent variables. Semester exams score for C++, JAVA and VB.Net were obtained from the Department with the participant’s consent to measure academic performance.

### 3.3 Research Design

This section gives the details of the approaches adopted by this work to analyse the joint effect of students’ self-motivation and demographic data on their academic performance.

#### 3.3.1 Machine Learning Algorithms

Machine-Learning is the science of making computers to learn and behave like humans, and with an autonomous-improvement in their learning over time, by feeding them data and information in the form of observations and real-world interactions (Faggella, 2018). There are various machine-learning algorithms. Nevertheless, the scope of this paper is not to describe the different types of machine-learning algorithm. This section offers a brief description of the two chosen machine-learning algorithms and how they were used in this study.

**Support Vector Machine (SVM):** SVM is a supervised machine learning algorithm for performing regression and classification task (Belgiu & Drăgu, 2016). SVM serves as the linear separator sandwiched between two data-points to detect two different classes in the multidimensional environs (Vaghela, Bhatt, & Patel, 2015). The SVM algorithm is as follows:

Let the dataset for training be denoted by  $\{(x_i, y_i, \dots, (x_n, y_n))\} \in X \times R$

where  $i = 1, 2, 3, \dots, n$

The SVM is optimised by

$$\min_{d, b, \epsilon} \frac{1}{2} W^T W + C \sum_{i=1}^n \epsilon_i \tag{3.1}$$

$$\text{subject to } y_i (W^T \phi(x_i + b)) \geq 1 - \epsilon_i \tag{3.2}$$

$$\epsilon_i > 0$$

With the function  $\phi$ , the vectors  $x_i$  (training dataset) are mapped into a dimension of higher space. At this dimension, the SVM finds a linear separating hyperplane with the best margin. The kernel function was formulated as  $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ . Four underlying kernels are available; however, for this study, the linear given in equation (3.3) was used.

$$\text{Linear: } K(x_i, x_j) = x_i^T x_j \quad (3.3)$$

**Random forest RF:** the RF is a supervised machine learning algorithm, combines the performance of the various decision-tree algorithm (Rodriguez-Galiano, Sanchez-Castillo, Chica-Olmo, & Chica-Rivas, 2015; Wakefield, 2013). The RF accepts an input of  $(u)$ , with  $u$  been a vector consisting of the variable of different evidential features examined for a given training area. Several regression or classification trees (N) are built by the RF and averages their results. Therefore, for N trees,  $\{T(x)\}^N$  the FR predictor is given as:

$$f_{rf}^N(u) = \frac{1}{N} \sum_{N=1}^N T(u) \quad (3.4)$$

The RF is used as a feature selection mechanism to select inputs that are highly correlated students' academic performance to feed the SVM

### 3.3.2 Data Collection and Integration Framework

Figure 2 shows the framework for data collection and integration. The model has four distinct phases, namely:

**Data collection and integration:** 216 records from students, offering CPS and ICT consist of their demographic data (Dataset A), self-motivation answered questionnaires (Dataset B), and past academic records (Dataset C) are integrated at this stage using SQL database server 2014. Thus and Dataset A + Dataset B = Dataset D and Dataset A + Dataset B + Dataset C = Dataset E.

**Data Transformation:** this stage took care of the data pre-processing. At this stage the integrated datasets (D and E) were passed through 3 stages, (i) data cleaning, which includes filling in missing values, smoothing noise, identification and removing of outliers where necessary and resolving data inconsistency. (ii) Data normalisation and aggregation (iii) data reduction, where the volume of data is reduced, but produce same or similar analytical results using feature selection and feature extraction. All this was carried out using python.

**Patterns Extraction:** Python, an open source software, was used to construct the predicting (RFSVM) model.

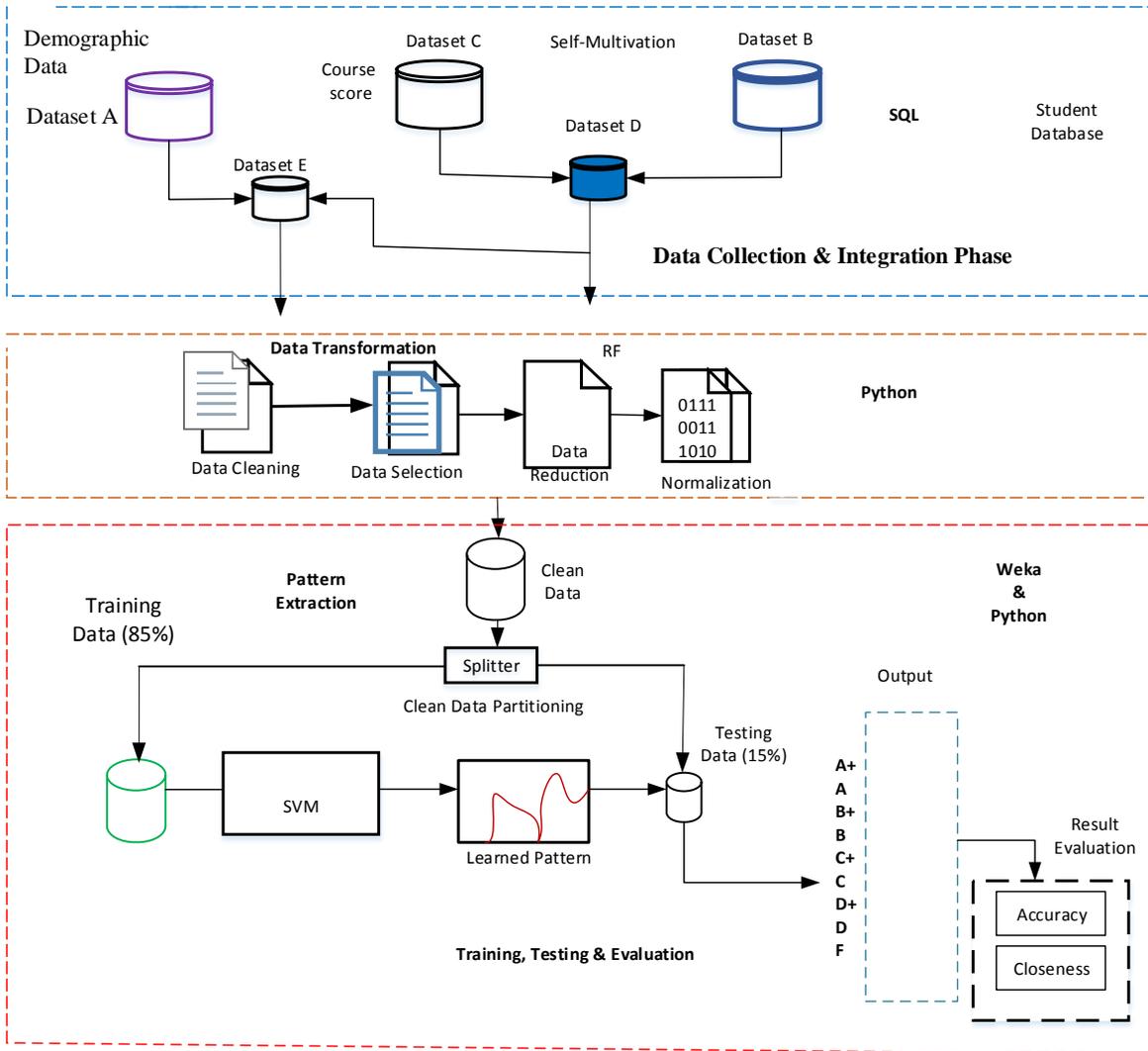


FIGURE 2: Study Framework for Data Collection and Integration.

**Feature Selection**

In a quest to use the feature that has high impact, we calculated feature impact by using feature importance ranking by (Breiman, 2001) to remove features of least importance. A single random forest (RF) was trained using the training dataset (D\_Train) to rank features.

**Feature selection Algorithm**

1. Train an FR using D\_Train (all  $k$  features)
2. Compute average RMSE of model for cross-validation data (C-V)
3. Rank performance by  $VI_k = \sum_{\beta}^{\varphi} \frac{(\epsilon_{\theta k} - \epsilon_{\theta, \pi k})}{\varphi \hat{\sigma}}$
4. **For** each subset of  $k$   $k_i = k - 1, k - 2, k - 3, \dots, 1$  **do**
5.     Train a new forest from  $k_i$  features, highest  $VI$
6.     Calculate the average RMSE of the model on C-V set
7.     Re-rank the features ( $k$ )
8. **End for**
9. Find  $j_j$  with smallest RMSE

### 3.3.3 Data Normalisation

Our selected feature was transformed within the range [0, 1] using equation (3.5) for better performance of the predictive model. Where:  $V'$  is the new scaled value obtained after normalisation;  $v$  the value to be normalised,  $v_{min}$  and  $v_{max}$  are the minimum and maximum value of the dataset.

$$V' = \frac{v - v_{min}}{v_{max} - v_{min}} \quad (3.5)$$

### 3.3.4 Training Techniques

The learning process adopted by this paper was by the supervised process, where the intended input variables (student self-motivation and demographic data) and outputs target (student test score) is entered into the network.

### 3.3.5 Evaluation Performance Metrics

Root mean squared error (RMSE), Mean absolute percentage error (MAPE) and Correlation coefficient (R) as defined in (Javed, Gouriveau, & Zerhouni, 2014; Rajashree, Dash, & Bisoi, 2014) and calculated as in question (3.6), (3.7) and (3.8) respectively, were used as performance metrics measure.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (t_i - y_i)^2} \quad (3.6)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{t_i - y_i}{t_i} \right| \quad (3.7)$$

$$R = \frac{\sum_{i=1}^n (t_i - \bar{t})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (t_i - \bar{t})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.8)$$

## 3.4 Procedure

Self-motivation Scale (SMS) designed by (Mind-Tools, 2018) was adopted for assessing students' self-motivation level. The questionnaire consisted of 4 parts (Self-Confidence & Self-Efficacy, Positive thinking, focus & Strong goals and Motivating environment) with 5 points Likert scale, where participants expressed their consent ranging from not at all to very often. The score for each phase was calculated, and a higher score indicates that a particular self-motivation attribute has a higher level. The overall Self-Motivation Index (SMI) was calculated based on (Mind-Tools, 2018), for this research participant with SMI within (12 – 27), (28 – 43) and (44 – 60) was interpreted as below-average, average and high self-motivation respectively.

## 3.5 Participants

Using purposeful sampling technique, 216 students from the Department Computer Science in Sunyani Technical University in Sunyani, Ghana, offering higher National Diploma (HND) in Information Communication and Technology (ICT) and Computer Science took part in this research. Participants of the current study completed an online questionnaire at their convenience within 30 days. Participants were up-to-date with the purpose of the study, their approval was sought, and participation was voluntary. A pre-test of the questionnaire was performed on a small group of students from Catholic University College in Sunyani to enhance questionnaire significance, simplicity, and specificity. Participant of the pre-test gave a verbal response to researchers expressing the easy in understanding each question and whether each respondent can provide the needed data. Results from the pre-test helped design the questionnaire to suit the expected purpose of the study.

#### 4. Experiments and Analysis of Results

The results obtained by the methodology described in this paper are presented in this section. Statistical tools such mean, standard deviation, variance skewness and kurtosis were applied, and results presented. The dataset for the experiment was divided into two sets, namely: training and testing sets, as shown in Table 2.

Total Dataset	Training Dataset	Testing Dataset
516	85%	15%

TABLE 2: Dataset Partitioning.

#### 4.1 Results and Discussions

The descriptive statistics were performed with Statistical Package for the Social Sciences on the collected data, while Python was used for the implementation of the proposed RFSVM classifier.

#### 4.2 Descriptive Statistics

Table 3 gives the gender analysis of respondents. The results reveal that a high percentage of 63.9% of respondents were males. This difference in gender is because of less female studying science and engineering related courses in higher education.

Gender	Frequency	Percent	Valid Percent	Cumulative Percent
Female	78	36.1	36.1	36.1
Male	138	63.9	63.9	100.0
Total	216	100.0	100.0	

TABLE 3: Gender Analysis of Respondents.

The age distribution of respondents is, as shown in Table 4. A higher percentage (83.3%) of the respondents were with the age bracket of (21 – 25), 11.1% with the brackets of (26 – 30) and 5.6% in the brackets of (15 – 20).

Age Group	Frequency	Percent	Valid Percent	Cumulative Percent
15 - 20	12	5.6	5.6	5.6
21 - 25	180	83.3	83.3	88.9
26 - 30	24	11.1	11.1	100.0
Total	216	100.0	100.0	

TABLE 4: Age Analysis of Respondents.

Table 5 shows the statistical results of the response obtained from respondents. The results reveal that the data distribution is moderately skewed. Again, the computed values of Kurtosis reveals that the data distribution is slightly leptokurtic and platykurtic.

Features	Statistic				Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	N	Range	Min.	Max.	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Age	216	11	15	26	21.22	.148	2.180	4.750	-.125	.166	3.421	.330
Focus	216	8	2	10	6.73	.121	1.772	3.139	-.304	.166	-.371	.330
Gender	216	1	0	1	.64	.033	.481	.232	-.582	.166	-	.330
Motivating	216	4	1	5	3.25	.068	1.000	1.000	.203	.166	1.676	.330
Positive	216	11	9	20	14.40	.174	2.555	6.529	-.021	.166	-.610	.330
Self-Efficacy	216	11	9	20	14.42	.196	2.881	8.300	.338	.166	-.558	.330
Thinking	216	6	4	10	7.80	.110	1.623	2.635	-.458	.166	-.602	.330
Valid N (listwise)	216										-.644	.330

**TABLE 5:** Descriptive Statistics of Respondents.

#### 4.2.1 Feature Ranking and Importance

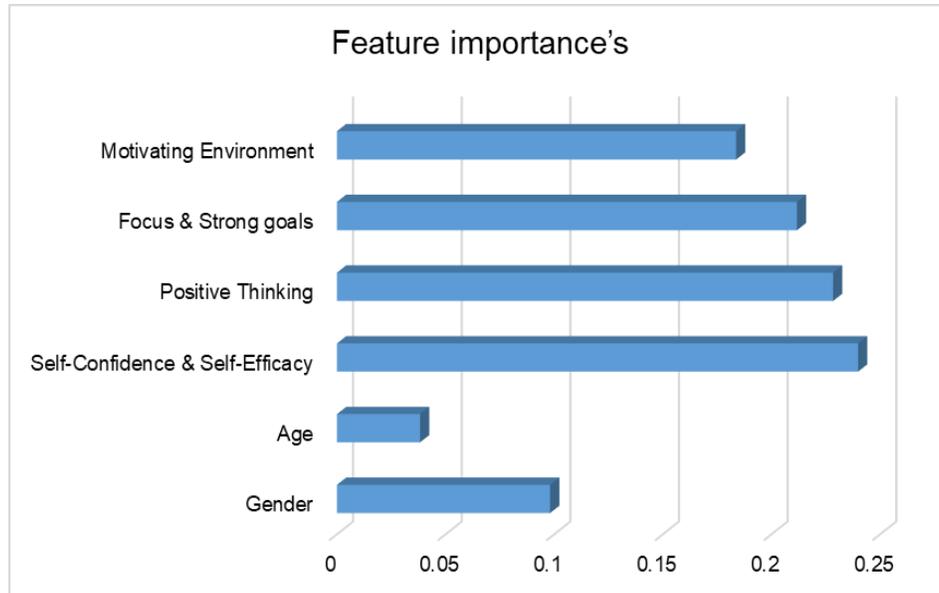
Table 6 and figure 3 shows the independent's input variables rankings and importance offered by the random forest algorithm. From figure 3, Self-Confidence & Self-Efficacy is the highest predictor of students' academic performance in programming courses with an importance score of (0.239825), which disagrees with (Talsma et al., 2018) arguments that self-efficacy is simply a mirror image of past performance, however, agrees with (AL-Baddareen et al., 2015) studies. Positive thinking was second with (0.228370) importance score. Focus and reasonable goals with an importance score of (0.211707) followed by Motivating Environment with an importance score of (0.183711).

The study revealed that demography data gender had a better score of importance (0.098140) which confirms (Chen et al., 2014; Musso et al., 2013; Oladokun et al., 2008) studies, as compared with age, which had little significance score of (0.038248), nevertheless, there is a relation with age and academic performance as argued by (Kieti, 2017; Musso et al., 2013; Oladokun et al., 2008).

S/N	Feature	Feature importance's
0	Gender	0.098140
1	Age	0.038248
2	Self-Confidence & Self-Efficacy	0.239825
3	Positive Thinking	0.228370
4	Focus & Strong goals	0.211707
5	Motivating Environment	0.183711

**TABLE 3:** Feature ranking by RF.

The result also revealed that is better to inculcate desirable and high personality traits such (self-confidence & self-efficacy, positive thinking and focus & strong goals) in students within the school environments, than instead creating a motivation-motivating environment. That is building these traits in them enable them to perform well academically than building a motivating environment around them, which confirms (Dweck et al., 2014) assessment. The age factor cannot be left out; the results reveal that age plays a little factor of importance in learning programming. Hence, programming as a course should be introduced early within the first year or during the first semester of the second year. Since students within the ages of 15 – 25 performed better than 26 – 30.

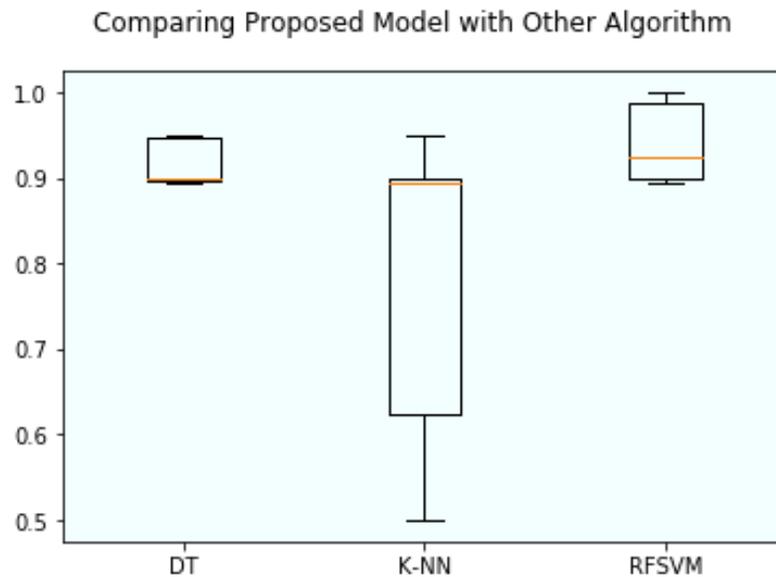


**FIGURE 3:** Feature Ranking and Importance.

However, the outcome clearly showed that self-motivation indication factors are of higher importance in predicting students' academic performance, and confirms studies by (Dweck et al., 2014; Lee et al., 2014). That is, for a student to be better academically in programming courses, such students must exhibit a high level of self-motivation. To measure how efficient is our proposed model, we benchmarked our RFSVM model with other known machine-learning algorithms such as Decision Tree (DT) and K-Nearest Neighbour and Table 7, shows the RMSE, MAE, Accuracy and  $R^2$  values. The proposed model obtained an accuracy of approximately 98% as compared with 95% for DT, 0.36% for K-NN. The error between actual values and predicted values of our proposed model was less (RMSE = 0.326401, MAE = 0.050909) and compared with the K-NN (RMSE = 2.671397, MAE = 1.954545) and DT models (RMSE = 0.426401, MAE = 0.090909).

Models	RMSE	MAE	R <sup>2</sup>	Accuracy
DT	0.426401	0.090909	0.944724	0.954545
K-NN	2.671397	1.954545	-1.169598	0.363636
RFSVM	0.326401	0.050909	0.974724	0.979945

**TABLE 4:** Accuracy Metric Comparison.



**FIGURE 4:** Boxplot Comparison of Models.

The outcome of the boxplot shown in figure 4, shows that the RFSVM model is the most suitable for students' academic performance prediction based on self-motivation and demographic data.

## 5. CONCLUSIONS

This paper proposed a hybridised machine learning technique to predict the academic performance of students in computer programming courses based on their self-motivation and demography data. Like all other studies on students' academic performance, we sought in identifying weak learners in computer programming courses at the earliest stage of their academic ladder, so that the necessary help can be offered to them by their respective academic counsellor and school management. The proposed model was compared with other machine learning algorithm base on model prediction efficiency, RMSE and MAE. All machine learning models (Decision Tree (DT), K-Nearest Neighbour (K-NN) and RFSVM) used for the study gained satisfactory results. However, the result revealed that the proposed hybrid model (RFSVM) performed better than DT and K-NN.

The proposed model yield and overall accuracy measure of 97.9945% and gave a better error value as compared with DT and K-NN. The proposed model was able to identify almost all students with self-motivation index of below average. The outcome confirms an agreement with past research work, that machine-learning techniques have the power to produce an efficient model for predicting students' academic performance.

Also, the study revealed that personality traits such as (Self-Confidence & Self-Efficacy, Positive Thinking and Focus & Strong goals) of a student has a higher impact on their academic as far as computer programming is concerned. Students with higher personal traits turn to perform well in computer programming. That is building these personality traits in them offer better academic performance than building a motivating environment around them. Age was observed to pose a little significance on the academic performance of students in programming; hence, it is concluded that programming should be introduced at the early stage of the educational ladder. If not primary education, secondary at most is better.

The weaker learners identified were directed to the school counselling units to work on their self-motivation area that is not strong. Though their performance afterwards has not yet been verified, it is believed that some improvement would be realised. With this outcome, we propose the

incorporation of building a better self-motivation into students' academic calendar, which we believe will help them academically. It is hoped that the proposed model is implemented in academic institutions to help promote academic excellence in programming languages.

## 6. LIMITATIONS AND FUTURE RESEARCH

The outcomes of the present study are limited to only the experimental sample, which should, in future, be extended across regions to make the outcome globally applicable. Future research work could feather include more demographic data such a father's occupation, mother's occupation, ethnicity, parents' marital status and family size as input features to impact prediction accuracy.

## 7. ACKNOWLEDGEMENT

Glory be to God Almighty for bringing us this far. We also express our sincere thanks and acknowledged the contribution and assistance of all staffs of Computer Science and Electrical & Electronic Engineering Department of STU.

Financial and Ethical Disclosures

**Funding:** Authors did not receive any fund anywhere.

**Conflict of Interest:** The authors of the current study declare that they have no conflict of interest.

## 8. REFERENCES

- [1] Adejo, O. W., & Connolly, T. (2018). Predicting student academic performance using multi-model heterogeneous ensemble approach. *Journal of Applied Research in Higher Education*, 10(1), 61–75. <https://doi.org/10.1108/JARHE-09-2017-0113>.
- [2] Affum-osei, E., Adom, E. A., Barnie, J., & Forkuoh, S. K. (2014). Achievement motivation, academic self-concept and academic achievement among high school students. *European Journal of Research and Reflection in Educational Sciences*, 2(2), 24–37. <https://doi.org/10.1016/B978-0-08-097086-8.92153-6>.
- [3] Agrawal, H., & Mavani, H. (2015). 10 Student Performance Prediction using Machine Learning, 4(03), 111–113.
- [4] Agustiani, H., Cahyad, S., & Musa, M. (2016). Self-efficacy and Self-Regulated Learning as Predictors of Students Academic Performance. *The Open Psychology Journal*, 9(1), 1–6. <https://doi.org/10.2174/1874350101609010001>.
- [5] AL-Baddareen, G., Ghaith, S., & Akour, M. (2015). Self-Efficacy, Achievement Goals, and Metacognition as Predicators of Academic Motivation. *Procedia - Social and Behavioral Sciences*, 191, 2068–2073. <https://doi.org/10.1016/j.sbspro.2015.04.345>.
- [6] Al-rahmi, W. M. (2013). The Impact of Social Media use on Academic Performance among university students: A Pilot Study. *Journal Of Information Systems Research and Innovation*, 1–10. <https://doi.org/http://seminar.utmspace.edu.my/jisri/>.
- [7] Al-rahmi, W. M., Zeki, A. M., Alias, N., & Saged, A. A. (2017). Social Media and its Impact on Academic Performance among University Students. *The Anthropologist*, 28(1–2), 52–68. <https://doi.org/10.1080/09720073.2017.1317962>.
- [8] Alkış, N., & Temizel, T. T. (2018). The Impact of Motivation and Personality on Academic Performance in Online and Blended Learning Environments. *Journal of Educational Technology & Society*, 21(3), 35–47.
- [9] Alwagait, E., Shahzad, B., & Alim, S. (2015). Impact of social media usage on students academic performance in Saudi Arabia. *Computers in Human Behavior*.

<https://doi.org/10.1016/j.chb.2014.09.028>.

- [10] Attuquayefio, NiiBoi, S., & Addo, H. (2014). Using the UTAUT model to analyze students' ICT adoption. *International Journal of Education & Development Using Information & Communication Technology*, 10(3), 75–86. Retrieved from <http://ezproxy.usq.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eht&AN=97923459&site=ehost-live>.
- [11] Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- [12] Bernard, K. J., & Dzandza, P. E. (2018). Effect of social media on academic performance of students in Ghanaian Universities: A case study of University of Ghana, Legon. *Library Philosophy and Practice*, 2018(February).
- [13] Bhardwaj, B. K., & Pal, S. (2011). Data Mining : A prediction for performance improvement using classification. *International Journal of Computer Science and Information Security*, 9(4), 136–140.
- [14] Bingham, T., & Conner, M. (2010). *The New Social Learning: A Guide to Transforming Organizations through Social Media*. Berrett-Koehler Store, San Francisco, CA.
- [15] Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- [16] Cetin, B. (2015). Academic Motivation And Self-Regulated Learning In Predicting Academic Achievement in College. *Journal of International Education Research (JIER)*, 11(2), 95–106. <https://doi.org/10.19030/jier.v11i2.9190>.
- [17] Chang, C. S., Liu, E. Z. F., Sung, H. Y., Lin, C. H., Chen, N. S., & Cheng, S. S. (2014). Effects of online college student's Internet self-efficacy on learning motivation and performance. *Innovations in Education and Teaching International*, 51(4), 366–377. <https://doi.org/10.1080/14703297.2013.771429>.
- [18] Chen, J. F., Hsieh, H. N., & Do, Q. H. (2014). Predicting student academic performance: A comparison of two meta-heuristic algorithms inspired by cuckoo birds for training neural networks. *Algorithms*, 7(4), 538–553. <https://doi.org/10.3390/a7040538>.
- [19] Cigan, V. (2014). Relationship between students' motivation and their socio-demographic characteristics. *Linguistica*, 54(1), 11–30. <https://doi.org/10.4312/linguistica.54.1.11-30>.
- [20] Cortez, P., & Silva, A. (2008). Using Data Mining To Predict Secondary School Student Performance. *5th Annual Future Business Technology Conference*, 2003(2000), 5–12. <https://doi.org/10.13140/RG.2.1.1465.8328>.
- [21] Devasia, T., Vinushree, T. P., & Hegde, V. (2016). Prediction of Students Performance using Educational Data Mining. 16.
- [22] Dogan, U. (2015). Student engagement, academic self-efficacy, and academic motivation as predictors of academic performance. *Anthropologist*, 20(3), 553–561. <https://doi.org/10.1080/09720073.2015.11891759>.
- [23] Dweck, S. carol, Walton, M. G., & Cohen, L. G. (2014). *Academic Tenacity*. United States and. Retrieved from [www.gatesfoundation.org](http://www.gatesfoundation.org).
- [24] Faggella, D. (2018). What is Machine Learning? Retrieved October 8, 2018, from <https://www.techemergence.com/what-is-machine-learning/>.

- [25] Ferguson, L. H. (2017). Mindset, Academic Motivation, And Academic Self-Efficacy As Correlates Of Academic Achievement Among Undergraduate Students in Communication Sciences and Andrews University. Retrieved from <https://digitalcommons.andrews.edu/dissertations/1648/>.
- [26] Fleischer, J. E. (2015). Information Communication Technology Usage Patterns in Second Cycle Schools: A Study of Two Selected Senior High Schools in Ghana.
- [27] Goga, M., Kuyoro, S., & Goga, N. (2015). A recommender for improving the student academic performance. *Procedia - Social and Behavioral Sciences*, 180(November 2014), 1481–1488. <https://doi.org/10.1016/j.sbspro.2015.02.296>.
- [28] Grunschel, C., Schwinger, M., Steinmayr, R., & Fries, S. (2016). Effects of using motivational regulation strategies on students' academic procrastination, academic performance, and well-being. *Learning and Individual Differences*, 49, 162–170. <https://doi.org/10.1016/j.lindif.2016.06.008>.
- [29] Holland, A. A., Hughes, C. W., Harder, L., Silver, C., Bowers, D. C., & Stavinoha, P. L. (2016). Effect of motivation on academic fluency performance in survivors of pediatric medulloblastoma. *Child Neuropsychology*, 22(5), 570–586. <https://doi.org/10.1080/09297049.2015.1023272>.
- [30] Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, 17, 63–84. <https://doi.org/10.1016/j.edurev.2015.11.002>.
- [31] Javed, K., Gouriveau, R., & Zerhouni, N. (2014). SW-ELM: A summation wavelet extreme learning machine algorithm with a priori parameter initialization. *Neurocomputing*, 123, 299–307. <https://doi.org/10.1016/j.neucom.2013.07.021>.
- [32] Jung, K. R., Zhou, A. Q., & Lee, R. M. (2017). Self-efficacy, self-discipline and academic performance: Testing a context-specific mediation model. *Learning and Individual Differences*, 60(October), 33–39. <https://doi.org/10.1016/j.lindif.2017.10.004>.
- [33] Khasanah, A. U., & Harwati. (2017). A Comparative Study to Predict Student's Performance Using Educational Data Mining Techniques. *IOP Conference Series: Materials Science and Engineering*, 215(1). <https://doi.org/10.1088/1757-899X/215/1/012036>.
- [34] Kieti, J. M. (2017). An Investigation into Factors Influencing Students' Academic Performance in Public Secondary Schools in Matungulu Sub-County, Machakos County. South Eastern Kenya University. <https://doi.org/10.1111/j.1469-7610.2010.02280.x>.
- [35] Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences*, 25, 67–72. <https://doi.org/10.1016/j.lindif.2013.01.005>.
- [36] Lee, W., Lee, M. J., & Bong, M. (2014). Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemporary Educational Psychology*, 39(2), 86–99. <https://doi.org/10.1016/j.cedpsych.2014.02.002>.
- [37] Mammadov, S., Cross, T. L., & Ward, T. J. (2018). The Big Five personality predictors of academic achievement in gifted students: Mediation by self-regulatory efficacy and academic motivation. *High Ability Studies*, 00(00), 1–23. <https://doi.org/10.1080/13598139.2018.1489222>.
- [38] Maya, k. G. (2015). Achievement scripts, media influences on Blacks students' academic performance, self-perceptions and carrier interests. *Journal of Black Psychology*, 42(3), 195–220. <https://doi.org/10.1177/0095798414566510>.

- [39] Mind-Tools. (2018). How Self-Motivated Are You: Taking Charge of Your Goals and Achievements. Retrieved from [https://www.mindtools.com/pages/article/newLDR\\_57.htm](https://www.mindtools.com/pages/article/newLDR_57.htm).
- [40] Musso, M. F., Kyndt, E., Cascallar, E. C., & Dochy, F. (2013). Predicting general academic performance and identifying the differential contribution of participating variables using artificial neural networks. *Frontline Learning Research*, 1(1), 42–71. <https://doi.org/10.14786/flr.v1i1.13>.
- [41] Muwonge, C. M., Schiefele, U., Ssenyonga, J., & Kibedi, H. (2017). Self-regulated learning among teacher education students: Motivational beliefs influence on the use of metacognition. *Journal of Psychology in Africa*, 27(6), 515–521. <https://doi.org/10.1080/14330237.2017.1399973>.
- [42] Okyeadie Mensah, S., Nizam, D. I., Mensah, O. S., & Nizam, I. (2016). the Impact of Social Media on Students' Academic. *International Journal Of Education, Learning & Training (IJELT)*, 1(1), 14–21. <https://doi.org/10.24924/ijelt/2016.11/v1.iss1/14.21>.
- [43] Oladokun, V. O., Adebajo, A. T., & Charles-Owaba, O. E. (2008). Predicting students' academic performance using artificial neural network: A case study of an engineering course. *The Pacific Journal of Science and Technology*, 9(1), 72–79.
- [44] Osmanbegovic, E., & Suljic, M. (2012). Data mining approach for predicting student performance. *Journal of Economics and Business*, X(1), 3–12.
- [45] Osmanbegović, E., Suljić, M., & Agić, H. (2014). Determining Dominant Factor for Students Performance Prediction by Using Data Mining. *Vitez-Tuzla-Zagreb-Beograd-Bucharest, XVII(34)*, 147–158.
- [46] Owusu-Acheaw, M., & Larson, A. G. (2015). Use of Social Media and its Impact on Academic Performance of Tertiary Institution Students: A Study of Students of Koforidua Polytechnic, Ghana. *Journal of Education and Practice*, 6(6), 94–101.
- [47] Perera, H. N., & DiGiacomo, M. (2015). The role of trait emotional intelligence in academic performance during the university transition: An integrative model of mediation via social support, coping, and adjustment. *Personality and Individual Differences*, 83, 208–213. <https://doi.org/10.1016/j.paid.2015.04.001>.
- [48] Rajashree, D., Dash, P. K., & Bisoi, R. (2014). A self adaptive differential harmony search based optimized extreme learning machine for financial time series prediction. *Swarm and Evolutionary Computation*, 19, 25–42. <https://doi.org/10.1016/j.swevo.2014.07.003>.
- [49] Rashid, T., & Asghar, H. M. (2016). Technology use, self-directed learning, student engagement and academic performance: Examining the interrelations. *Computers in Human Behavior*, 63, 604–612. <https://doi.org/https://doi.org/10.1016/j.chb.2016.05.084>.
- [50] Rimfeld, K., Kovas, Y., Dale, P. S., & Plomin, R. (2016). True grit and genetics: Predicting academic achievement from personality. *Journal of Personality and Social Psychology*, 111(5), 780–789. <https://doi.org/10.1037/pspp0000089>.
- [51] Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, 71, 804–818. <https://doi.org/10.1016/j.oregeorev.2015.01.001>.
- [52] Ross, M., Perkins, H., & Bodey, K. (2016). Academic motivation and information literacy self-efficacy: The importance of a simple desire to know. *Library and Information Science Research*, 38(1), 2–9. <https://doi.org/10.1016/j.lisr.2016.01.002>.

- [53] Salomon, A., & Ben-David, K. Y. (2016). High-school students' perceptions of the effects of non-academic usage of ICT on their academic achievements. *Computers in Human Behavior*, 64, 143–151. <https://doi.org/https://doi.org/10.1016/j.chb.2016.06.024>.
- [54] Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., & Zorzie, M. (2009). Prediction of 4-Year College Student Performance Using Cognitive and Noncognitive Predictors and the Impact on Demographic Status of Admitted Students. *Journal of Applied Psychology*, 94(6), 1479–1497. <https://doi.org/10.1037/a0016810>.
- [55] Schwartz, T. D. B. (2017). The Effect of Student-Led Conferencing at School and at Home on Goal-Setting , Goal-Fulfillment , Effort , Achievement , Intrinsic Motivation , and Satisfaction for Montessori Lower Elementary 3rd Year Students. The St. Catherine University. Retrieved from <https://sophia.stkate.edu/maed/224>.
- [56] Stajkovic, A. D., Bandura, A., Locke, E. A., Lee, D., & Sergent, K. (2018). Test of three conceptual models of influence of the big five personality traits and self-efficacy on academic performance: A meta-analytic path-analysis. *Personality and Individual Differences*, 120(February 2017), 238–245. <https://doi.org/10.1016/j.paid.2017.08.014>.
- [57] Sudha, S., & Kavitha, E. S. (2016). The Effect of Social Networking on Students ' Academic Performance : the Perspective of Faculty Members of Periyar. *Library Philophy and Practice*, 14(55). Retrieved from <http://digitalcommons.unl.edu/libphilprac%0Ahttp://digitalcommons.unl.edu/libphilprac/1455>.
- [58] Suprayogi, M. N., & Valcke, M. (2014). Academic Efficacy, Goal Orientation, and Academic Achievement; A Comparative Study between Belgium and Indonesian Students.
- [59] Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences*, 61(October 2017), 136–150. <https://doi.org/10.1016/j.lindif.2017.11.015>.
- [60] Trolian, T. L., Jach, E. A., Hanson, J. M., & Pascarella, E. T. (2017). Influencing Academic Motivation : The Effects of Student – Faculty Interaction Influencing Academic Motivation : The Effects of Student – Faculty Interaction. *Journal of College Student Development*, 57(7), 810–826.
- [61] Tuen, W., Leung, V., Yee, T., Pan, W., Wu, C., Lung, S. C., & Spengler, J. D. (2019). Landscape and Urban Planning How is environmental greenness related to students ' academic performance in English and Mathematics ? *Landscape and Urban Planning*, 181(1), 118–124. <https://doi.org/10.1016/j.landurbplan.2018.09.021>.
- [62] Vaghela, C., Bhatt, N., & Patel, P. U. (2015). A Survey on Various Classification Techniques for Clinical Decision Support System. *International Journal of Computer Applications*, 116(23), 975–8887.
- [63] Wakefield, K. (2013). A guide to machine learning algorithms and their applications. Retrieved August 1, 2018, from [https://www.sas.com/en\\_gb/insights/articles/analytics/machine-learning-algorithms.html](https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html).
- [64] Wentworth, D. K., & Middleton, J. H. (2014). Technology use and academic performance. *Computers and Education*, 78(September 2014), 306–311. <https://doi.org/10.1016/j.compedu.2014.06.012>.
- [65] Wilson, K., & Narayan, A. (2016). Relationships among individual task self-efficacy, self-regulated learning strategy use and academic performance in a computer-supported collaborative learning environment. *Educational Psychology*, 36(2), 236–253.

<https://doi.org/10.1080/01443410.2014.926312>.

- [66] Yadav, S. K., & Pal, S. (2012). Data Mining : A Prediction for Performance Improvement of Engineering Students using Classification. *World of Computer Science and Information Technology Journal WCSIT*, 2(2), 51–56. [https://doi.org/10.1142/9789812771728\\_0012](https://doi.org/10.1142/9789812771728_0012).
- [67] Yeboah, Y. K. (2014). Investigating the Low Performance of Students' English in the Basic Education Certificate Examination in the Sunyani Municipality. UNIVERSITY OF GHANA, LEGON. <https://doi.org/10.1038/253004b0>.
- [68] Yusif, H. M., Yussof, I., & Noor, A. H. S. M. (2011). Determinants of Students Academic Performance in Senior High Schools: a Binary Logit Approach, 31(3), 107–117. <https://doi.org/10.4314/just.v31i3.12>.
- [69] Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology*, 39(2), 145–155. <https://doi.org/10.1016/j.cedpsych.2014.03.004>.