LEARNER ANALYTICS DASHBOARD FOR PREDICTING STUDENTS PERFORMANCE: A CASE STUDY OF TECHNICAL INSTITUTIONS

By

CHEGE LILIAN MURINGI

Master of Science in Data Analytics

KCA UNIVERSITY

2016

FACULTY OF COMPUTING & INFORMATION MANAGEMENT

RESEARCH PROJECT

ON

LEARNER ANALYTICS DASHBOARD FOR PREDICTING STUDENTS PERFORMANCE: A CASE STUDY OF TECHNICAL INSTITUTIONS

By

CHEGE LILIAN MURINGI

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE IN DATA ANALYTICS IN THE FACULTY OF COMPUTING AND INFORMATION MANAGEMENT AT KCA UNIVERSITY

NOVEMBER, 2016

DECLARATION

I Lilian Muringi Chege, the undersigned, hereby declare that this research proposal is my original work and that it has not been presented to any other University, college or institution for higher learning other than KCA University.

………………………………………………….. ………………………………

Lilian Muringi Chege Date

Student No: KCA/14/04634

This project has been presented for examination with my approval as the appointed supervisor.

Signed ………………………………………… Date ………………………..

Prof. Ddembe Williams (Supervisor)

Faculty of Computing & Information Management KCA University

Signed ………………………………………… Date ………………………..

Mr. Peter Kiarie (Supervisor)

Faculty of Computing & Information Management KCA University

Signed ……………………………………….… Date ………………………...

Dean,

Faculty of Computing & Information Management KCA University

ABSTRACT

Learning analytics is the interpretation of a wide range of data produced by and gathered on behalf of students in order to assess academic progress, predict future performance and spot potential issues. It emphasizes that institutions should undertake and understand data measurement and collection activities and focus on the analysis and reporting of the data. The information can be displayed via a dashboard which is a visual display of the most important information needed to achieve one or more objectives. Student achievement is the main agenda of any learning institution. The school administration has a responsibility to ensure students learn what they need to know depending on the course they are undertaking to prepare them for work. This requires access to a wide range of data and tools to continuously monitor student learning and performance. The need arises to manage and integrate data into a dashboard for effective decision making on matters concerning the student's performance. Therefore, the aim of the study is to develop a learner analytics dashboard for predicting student performances: a case study of technical institutions. The objectives of the study are to: establish the current methods of predicting performance in technical institutions; determine the requirements for the dashboard system; design a learner analytics dashboard; implement the learner analytics dashboard system for use in technical institutions; test and validate the learner analytics dashboard.

DEDICATION

To my sons' Lenny and Levin.

ACKNOWLEDGEMENTS

My sincere gratitude goes to my supervisors Prof. Ddembe Williams and Mr. Peter Kiarie who have reviewed my work, edited it and kept me on the right track with their intellectual guidance, advices and communications.

I am also grateful to my colleagues for their support, company, sharing of ideas and resources.

May God Bless You All.

Table of contents

List of Figures

List of Tables

1 Chapter One: Introduction

According to Shahiri et.al (2015), student's performance is an essential part in higher learning institutions. This is because one of the criteria for a high quality university is based on its excellent record of academic achievement. Final grades are used to evaluate student's performance. They could be based on course structure, assessment mark, final exam score and also extracurricular activities. The evaluation is important to maintain student's performances and the effectiveness of learning process. By analyzing student's performance, a strategic program can be well planned during their period of studies in an institution.

According to Stapel et.al (2016), performance prediction is one cornerstone of a fully personalized learning environment and also an important component of the efforts to deliver quality education. Higher education institutes, for example, are striving to incorporate predictive elements into their educational processes to better support students.

1.1. Learning Analytics

According to Bienkowski & et.al (2012), learning analytics as defined by Johnson et al. (2011), is the interpretation of a wide range of data produced by and gathered on behalf of students in order to assess academic progress, predict future performance and spot potential issues. Data are collected from explicit student actions, such as completing assignments and taking exams, and from tacit actions, including online social interactions, extracurricular activities, posts on discussion forums, and other activities that are not directly assessed as part of the student's educational progress. The goal is to enable teachers and schools to tailor educational opportunities to each student level of need and ability. Learning analytics emphasizes that institutions should undertake and understand data measurement and collection activities and focus on the analysis and reporting of the data. It addresses application of known methods and models to answer important questions that affect student learning and organizational learning systems.

In summary, learning analytics systems apply models to answer such questions as: When are students ready to move on to the next topic? ; When are students falling behind in a course?; When is a student at risk for not completing a course?; What grade is a student likely to get without intervention?; What is the best next course for a given student?; Should a student be referred to a counselor for help?

Learning analytics can be both descriptive and predictive. From a descriptive perspective, learning analytics can help us answer such questions as: "What happened?", "Where was the problem?", and "What actions are needed?" Learning analytics can also help us to predict and prescribe by answering such questions as: "Why is this happening?" "What if these trends continue?", "What will happen next?", and "What is the best that can happen?" (Dietz-Uhler & Hurn, 2013)

1.2. Benefits of Using Learning Analytics for Higher Education.

At administrative level, learning analytics improve decision-making and informing resource allocation, highlighting an institution's successes and challenges, and increasing organizational productivity. From the perspective of a faculty member, learning analytics can help faculty identify at-risk learners and provide interventions, transform pedagogical approaches, and help students gain insight into their own learning. Having data at hand and knowing what to do with it can allow us to realize these benefits, (Dietz-Uhler & Hurn, 2013).

1.2.1. Goals of Learning Analytics

According to Dietz-Uhler & Hurn (2013), the following are some of the reasons for learning analytics:

- Increased accountability in all levels of education. Educational institutions around the country are feeling increased pressure to account for what and how their students are learning.
- Learning analytics provides one of many methods to not only document student performance but also to provide tools that encourage the types of continuous improvement that accrediting bodies are seeking.
- Institutions of higher education are experiencing greater demands to retain students. Learning analytics can assist with this goal by providing a more personalized learning experience through the use of data to respond to students' needs. These kinds of personalization will likely lead to greater success in the classroom.
- Predicting learner performance can suggest to learners relevant learning resources, increased reflection and awareness on the part of the learner, detection of undesirable learning behaviors, and detecting affective states e.g., boredom, frustration of the learner.

1.3. Business Intelligence

According to Rouse (2015), Business intelligence (BI) is a technology-driven process for analyzing data and presenting actionable information to help corporate executives, business managers and other end users make more informed business decisions. BI encompasses a variety of tools, applications and methodologies that enable organizations to collect data from internal systems and external sources, prepare it for analysis, develop and run queries against the data, and create reports, dashboards and data visualizations to make the analytical results available to corporate decision makers as well as operational workers.

1.4. Dashboards

According to Few (2006), a dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance. Just as the dashboard of a car provides critical information needed to operate the vehicle at a glance, a BI dashboard serves a similar purpose, whether you're using it to make strategic decisions for a huge corporation, run the daily operations of a team, or perform tasks that involve no one but yourself. The means is a single-screen display, and the purpose is to efficiently monitor the information needed to achieve one's objectives.

The information on a dashboard is presented visually, usually as a combination of text and graphics, but with an emphasis on graphics. Dashboards are highly graphical because graphical presentation often communicates with greater efficiency and richer meaning than text alone.

Dashboards display the information needed to achieve specific objectives. To achieve even a single objective often requires access to a collection of information that is not otherwise related, often coming from diverse sources related to various business functions. It is not a specific type of information, but information of whatever type that is needed to do a job. It is not just information that is needed by executives or even by managers; it can be information that is needed by anyone who has objectives to meet.

Despite the fact that information about almost anything can be appropriately displayed in a dashboard, there is at least one characteristic that describes almost all the information found in dashboards: it is abbreviated in the form of summaries or exceptions. This is because you cannot monitor at a glance all the details needed to achieve your objectives. A dashboard must be able to quickly point out that something deserves your attention and might require action. It need not provide all the details necessary to take action, but if it does not, it ought to make it as easy and seamless as possible to get to that information. Getting there might involve shifting to a different display beyond the dashboard, using navigational methods such as drilling down. The dashboard does its primary job if it tells you with no more than a glance that you should act. It serves you superbly if it directly opens the door to any additional information that you need to take that action.

Dashboards and visualization are cognitive tools that improve your "span of control" over a lot of business data. These tools help people visually identify trends, patterns and anomalies, reason about what they see and help guide them toward effective decisions. As such, these tools need to leverage people's visual capabilities. With the prevalence of scorecards, dashboards and other visualization tools now widely available for business users to review their data, the issue of visual information design is more important than ever.

1.5. Types of Dashboards

According to Stanley (2011), there are three types of dashboards:

Strategic dashboards provide managers and executives at all levels of the organization the information they need understand the health of the organization and help identify potential opportunities for expansion and improvement. Strategic dashboards do not provide all the detailed information needed to make complex decisions, but instead help executives identify opportunities for further analysis.

Analytical dashboards provide users with the data they need to understand trends and why certain things are happening by making comparisons across time and multiple variables. Analytical dashboards often contain more information per square inch than both strategic and operational dashboards. Since understanding is the goal analytical dashboards can be more complex than strategic or operations dashboards. Also, while analytical dashboards should facilitate interactions with the data, including viewing the data in increasing detail, it is important to maintain the ability to compare data across time and multiple variables. If you lose the ability to compare data then an analytical dashboards is no longer able to accomplish the goal of allowing users to understand trends and why things are happening.

Operational dashboards are used to monitor real time operations and alert the users to deviations for the norm. This often means that operational dashboards need to be updated frequently if not in real time, contain less information than analytical or strategic dashboards, and make it nearly impossible to avoid or misunderstand an alert when something deviates from the acceptable standards. Operational dashboards should provide users with specific alerts and provide them with exactly what information they need to quickly get operations back to normal.

N/B: The dashboard being developed will be both analytical and operational as it addresses performance issues.

1.6. Information Visualization

Dashboards convey information through a process referred to as visualization. Information visualization refers to the "use of interactive visual representations of abstract, non physically based data to amplify cognition". Information visualization is a visual user interface to information, the goal providing users with information insight (Velcu-Laitinen & Yigitbasioglu, 2012).

1.7. Technical Vocational Education and Training Authority

Training in technical institutions is controlled by Technical and Vocational Education and Training Authority. According to tvetact_no29 of 2013, some of the functions of the authority are to: regulate and coordinate training under this Act; accredit and inspect programmes and courses; determine the national technical and vocational training objectives; promote access and relevance of training programmes within the framework of the overall national socio-economic development plans and policies; prescribe the minimum criteria for admission to training institutions and programmes in order to promote access, equity and gender parity; establish a training system which meets the needs of both the formal and informal sectors as provided under this Act; collect, examine and publish information relating to training; inspect, license, register and accredit training institutions; undertake, or cause to be undertaken, regular monitoring, evaluation and inspection of training and institutions to ensure compliance with set standards and guidelines; among others.

1.8. The Technical and Vocational Education and Training Act, 2013

An Act of Parliament to provide for the establishment of a technical and vocational education and training system; to provide for the governance and management of institutions offering technical and vocational education and training; to provide for coordinated assessment, examination and certification; to institute a mechanism for promoting access and equity in training; to assure standards, quality and relevance; and for connected purposes (NationalCouncilforLawReporting, 2013).

1.9. Levels of Technical and Vocational Education and Training

Currently TVET in Kenya is offered at four levels namely:

i. Artisan level in Youth Polytechnics and on-the-job training in the formal sector and informal sector (Jua Kali apprentices).
ii. Craft level in Technical Training Institutes (TTIs) and Institutes of Technology (ITs).
iii. Technician level in National Polytechnics (NPs) and a few selected TTIs and ITs.
iv. Technologist in National Polytechnics and University.

Examinations offered at these four levels are offered by the Kenya National Examination Council (KNEC).

1.10. Problem Statement

Student achievement is the main agenda of any learning institution. The school administration has a responsibility to ensure students learn what they need to know depending on the course they are undertaking to prepare them for work. This requires access to a wide range of data and tools to continuously monitor student learning and performance. The administration needs to view a variety of indicators measuring student performance both internal and external factors to a students learning environment and have them presented in a way that can guide them in carrying out their management tasks. Therefore, a need arises to manage and integrate data into a dashboard for effective decision making on matters concerning the learners. This would be achieved by developing a learner analytics dashboard that can be used to predict students performances based on previous performances and give recommendations that would prompt for appropriate decisions to be made to improve on the performance.

1.11. Purpose of the Research

The aim of the research is to develop a learner analytics dashboard for predicting student performances: a case study of technical institutions.

1.12. Specific Objectives

The objectives of the study are to:

i. To establish the current methods of predicting performance in technical institutions.
ii. To determine the requirements for the dashboard system.
iii. To design a learner analytics dashboard.

iv. To implement the learner analytics dashboard system for use in technical institutions.

v. To test and validate the learner analytics dashboard.

1.13. Research Questions

i. Describe the current methods of predicting performance in technical institutions?

ii. What are the requirements needed for the dashboard system?

iii. Describe how the learner analytics dashboard will be designed?

iv. Describe how the learner analytics dashboard system will be implemented?

v. How will the learner analytics dashboard be tested and validated?

1.14. Justification/Significance of the Research

The dashboard will help identify unfavorable trends in student's performance and effort will be put to improve on weak areas. It will aid in monitoring student's performances to ascertain what is working well and where additional attention may be required.

The dashboard will also help the lecturers in being able to assist their students academically. Depending on individual performances, they will be able to grade their students on the merit of performers, average performers and poor performers. They will then be able to advice and guide them accordingly. The teachers will also be able to reflect on the impact of their teaching method on the learning behavior and performance of their students, and draw first conclusions about the effectiveness of their teaching.

The administration too will be able to monitor their students and based on courses, they can be able to tell which courses are doing well and those that are not. Working in hand with the lecturers they will be able to know the right administrative measures to put in place like student recruitment policies, adjusting course planning and determining hiring needs.

All technical institutions country wide will be able to use the system to predict their student performances. From analyzing the performances, the different institutions can come up together and incase they have any issues they can be able to consult with the examination body – KNEC.

1.15. Scope

The area of study is not generalized to all technical institutions in the country. It will be limited to technical institutions in Nairobi.

There are many national examining bodies in the country but the study will focus on examination results offered by one main national body – Kenya National Examination Council.

There are a number of artificial intelligence systems used in predicting, but the study only employs artificial neural networks. The system is a learner analytics dashboard for predicting student performances based on neural networks to do the performance prediction: a case study of technical institutions.

2 Chapter Two: Literature Review

This chapter focuses on the literature review relating to the research of learner analytic dashboards. The state of the art of literature related to learner analytics dashboards is discussed. The state of practice and technological advancement is also discussed and a critique of the related earlier works is highlighted.

2.1. State of the Art

There have been a lot of studies related to prediction of student performances in higher learning institutions. Different tools have been used to predict student performance. Also different models have been used to describe the learning analytics process.

2.2. Learning Analytics Process

According to Chatti et al. (2012), learning analytics process is often an iterative cycle and is generally carried out in three major steps:

i. data collection and pre-processing,
ii. analytics and action,
iii. post- processing.

Data collection and pre-processing:

Educational data is the foundation of the learning analytics process. First collect data from various educational environments and systems. This helps in the successful discovery of useful patterns from the data. The data collected may be too large and/or involve many irrelevant attributes, which call for data pre-processing. Data pre-processing allows transformation of the data into a suitable format that can be used as input for a particular learning analytic method.

Analytics and action:

Depending on the pre-processed data and the objective of the analytics exercise, different learning analytics techniques such as clustering, association mining rule, and social network analysis can be applied to explore the data in order to discover hidden patterns that can help to provide a more effective learning experience. The results of the mining process can be presented as a widget, which might be integrated in this case into a dashboard.

The analytics step does not only include the analysis and visualization of information, but also actions on that information - teachers are supposed to be able to more quickly interpret the visualized information, reflect on the impact of their teaching method on the learning behavior and performance of their students, and draw first conclusions about the effectiveness of their teaching, i.e., consider if their goals have been reached. Taking actions is the primary aim of the whole analytics process. These actions include monitoring, analysis, prediction, intervention, assessment, adaptation, personalization, recommendation, and reflection.

Post-processing:

Post-processing involves compiling new data from additional data sources, refining the data set, determining new attributes required for the new iteration, identifying new indicators/metrics, modifying the variables of analysis, or choosing a new analytics method. It is illustrated in the diagram below.

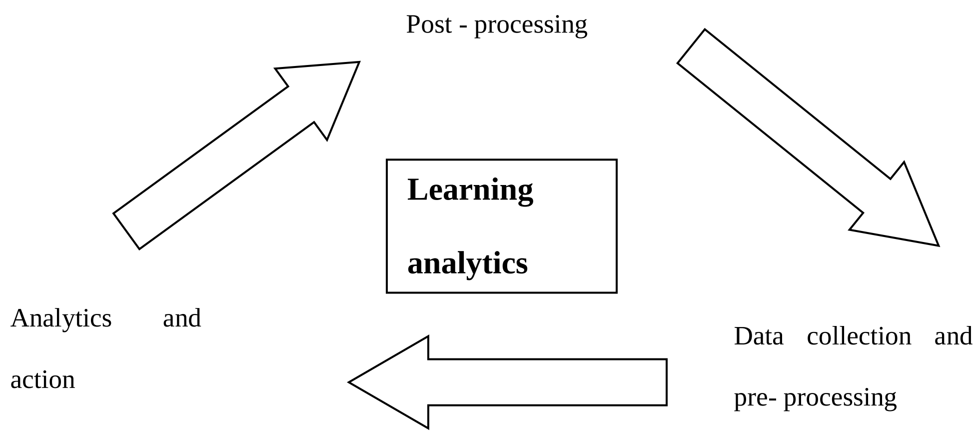


Figure 2.2-1: Learning Analytics Process (Chatti et. al, 2012)

2.3. Learning Analytics Reference Model

According to Chatti et al. (2012), the reference model for learning analytics is based on four dimensions and identifies various challenges and research opportunities in relation to each dimension. The four dimensions of the proposed reference model for learning analytics are:

What? What kind of data does the system gather, manage, and use for the analysis?

Who? Who is targeted by the analysis?

Why? Why does the system analyze the collected data?

How? How does the system perform the analysis of the collected data?

The reference model is illustrated in the diagram below.

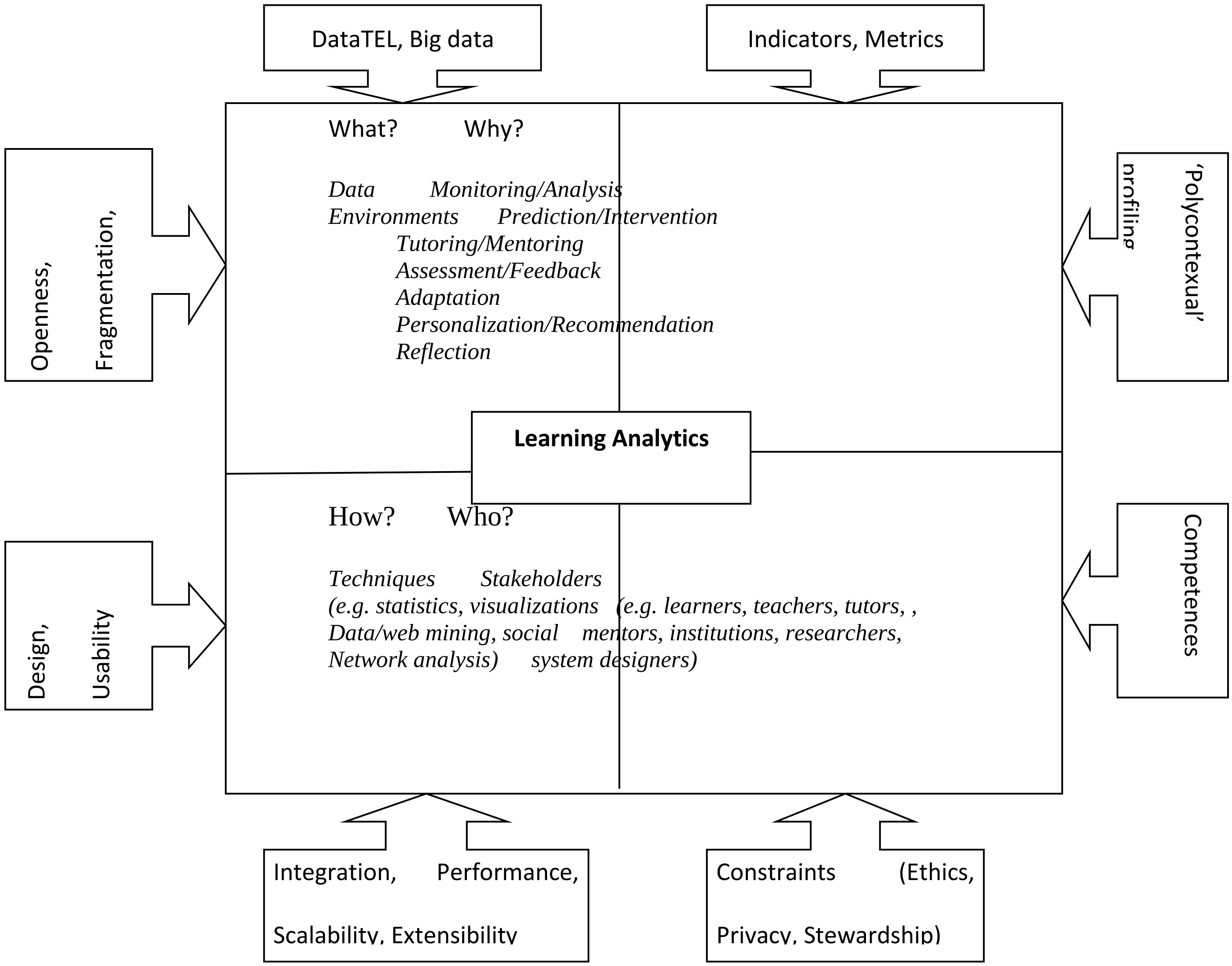


Figure 2.3-1: Learning Analytics Reference Model (Chatti et. al, 2012)

Data and Environments (What?)

Learning analytics is a data-driven approach which uses varied sources of educational data. These sources are categorized into centralized educational systems and distributed learning environments. Centralized educational systems are well represented by learning management systems (LMS). LMS accumulate large logs of data of the students' activities and interaction data, such as reading, writing, accessing and uploading learning material, taking tests, and sometimes have simple, built-in reporting tools. LMS are often used in formal learning settings to enhance traditional face-to-face teaching methods or to support distant learning.

User-generated content, facilitated by ubiquitous technologies of production and cheap tools of creativity has led to a vast amount of data produced by learners across several learning environments and systems. With the growth of user-generated content, LA based on data from distributed sources is becoming increasingly important and popular. Open and distributed learning environments are well represented by the personal learning environment (PLE) concept. As learning tools and resources are increasingly moving into the cloud, the challenge is how to aggregate and integrate raw data from multiple, heterogeneous sources, often available in different formats, to create a useful educational data set that reflects the distributed activities of the learner; thus leading to more precise and solid LA results.

Stakeholders (Who?)

The application of LA can be oriented toward different stakeholders, including students, teachers, (intelligent) tutors/mentors, educational institutions (administrators and faculty decision makers), researchers, and system designers with different perspectives, goals, and expectations from the LA exercise.

- Students will probably be interested in how analytics might improve their grades or help them build their personal learning environments.

- Teachers might be interested in how analytics can augment the effectiveness of their teaching practices or support them in adapting their teaching offerings to the needs of students.
- Educational institutions can use analytics tools to support decision making, identify potential students “at risk”, improve student success (i.e. student retention and graduation rates) develop student recruitment policies, adjust course planning, determine hiring needs, or make financial decisions

Objectives (Why?)

Possible objectives of LA include monitoring, analysis, prediction, intervention, tutoring/mentoring, assessment, feedback, adaptation, personalization, recommendation, and reflection.

- *Monitoring and analysis*: In monitoring the objectives are to track student activities and generate reports in order to support decision-making by the teacher or the educational institution.

  Monitoring is also related to instructional design and refers to the evaluation of the learning process by the teacher with the purpose of continuously improving the learning environment. Examining how students use a learning system and analyzing student accomplishments can help teachers detect patterns and make decision on the future design of the learning activity.
- *Prediction and intervention*: In prediction, the goal is to develop a model that attempts to predict learner knowledge and future performance, based on his or her current activities and accomplishments. The effective analysis and prediction of the learner performance

can support the teacher or the educational institution in intervention by suggesting actions that should be taken to help learners improve their performance.

- *Tutoring and Mentoring*: Tutoring is mainly concerned with helping students with their learning (assignments), often very domain-specific and limited to the context of a course. In tutoring processes the control is with the tutor and the focus is rather on the teaching process. In contrast, mentoring goes beyond tutoring and focuses on supporting the learner throughout the whole process – ideally throughout the whole life, but in reality limited to the time that both mentor and learner are part of the same organization. As part of this support, mentors provide guidance in career planning, supervise goal achievement, help preparing new challenges, etc. In mentoring processes the control lies rather with the learners and the focus is on the learning process.
- *Assessment and feedback*: The objective is to support the (self-)assessment of improved efficiency and effectiveness of the learning process. Important is also to get intelligent feedback to both students and teachers/mentors. Intelligent feedback provides interesting information generated based on data about the user's interests and the learning context.
- *Adaptation*: Adaptation is triggered by the teacher/tutoring system or the educational institution. The goal of LA here is to tell learners what to do next by adaptively organizing learning resources and instructional activities according to the needs of the individual learner.
- *Personalization and recommendation*: In personalization, LA is highly learner-centric, focusing on how to help learners decide on their own learning and continuously shape their PLEs to achieve their learning goals.

- *Reflection*: Analytics can be a valuable tool to promote reflection. Students and teachers can benefit from data compared within the same course, across classes, or even across institutions to draw conclusions and (self-) reflect on the effectiveness of their learning or teaching practice.

  Learning by reflection (or reflective learning) offers the chance of learning by returning to and evaluating past work and personal experiences in order to improve future experiences and promote continuous learning.

The above mentioned various objectives in LA are sometimes difficult to measure and need a tailored set of performance indicators and metrics. Moreover, we need to define new metrics beyond grades and graduation rates in order to support different types of learning including self organized learning, network learning, informal learning, professional learning, and lifelong learning. The challenge is thus to define the right Objective / Indicator / Metric (OIM) triple before starting the LA exercise.

Methods (How?)

LA applies different techniques to detect interesting patterns hidden in educational data sets. Some of these techniques are:

- *Statistics*: Most existing learning management systems implement reporting tools that provide basic statistics of the students' interaction with the system. Examples of usage statistics include time online, total number of visits, number of visits per page, distribution of visits over time, frequency of student's postings/replies, percentage of material read. These statistical tools often generate simple statistical operations such as average, mean, and standard deviation

- *Data Mining (DM)*: Data mining, also called Knowledge Discovery in Databases (KDD), is defined as "the process of discovering useful patterns or knowledge from data sources, e.g., databases, texts, images, the Web".
- *Social Network Analysis (SNA)*: As social networks become important to support networked learning, tools that enable to manage, visualize, and analyze these networks are gaining popularity. By representing a social network visually, interesting connections could be seen and explored in a user-friendly form.

2.4. Dashboard Theories

According to Velcu-Laitinen & Yigitbasioglu (2012), dashboards draw on theories from a variety of disciplines including information systems, accounting, and cognitive psychology. They communicate complex data to the decision maker through a process called visualization. Visualization refers to interactive visual representations of abstract, non-physically based data to amplify cognition. The motivation for using visualization in dashboards can probably be best explained through:

2.4.1 Cognitive Fit Theory

Cognitive fit theory focuses on the fit between an individual’s decision-making skills, the information presentation format and the task at hand, providing useful guidelines with regard to the choice of presentation format to be applied. For example, graphs are well suited for spatial tasks that involve forecasting and comparisons as well as for tasks that require multidimensional data analysis and pattern recognition. On the other hand, tabular information is more appropriate for symbolic tasks and for users who are more numerical like accountants and financial analysts. Cognitive fit theory postulates that decisions made under the fit condition are superior compared to the non-fit. Since tasks related to performance monitoring involve comparisons of a variety of measures (Key Performance Indicators), the presentation of information in graphical format is well justified. Nevertheless, because of the diversity of dashboard user's knowledge, skills, and cognitive profiles, it is recommended that dashboards come with some level of flexibility in terms of drill down capabilities and presentation format flexibility (Velcu-Laitinen & Yigitbasioglu, 2012).

2.4.2. Task-technology Fit Theory

It focuses on the fit between individual abilities, task requirements, and the functionality of the technology. For example, certain kinds of tasks that require information from different organizational units require technological functionality under the form of integrated databases throughout an organization. Thus dashboards need to be designed in the light of the tasks and users that it is required to support (Velcu-Laitinen & Yigitbasioglu, 2012).

In the context of information systems research, technology refers to computer systems (such as hardware, software, and data) and user support services (such as training and help lines). Technologies are viewed as tools used by individuals in carrying out their tasks. Tasks are defined as "the actions carried out by individuals in turning inputs into outputs. In order for an information system to have a positive impact on individual performance the technology must be utilized and there must be a good fit with the tasks the technology supports. If either the task-technology fit of the technology or its utilization is lacking, the technology will not improve performance (Irick, 2008).

2.5. Business Intelligence Technology

According to Ranjan (2009), business intelligence provides organizational data in such a way that the organizational knowledge filters can easily associate with this data and turn it into information for the organization. Persons involved in business intelligence processes may use application software and other technologies to gather, store, analyze, and provide access to data, and present that data in a simple, useful manner. The software aids in Business performance management, and aims to help people make "better" business decisions by making accurate, current, and relevant information available to them when they need it. Some businesses use data warehouses because they are a logical collection of information gathered from various operational databases for the purpose of creating business intelligence.

In order for BI system to work effectively there must be some technical constraints in place. BI technical requirements have to address the following issues: security and specified user access to the warehouse; data volume (capacity); how long data will be stored (data retention); and benchmark and performance targets

People working in business intelligence have developed tools that ease the work, especially when the intelligence task involves gathering and analyzing large quantities of unstructured data. Each vendor typically defines Business Intelligence their own way, and markets tools to do BI the way that they see it. BI often uses Key performance indicators (KPIs) to assess the present state of business and to prescribe a course of action. More and more organizations have started to make more data available more promptly.

2.6. Business Intelligence Framework

According to Parker (2008), the following sections contain directions for how to create the Business Intelligence framework, from source data to finished dashboards and scorecards. The following diagram illustrates the major components of the framework, and the flow of data from source data to the finished dashboards. Data processes are numbered one through six in the diagram, with detailed descriptions to follow.

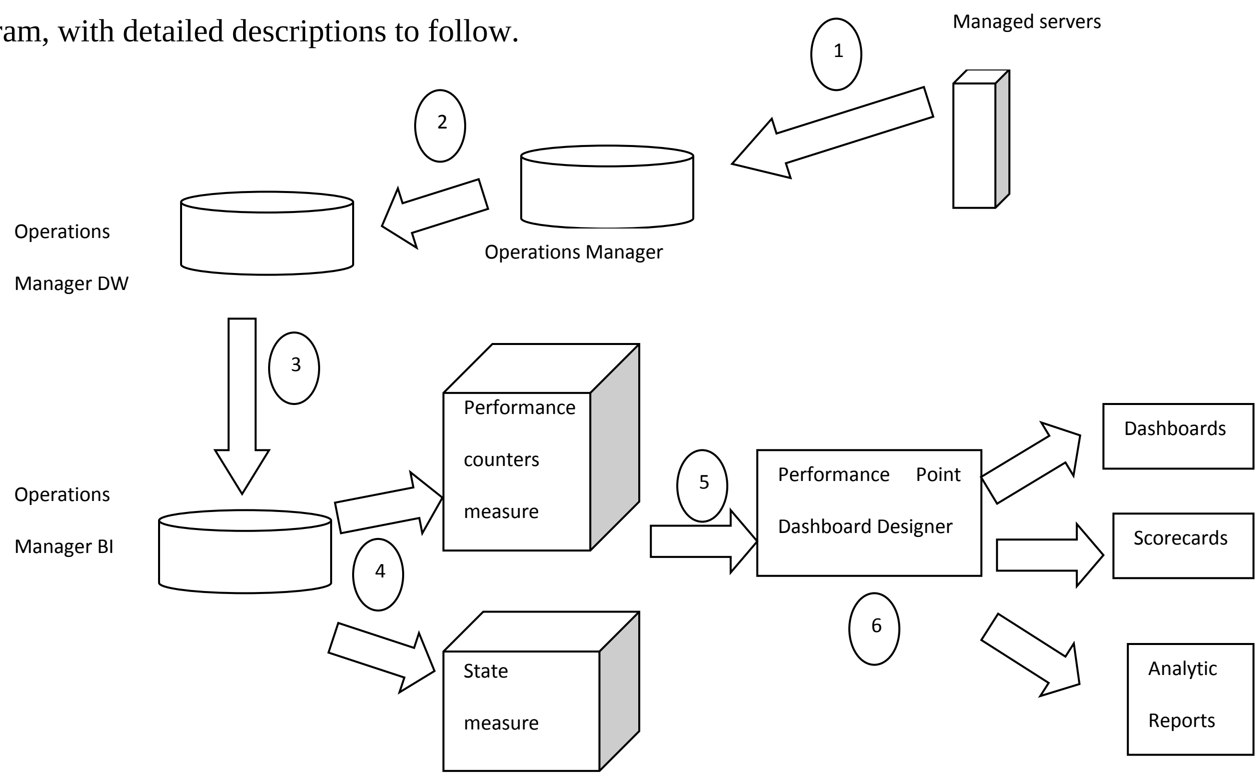


Figure 2.6-1: Data flow diagram (Parker, 2008)

1. System Center Operations Manager managed servers collect monitor state and performance counter data saved in the Operations Manager operational database.

2. The Operations Manager data warehouse (Operations Manager DW) collects data from the operational database. Data is pre-aggregated and stored in tables architected to support production reporting requirements.

3. Collect the Source Data - A small subset of data in the Operations Manager data warehouse is transformed and loaded into the Business Intelligence framework database (Operations Manager BI). This database contains the star schemas for the Analysis Services OLAP cubes.

4. Build OLAP Cubes - Analysis Services OLAP cubes are built and processed from data stored in the Operations Manager BI database.

5. Put it All Together - Create Office Performance Point Server 2007 Dashboards, Scorecards, and Analytic Reports - Data from the OLAP cubes is used to populate Performance Point Server scorecards, dashboards, and analytic reports. These components are originally created using the Office Performance Point Server 2007 Dashboard Designer.

6. Deploy Dashboards, Scorecards, and Analytic Reports in SharePoint - Scorecards, dashboards, and analytic reports are made available to the user community through SharePoint Server. Once the scorecards, dashboards, and analytic reports are initially created and deployed, they should not need to be deployed again – these components will be refreshed as new data becomes available in the OLAP cubes.

2.7. Artificial Intelligence (AI) in Predicting.

According to Hall (2002), AI is defined as a computer-based analytical process that exhibits behavior and actions that are considered “intelligent” by human observers. Forecasting is the process of estimating future events, and it is fundamental to all aspects of management. The goals of forecasting are to reduce uncertainty and to provide benchmarks for monitoring actual performance.

2.8. Methods/Tools Used for Predicting Student Performance

In educational data mining method, predictive modeling is usually used in predicting student performance. In order to build the predictive modeling, there are several tasks used, which are classification, regression and categorization. The most popular task to predict student's performance is classification. There are several algorithms under classification task that have been applied to predict student's performance. Among the algorithms used are:

2.8.1. Decision Tree

Decision Tree is one of a popular technique for prediction. A decision tree is a classifier expressed as a recursive partition of the instance space. The decision tree consists of nodes that form a rooted tree. All other nodes have exactly one incoming edge. A node with outgoing edges is called an internal or test node. All other nodes are called leaves (also known as terminal or decision nodes). In a decision tree, each internal node splits the instance space into two or more sub-spaces according to a certain discrete function of the input attributes values. In the simplest and most frequent case, each test considers a single attribute, such that the instance space is partitioned according to the attribute's value, (Rokach & Maimon, 2005).

According to Shahiri et.al (2015), most of researchers have used this technique because of its simplicity and comprehensibility to uncover small or large data structure and predict the value. The decision tree models are easily understood because of their reasoning process and can be

directly converted into set of IF-THEN rules. Examples of previous studies using Decision Tree method are predicting drop out features of student's data for academic performance, predicting third semester performance of MCA students and also predicting the suitable career for a student through their behavioral patterns. The students performance evaluation is based on features extracted from logged data in an
education web-based system. The examples of dataset are student's final grades, final cumulative grade point average (CGPA) and marks obtained in particular courses.

2.8.2. Neural Network

Neural network is another popular technique used in educational data mining. According to Haykin (2005), a neural network is a massively parallel distributed processor made up of simple processing units which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects: knowledge is acquired by the network from its environment through a learning process; and inter neuron connection strengths known as synaptic weights are used to store the acquired knowledge. The purpose of learning process is to modify the synaptic weights of the network in an ordered way so as to be able to attain a desired design objective.

According to Shahiri et.al (2015), the advantage of neural network is that it has the ability to detect all possible interactions between predictor's variables. Neural network could also do a complete detection without having any doubt even in complex nonlinear relationship between dependent and independent variables. Therefore, neural network technique is selected as one of the best prediction method.

2.8.3. Naive Bayes

Naive Bayes algorithm is also an option for researchers to make a prediction. Naive Bayes is the simplest form of Bayesian network, in which all attributes are independent given the value of the class variable. This is called conditional independence. It is obvious that the conditional independence assumption is rarely true in most real-world applications. A straightforward approach to overcome the limitation of naive Bayes is to extend its structure to represent explicitly the dependencies among attributes, to extend its structure to represent explicitly the dependencies among attributes (Zhang, 2004).

2.8.4. K-Nearest Neighbor

According to Sweeney (2016), the k-Nearest Neighbors (kNN) algorithm is a classic method for clustering samples based on similarity. These clustering's can be used for regression in the following manner. First a pair wise distance metric is used to identify the k most similar neighbors among all dyads in the training set. The grade is then predicted using local interpolation of the targets associated with these neighbors.

According to Shahiri et.al (2015), K-Nearest Neighbor gave the best performance with the good accuracy. As stated by Bigdoli et al. (2003), K-Nearest Neighbor method had taken less time to identify the student's performance as a slow learner, average learner, good learner and excellent learner. K-Nearest Neighbor gives a good accuracy in estimating the detailed pattern for learner's progression in tertiary education.

2.8.5. Support Vector Machine

According to Shahiri et.al (2015), Support Vector Machine is a supervised learning method used for classification. Support Vector Machine is suited well in small datasets. Sembiring et al.

(2011) stated that Support Vector Machine has a good generalization ability and faster than other methods. Meanwhile, the study done by Gray et al (2014) demonstrated that Support Vector Machine method has acquired the highest prediction accuracy in identifying students at risk of failing.

2.9. Neural Network Architecture

According to DTREG (2016), all neural networks have an input layer and an output layer, but the number of hidden layers may vary. When there is more than one hidden layer, the output from one hidden layer is fed into the next hidden layer and separate weights are applied to the sum going into each layer.

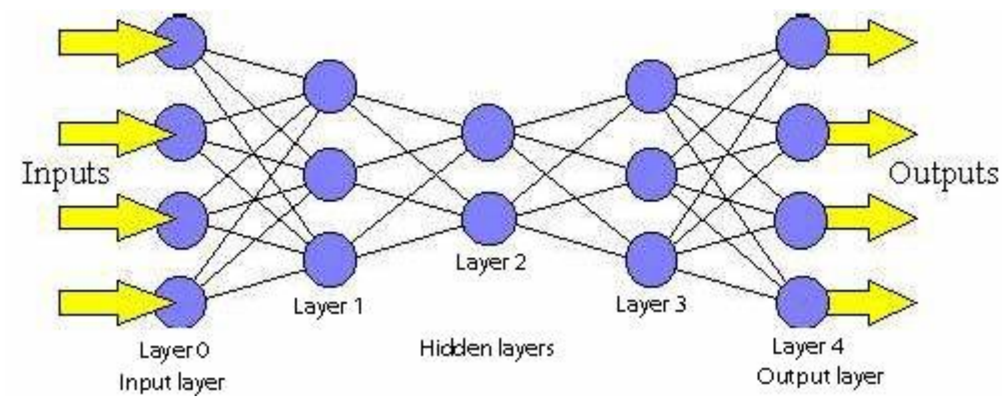


Figure 2.9-1: Feed forward network (Roberts, 2000)

According to Roberts (2000), in feed-forward networks, pairs of input and output values are fed into the network for many cycles, so that the network 'learns' the relationship between the input and output.

2.10. State of Practice – Case Studies

There have been a number of systems developed to predict student's performance. This are:

2.10.1. Learning Online Network with Computer-Assisted Personalized Approach (LON-CAPA)

According to Minaei-Bidgoli et.al (2003), two large databases have been developed in LON-CAPA. The first contains educational resources such as web pages, demonstrations, simulations, and individualized problems designed for use on homework assignments the Learning Online Network with Computer-Assisted Personalized Approach (LON-CAPA) quizzes, and examinations. As more instructors develop educational materials for their courses to use with the LON- CAPA system, the content of this database grows.

The second database contains information about student users of LON-CAPA. This database stores a wide range of variables including when, for how long, and how many times they access each resource, the number of correct responses they give on assigned problems, their pattern of correct and incorrect responses, and so on.

2.10.2. SQL-Tutor Intelligent Tutoring system

An intelligent problem-selection agent, which identifies the appropriate problem for a student in two stages. It firstly predicts the number of errors the student will make on a set of problems, and then in the second stage decides on a suitable problem for the student. In order to develop such an agent, a feed-forward, back propagation neural network was trained to predict the number of errors a student will make. The achieved prediction accuracy is high, showing that a neural network is capable of making such predictions, (Wang & Mitrovic, 2002).

2.10.3. School-Level Performance Assessments Dashboard

The principal in each school building has a lot to manage with regard to its many educators, administrators and health professionals, as each individual contributor will directly affect student performance outcomes. A school level-assessment dashboard will allow a principal to track teacher performance data from observations, certifications, years of experience, and parent surveys. Even more importantly, it will allow for an assessment at the school level into student behavior, demographics, discipline, and performance on standardized assessments from year to year. Commonly used at the school level, principal dashboards, will allow for school buildings within a district to be ranked against each other, which aids in quickly identifying low performing schools, and even drop-out factories. All of this data together can bring forth early indicators so principals and teachers can pinpoint struggling students and take action for improvement, (Tableau, 2016).

2.10.4. Student Insights for Teachers Dashboard

Because teachers are the front line to keeping students on a trajectory toward graduation, it is crucial to empower them with all of the available student data. For teachers, need-to-know information goes beyond standardized tests scores, attendance and discipline. Student demographics, zip codes, social-emotional learning data and parent survey data, all have need-to-know insights for combating poor performance. Teachers use a common assessment dashboard for each individual student in conjunction with a failing grades report to spot students in need of intervention. Using data from how students performed on tests in the past, teachers can help individuals set growth goals for improving scores and grades throughout the year. By including predictions of student scores, teachers get detailed information about what is most actionable for curriculum and student improvement, (Tableau, 2016).

2.10.5. Grades, Questions, and Concerns Answered for Parents Dashboard

Mounting evidence in data shows that when schools and teachers partner with parents, not only do students stay in school longer, they perform better too. So when parents step in to act as advocates for their children, schools want to support that behavior by giving them correct, actionable information. Using a dashboard as a student portfolio, teachers can share information directly with parents about individualized student progress. Further enhancing the parent-teacher/administrator relationship, data visualizations can act as sounding boards for when parents have specific and immediate questions about grades or anything else related to their student's success.

2.10.6. Infrastructure and Facilities Decision-Making Dashboard

The path to student success has hidden obstacles in things like access to comfortable classrooms and up-to-date lab instruments. If students are not equipped with the appropriate materials, technologies and environments with which to learn, they are less likely to succeed. Data visualizations for facilities, operations and finances at educational institutions can shed light on how seemingly small decisions can have big negative or positive impacts. A technology asset dashboard is used to have a clear understanding of all of the assets at a particular campus, how old they are, and how they were acquired. From laptops and tablets to projectors and document cameras, principals and administrators use this dashboard to re-examine and prioritize technological needs each year.

2.11. Critique of the Related Work

According to Shahiri et.al (2015), several factors were used on decision trees, neural network, naïve bayes, K-Nearest Neighbor and Support Vector Machine to predict student's performance. This included internal assessment, psychometric factors, external assessment, social network interaction, CGPA, student demographic and extra- curricular activities. Neural Network was seen to have the highest prediction accuracy by (98%) followed by Decision Tree by (91%). Next, Support Vector Machine and K-Nearest Neighbor gave the same accuracy, which is (83%). Lastly, the method that has lower prediction accuracy is Naive Bayes by (76%).

The result on prediction accuracy is depending on the attributes or features that were used during the prediction process. Neural Network method gave the highest prediction accuracy because of the influence from main attributes. This is as illustrated in the figure below:

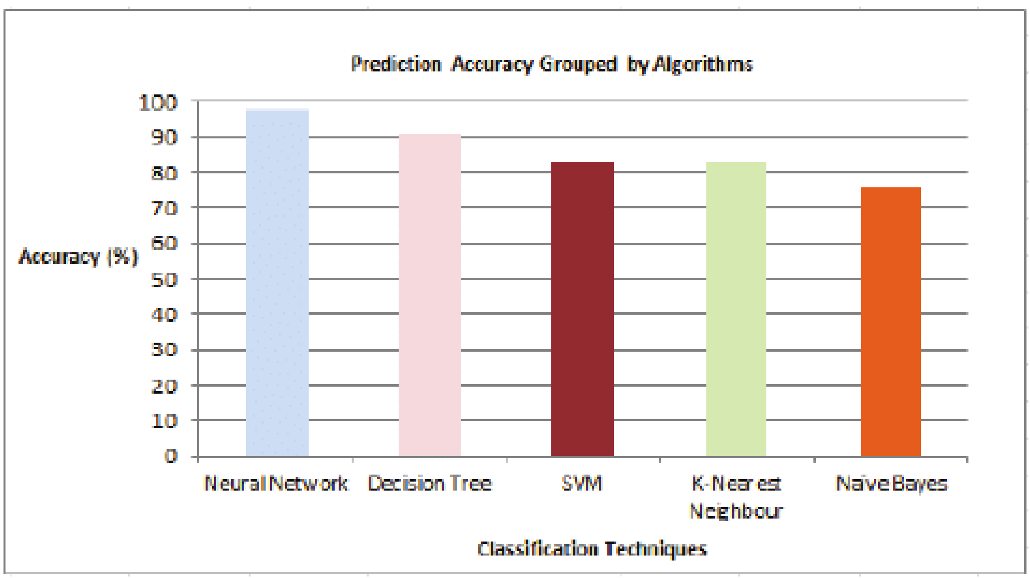


Figure 2.11-1: Prediction accuracy grouped by algorithms since 2002 – 2015 (Shahiri et.al, 2015)

This attributes are hybridization of two features, which are internal and external assessments. With the use of only one variable, which is external assessments, the accuracy is decreased by (1%). The third most used variable is internal assessments that gave the result of (81%) accuracy.

It shows that external assessment, which is the marks obtained in final examination, plays an important role in predicting student's performance. While, significant variable that gave the least impact on student performance is psychometric factors with the accuracy of only (69%). The psychometric factors usually use qualitative data, so it is difficult for Neural Network algorithm to make a prediction instead of using quantitative data. However, Neural Network method still has less maximum error prediction. The maximum error of prediction is less than (10%). Another advantage of Neural Network is the ability to capture nonlinear relationships easily. It is also referred as adaptive system due to its ability to readily update the historical data like a human brain. So, the model always functions beyond the knowledge base. In addition, the strength of neural network is the ability to learn from a limited set of data.

3. Chapter Three: Methodology

How successful a research study is depends on the methodology used for the given study. There are different types of methodologies but the methodology chosen depends on the type of research under study.

3.1. Existing Methodologies

Theoretical Methodology

Theories are formulated to explain, predict, and understand phenomena and, in many cases, to challenge and extend existing knowledge within the limits of critical bounding assumptions. The theoretical framework is the structure that can hold or support a theory of a research study. The theoretical framework introduces and describes the theory that explains why the research problem under study exists, (Swanson & Chermack, 2013).

Experimental Methodology

According to McLeod (2012), an experiment is an investigation in which a hypothesis is scientifically tested. In an experiment, an independent variable (the cause) is manipulated and the dependent variable (the effect) is measured; any extraneous variables are controlled. An advantage is that experiments should be objective. There are three types of experiments:

- Laboratory / Controlled Experiments - This type of experiment is conducted in a well-controlled environment – not necessarily a laboratory – and therefore accurate measurements are possible. The researcher decides where the experiment will take place, at what time, with which participants, in what circumstances and using a standardized procedure. Participants are randomly allocated to each independent variable group.

- Field Experiments - Field experiments are done in the everyday (i.e. real life) environment of the participants. The experimenter still manipulates the independent variable, but in a real-life setting (so cannot really control extraneous variables).

- Natural Experiments - Natural experiments are conducted in the everyday (i.e. real life) environment of the participants, but here the experimenter has no control over the IV as it occurs naturally in real life.

Case Study Methodology

Case study method enables a researcher to closely examine the data within a specific context. In most cases, a case study method selects a small geographical area or a very limited number of individuals as the subjects of study. Case studies, in their true essence, explore and investigate contemporary real-life phenomenon through detailed contextual analysis of a limited number of events or conditions, and their relationships, (Zainal, 2007).

Evolutionary Prototyping Methodology

Evolutionary prototyping is a software development method where the developer or development team first constructs a prototype. After receiving initial feedback from the customer, subsequent prototypes are produced, each with additional functionality or improvements, until the final product emerges, (Sherrel, 2013).

3.2. Evaluation of the Methodologies

From the different methodologies described above, the study will employ evolutionary prototyping methodology. A prototype of the prediction system will first be developed then enhanced to meet user needs. It will be combined with case study methodology which will help define the scope of data collection.

3.3. Evolutionary Prototyping Methodology

According to Sherrel (2013), evolutionary prototyping is a software development method where the developer or development team first constructs a prototype. After receiving initial feedback from the customer, subsequent prototypes are produced, each with additional functionality or improvements, until the final product emerges. Below is a diagram showing the stages of evolutionary prototyping.

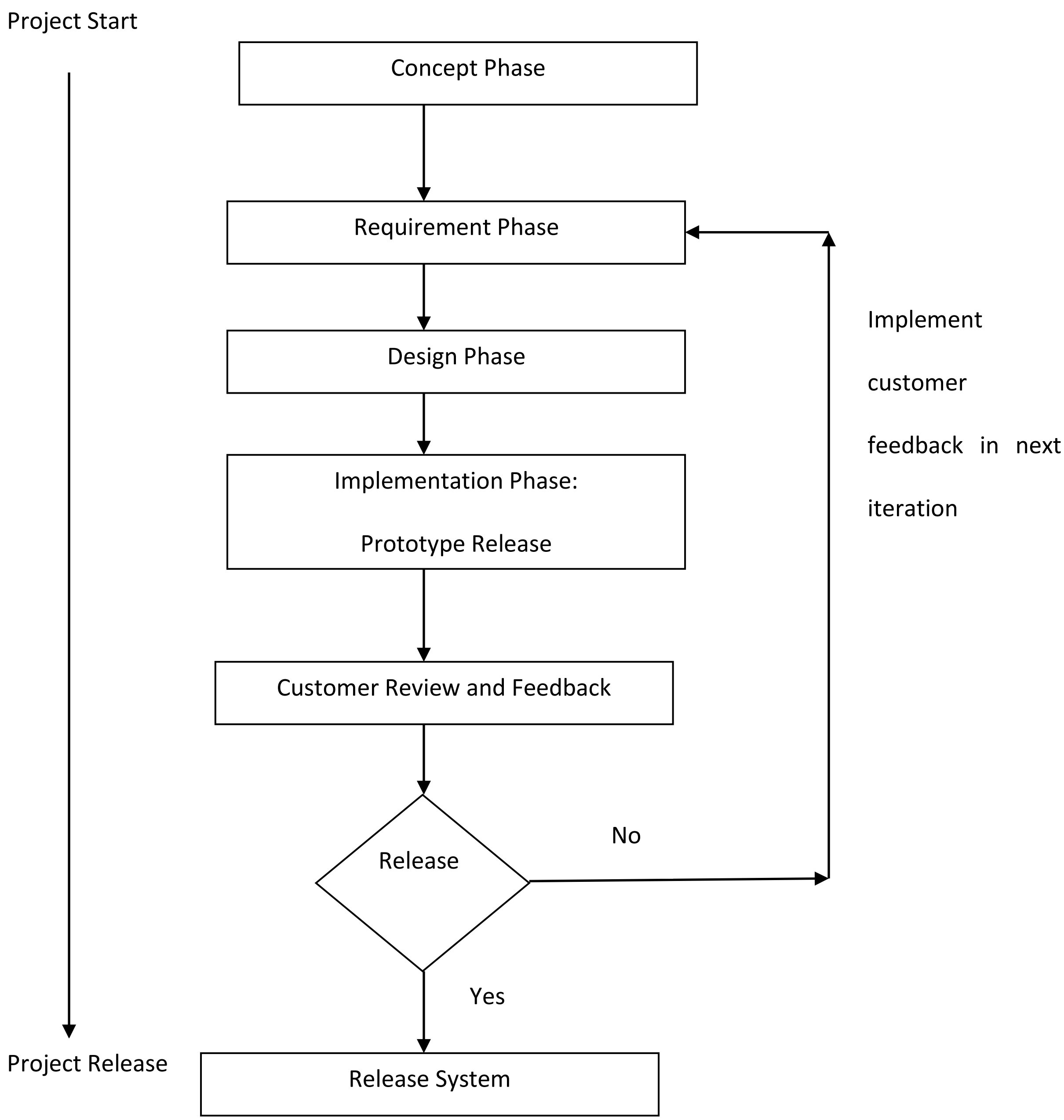


Figure 3.3-1: Evolutionary prototyping model (Lenz & Moeller, 2008)

Concept phase

The researcher interacted with the administration and lecturers of technical institutions to be able to know courses done in the institution and the examining bodies, how student information is stored, identify factors affecting student’s performance and how they track student performance. The researcher also interacted with students to know about their academic performances.

Requirements phase

The researcher interacted with the administration, lecturers and students who constituted the target population for the study. From the entire population a sample size was determined using purposive sampling technique. The study employed semi structured interviews and record inspection as the main tools for collecting data.

Record inspection was done on files in which students personal details and result slips are stored. The selection of these instruments was guided by the nature of data to be collected and goals of the study. The main aim was to gather data that would help in design and development of a dashboard for predicting student performance while ensuring user requirements are well met.

Semi structured interviews were used to collect data related to the research questions. They were preferred because they allowed the respondents to talk in detail and in depth on the research topic. They also allowed the researcher/interviewer to probe the respondents more depending on the answers they gave. In this interview the researcher was the only interviewer. The interviewees included the administration, lecturers and students at the technical institutions. The interview guides are in appendices I, II and III. The researcher was able to give the respondents time and scope to talk about courses done in the institution, how student information is stored, factors affecting students performance, how student performance is tracked and students academic performances.

Design Phase

It's a phase that systematically gives a blue print of the system being developed, identifying modules and components, interfaces of the components and the data that goes through the system. The design specification was used in developing the prototype. The neural network was developed using Encog library. It was built as described in section 2.11. The neural network is based on the different courses as each course has different units being done. The neural networks have three layers – input, hidden and output layers. The code for the neural network is in appendix VII.

An input interface was also developed to facilitate input data into the system. Output of the neural network is represented by the dashboard.

Implementation phase – prototype release

The design specification was discussed with the system users to check if it captured all the requirements of the system. A prototype was then developed through iterative refinements to ensure that a working system is delivered to the users. The refinements were based on the user's review of the model prototype and modifications made from their feedback on expectations of the system.

Customer review and feedback

The system was presented to the users at the institution to test if it was working to their satisfaction – can it be used to predict student performance given the input parameters. All their comments were noted and were used to enhance the system.

Release

Once the system was considered to have met all the requirements and was working as per the expectations it was released to the potential users.

3.4. Population and Sampling

The target population was the administration, lecturers and students of technical institutions in Nairobi County. Cluster sampling was used in determining the sample groups to be interviewed. The sample size was then determined by purposive sampling technique (table 3:4-1). Thirteen administrators, sixty eight lecturers and two hundred and fifty five students were chosen and interviewed on academic issues that pertain to student's performance.

Institution	Administrators		Lecturers		Students	
	Population	Sample	Population	Sample	Population	Sample
NYS Institute of Business Studies	5	4	35	20	1000	80
NYS Engineering School	5	4	30	18	1200	50
Kabete Technical Training Institute	8	3	80	20	2500	85
Railways Training Institute	6	2	42	10	865	40
TOTAL	**24**	**13**	**187**	**68**	**5565**	**255**

Table 3:4-1: Population and sample

4 Chapter Four: Data and Requirements Analysis

4.1. Introduction

According to Dey (2005), data analysis is the process of resolving data into its constituent components, to reveal its characteristic elements and structure. It follows collection of information and precedes interpretation of that information. Requirement analysis aids in understanding what the user wants. The system has been developed for use by the administration, teaching staff and students in technical institutions to aid in predicting student performance.

4.2. Description of the Basic Data Set

4.2.1. Data Source

The data was obtained from technical institutions in Nairobi County. This was student's module I, II and III result slips that depended on the course the student was undertaking.

4.2.2. Dataset Understanding

The results slips obtained for module I, II and III indicated grades ranging from distinction 1 to 8. A sample of the result slip is as shown in appendix IV. The grading is as shown in the table below:

Grade	Description
1	Distinction
2	Distinction
3	Credit
4	Credit
5	Pass
6	Pass
7	Refer
8	Fail
X	CRNM (Course not complete)

Table 4.2-1: Grading for business exams

4.3. Data Analysis and Results

The results on the result slips were keyed in to Ms Excel from where they were fed into the system as .csv files. Sample data for two courses – Diploma ICT module III and craft supply chain management module I results are as shown in appendix V and VI respectively..

4.4. Requirement Specification

According to Rouse (2007), a software requirements specification (SRS) is a comprehensive description of the intended purpose and environment for software under development. The SRS fully describes what the software will do and how it will be expected to perform. It addresses the 'what' of the system.

The following are the requirement for the developed system:

i. To identify factors/parameters that affect student performance.

ii. To collect students examination data and relate it to parameters that affect individual performance.

iii. To predict students performance using the performance dashboard.

4.4.1. User Requirement Specification

To identify user requirements, interviews and record inspection were used. The interview guides used are in Appendix I, II and III. They were used to gather data from administration, lecturers and students at technical institutions. Face to face interviews were used so that they aid in getting data on procedures and operations, verify understanding of the system with the user, and validate different aspects of the proposed system and also to build user confidence in the new system. Record inspection was used to collect data on student's results in KNEC examinations.

4.4.2. Functional Requirements

These requirements define functions of a given system or its components describing an activity or process that the system must perform.

- To enter into the system the students data.
- To predict students expected grades in the current module based on the previous module and give appropriate recommendations.
- To graphically display output on a dashboard.

4.4.3. Non-functional Requirements

According to Sommerville (2004), non-functional requirement specifies systems properties and constraints.

The proposed system is required to meet non-functional requirements categorized as follows:

1. Product requirements – They specify that the delivered product must behave in a particular way for instance:-
   - Performance – data throughput, speed and accuracy of the server processing student's examinations data will be as fast as possible. Response time for client-side application requests for data will also be fastened.
   - Efficiency – error rates will be minimized as much as possible with high speeds of system operation.
   - Maintainability – the system will be made in such a way that repairs and new functionalities will be easily acceptable.
   - Reliability – the dashboard should run with minimal and/or no system outages.
   - Compatibility with the Windows operating system while the client-side application should be browser independent.
   - Other aspects that will be put into consideration include: extensibility, flexibility, usability and portability. It should be possible to add new features to the prediction system. The interface of the system should be easy to use.

4.4.4. Entity relationship diagram (ERD)

According to Dennis et.al (2012), an entity relationship diagram (ERD) is a graphical representation which shows the information that is created, stored, and used by a business system. It shows the individual pieces of information in a system and how they are organized and related to each other that is logical relationship of entities so as to create a database. The following is the ERD for the system:

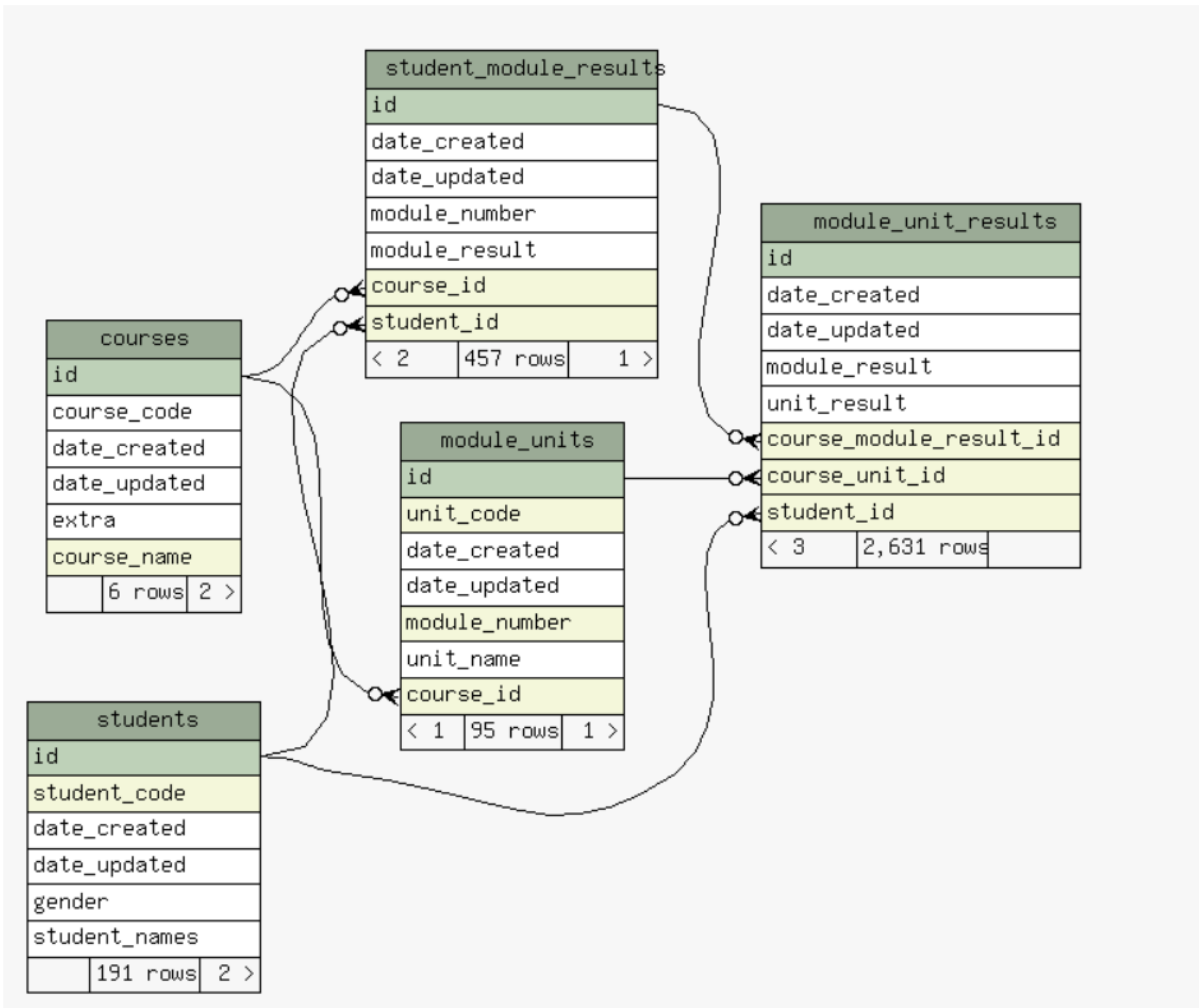


Figure 4.21: Entity relationship diagram for the system

4.5. Conceptual Framework

The following is the conceptual framework for the dashboard system.

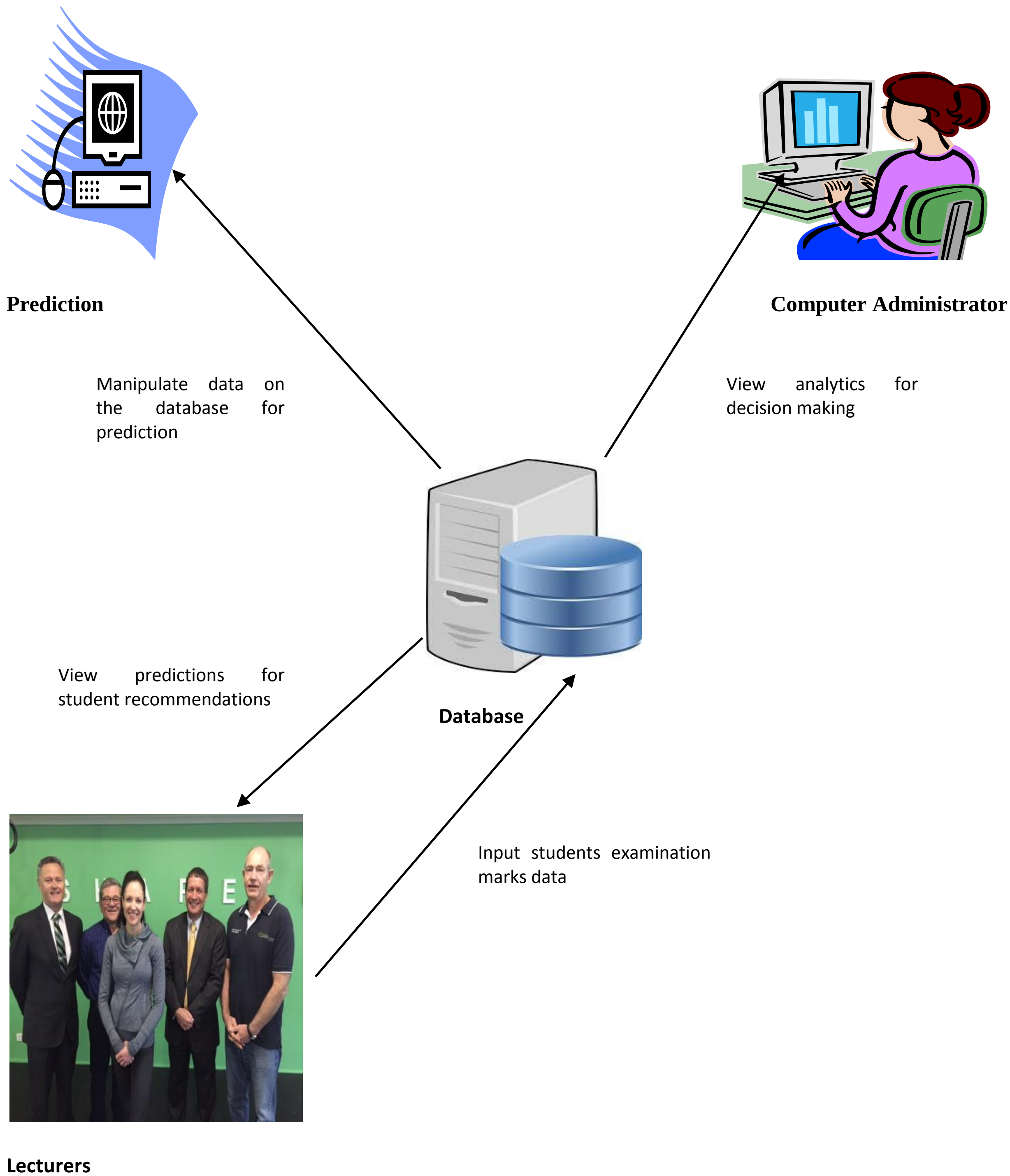


Figure 4.5-1: Conceptual framework

5 Chapter Five: Implementation

5.1. Implementation Environment

The system was implemented using the following tools:

- Languages - Java Development Kit (JDK), Java Runtime Environment (JRE) and Netbeans.
    a. Java Development Kit (JDK) is a development environment that was used to build and compile applications using Java programming language.
    b. Java Runtime Environment (JRE) was installed in the user's machine to run the application. It establishes a connection between popular browsers and the Java platform.
    c. Net beans IDE is a free and open source integrated development environment that was used to develop the project.
    d. Embedded Jetty server to avail the web server.
- Libraries:
    a. Encog library is an advanced Artificial Intelligence framework that was used for building the neural network application.
- MySQL database
- Ms Excel – the students results were presented in Ms Excel.

5.2. System Development

A learner analytics dashboard for predicting student performances was accurately developed that predicted student performances. The user keys in student's details – name, code, gender; course details – course code, name, department name; and module details – module, module units, results in each module.

To view individual student performance, select the course the student is taking and select the individual student you want to view their performance. The student's code, gender and name will be displayed. If a student is taking a diploma course, their results will be displayed for each unit in module one, two and three. If the student is taking a craft course, their results will be displayed for each unit in module one and two. The results are displayed as D1 and D2 for a distinction; C3 and C4 for a credit; P5 and P6 for a pass; R7 and R8 for a refer; and X for a unit not done.

To view performance for the different courses done in the institution, select the course you want to view. The course code is displayed showing the total number of students in that course/class both male and female. Their performance is displayed using bar graphs and pie charts showing the number of students who scored what grade – distinction, credit, pass, refer, absent or referred. It also displays percentage score for each grade using pie charts. Percentage score of each grade is also shown using pie charts based on gender.

To predict student performance, enter name of the student, select the course the student is taking. Select the module for which you want to do prediction. One will be promoted to enter the results for the previous module. Results will then be displayed for each unit, overall score and a probable recommendation for the student scoring that grade.

5.3 Screen Shots

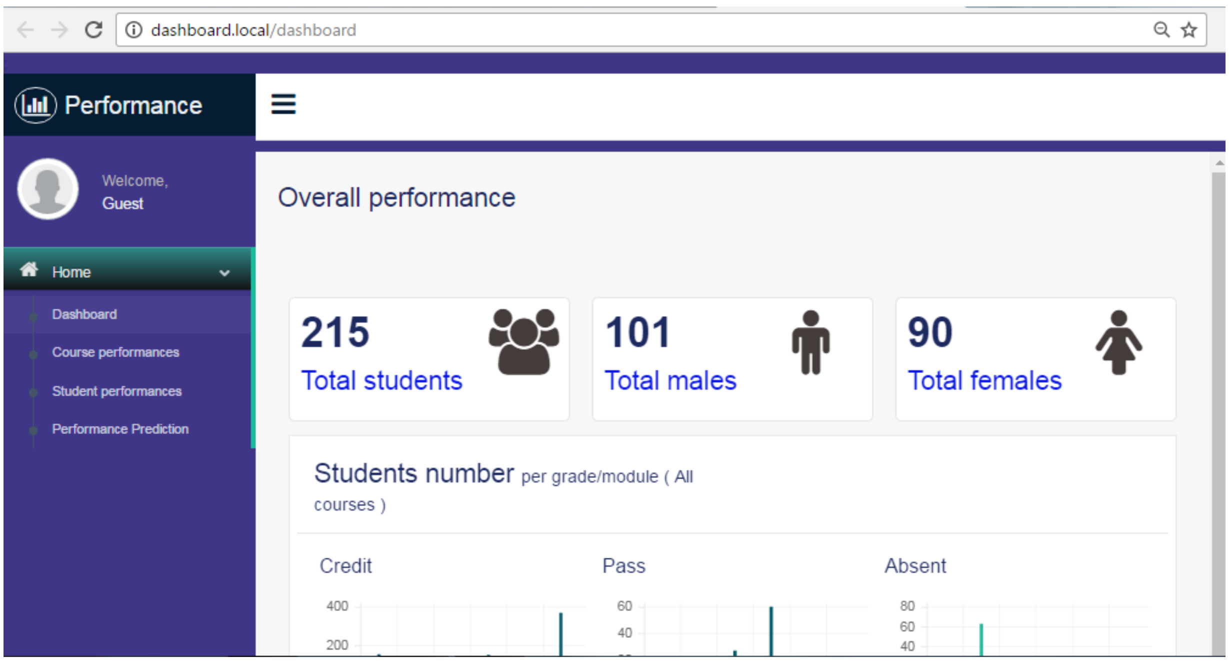


Figure 5.3-1: Dashboard Interface

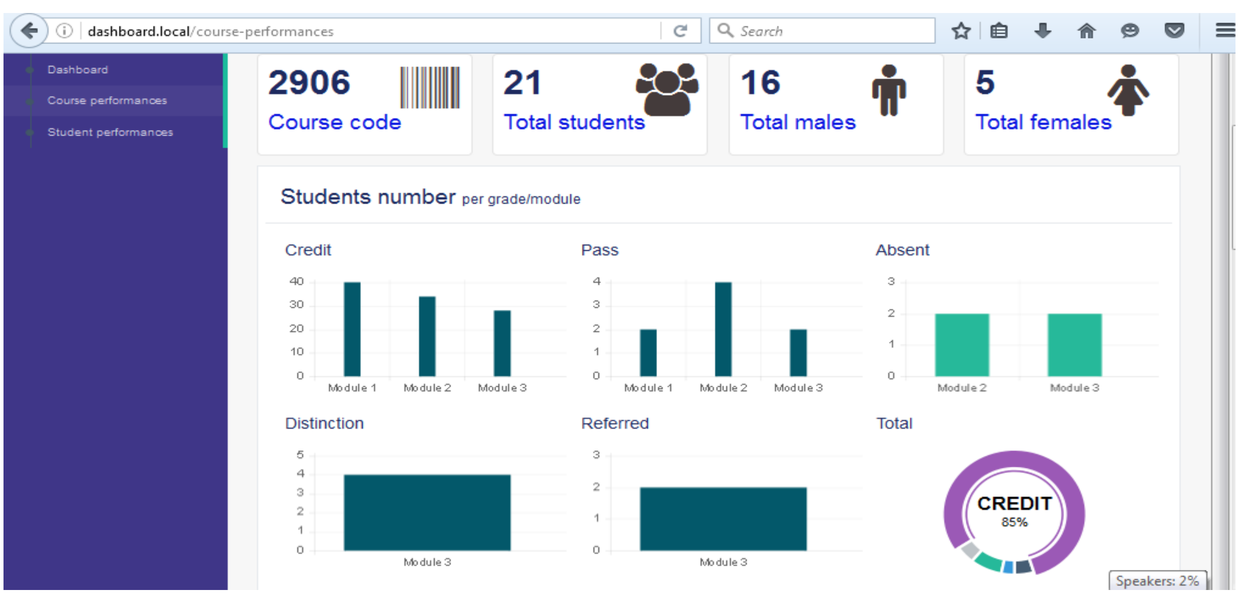


Figure 5.3-2: Course performance details showing number of students for each grade in the given module

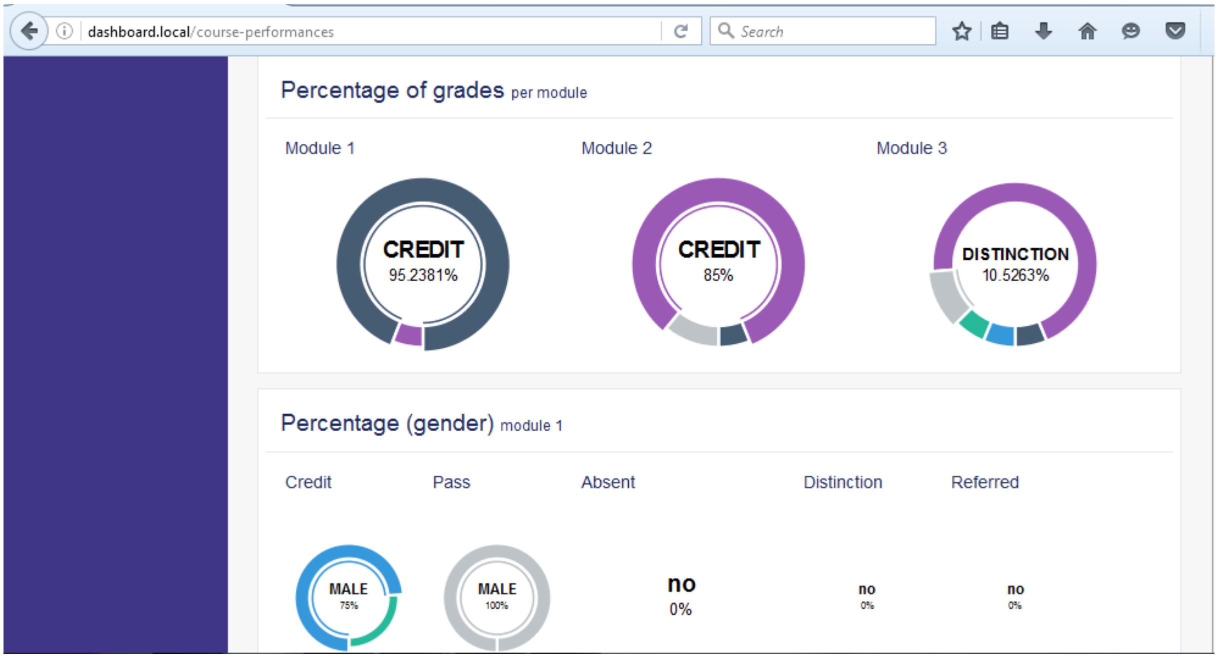


Figure 5.3-3: Course performance details showing percentage score of students for each grade in the given module

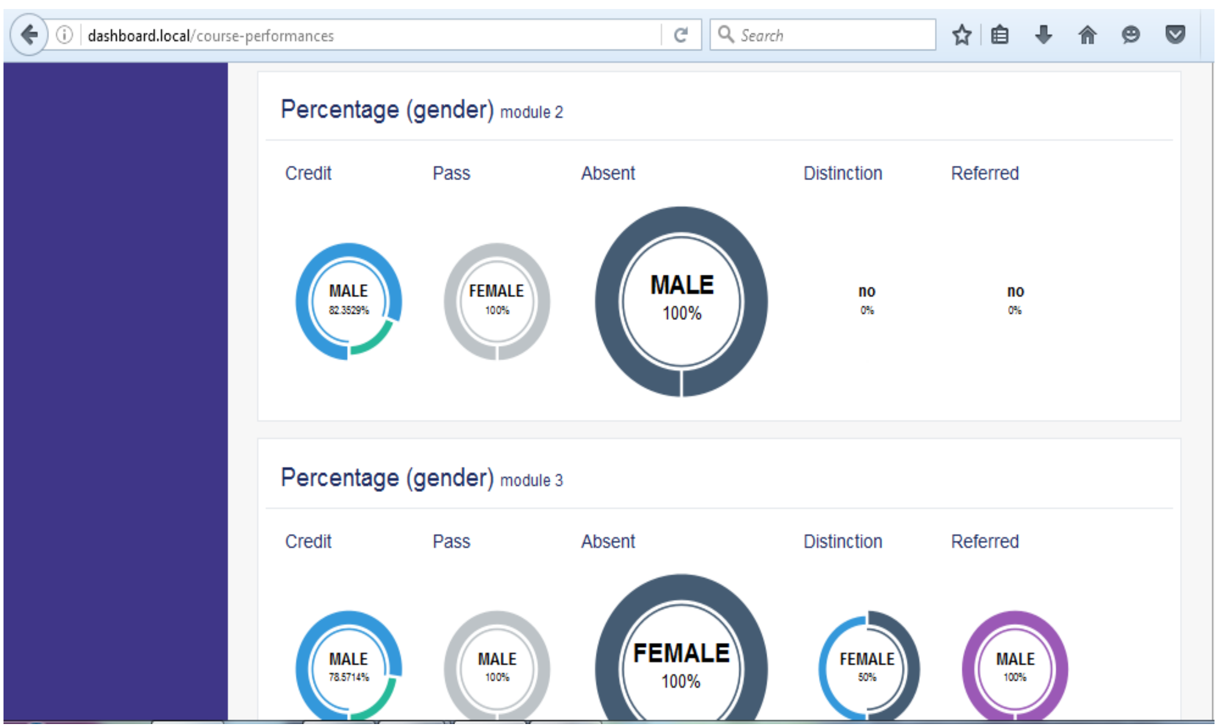


Figure 5.3-4: Course performance details showing percentage score of students by gender for each grade in the given module

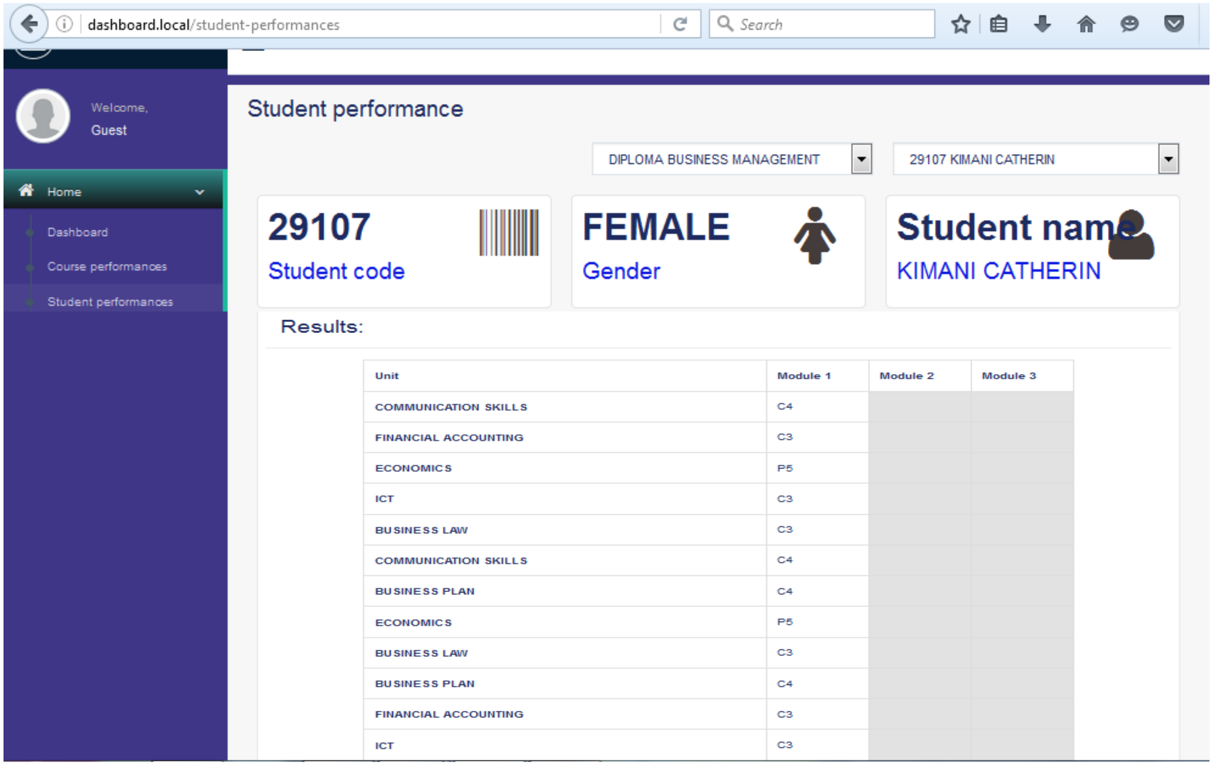


Figure 5.3-5: Individual student performance – module I

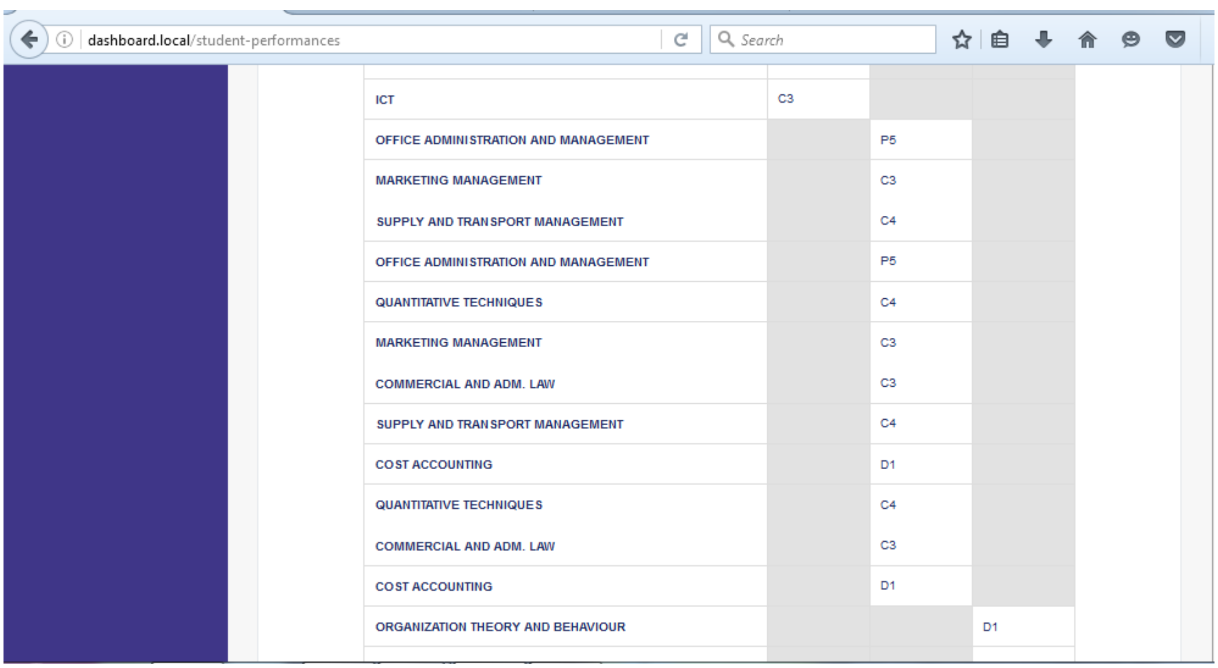


Figure 5.3-6: Individual student performance – module II

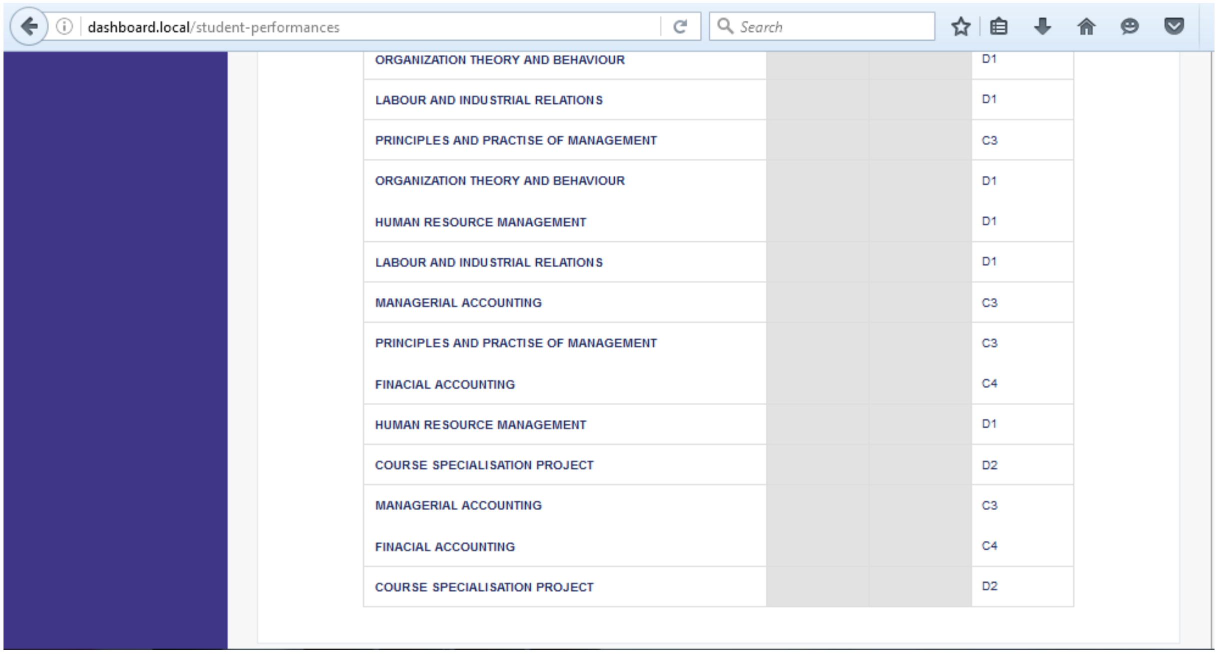


Figure 5.3-7: Individual student performance – module III

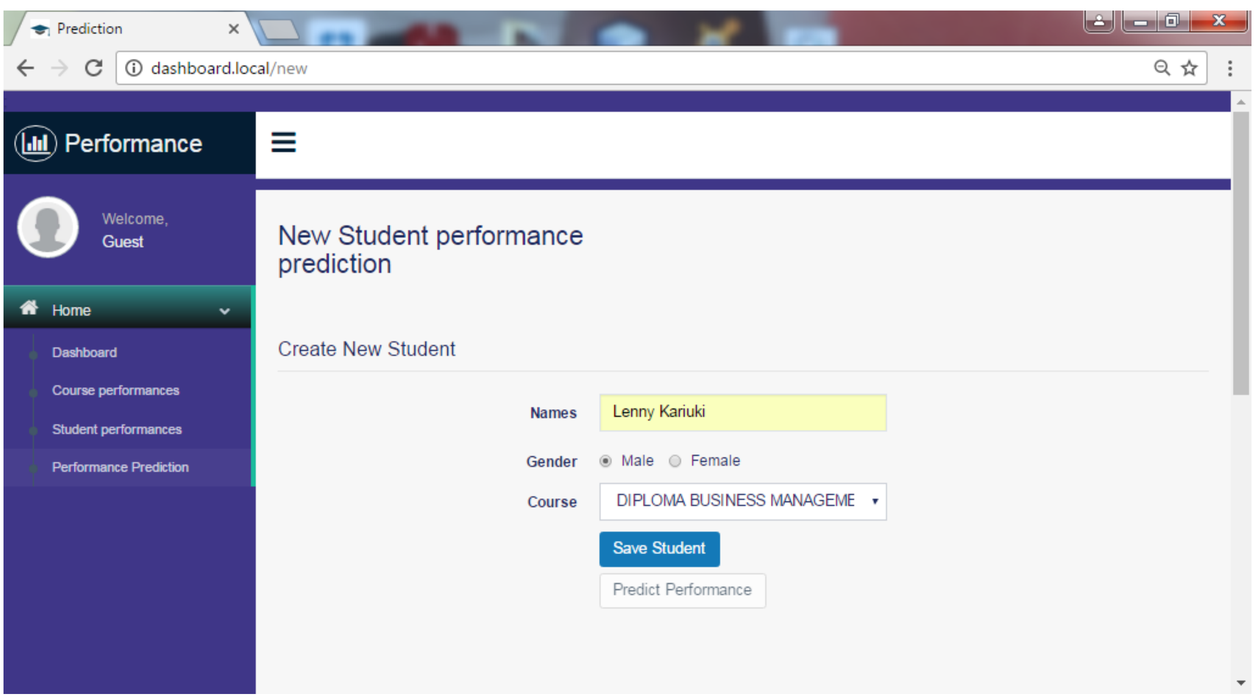


Figure 5.3-8: Prediction of student performance – key in details of student to predict performance for

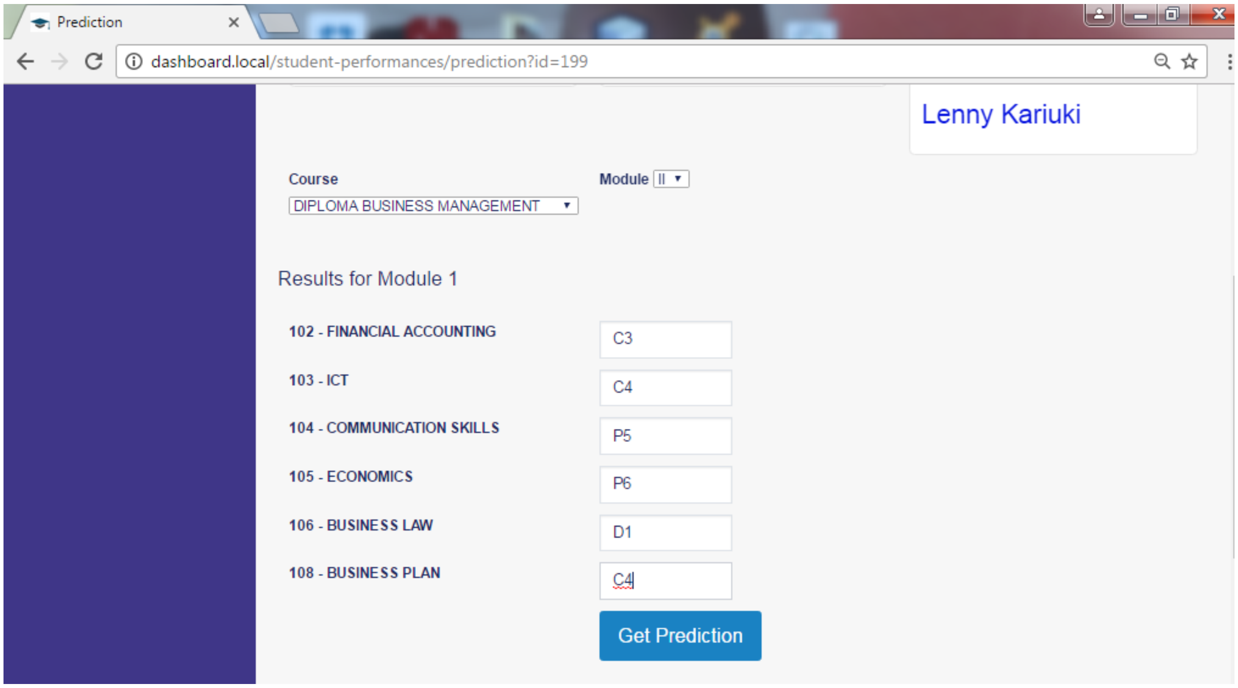


Figure 5.3-9: Predict student performance – details of previous module

Figure 5.3-10: Predicted results

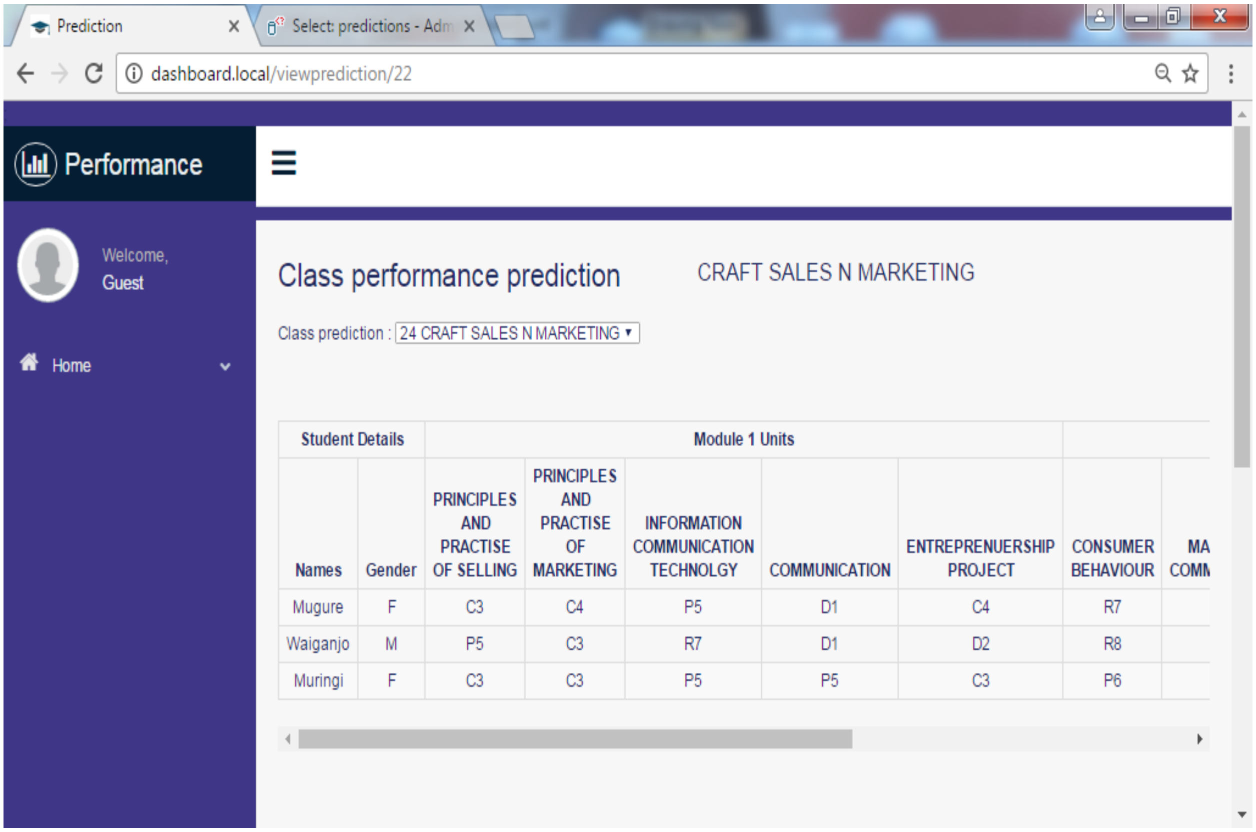


Figure 5.3-21: Group prediction of an entire class

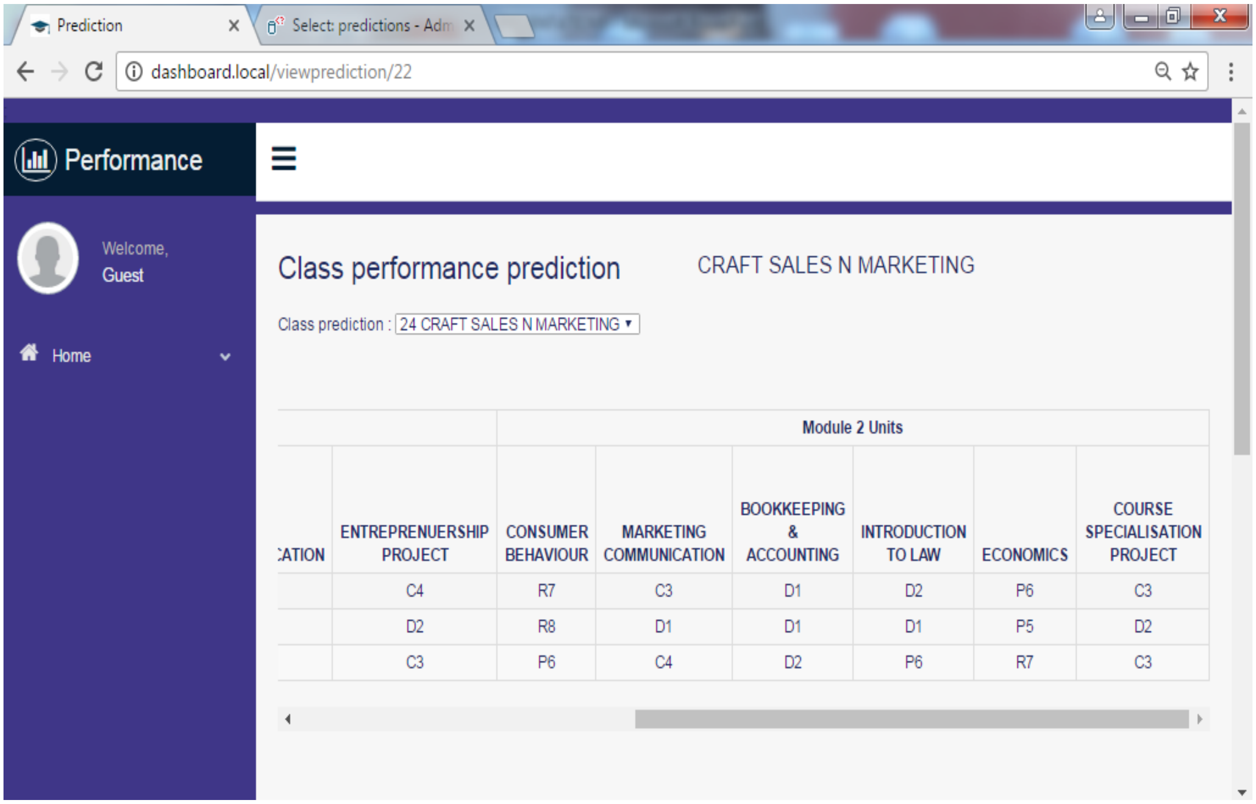


Figure 5.3-32: Continued group prediction

5.4 System testing

The types of testing done were unit, integration and acceptance testing. The system was tested using dummy and real data. This is illustrated using screen shots in appendix VII.

Unit testing

In unit testing the modules of the system were tested as they were developed. This was done in parallel with the coding process.

Integration testing

Integration testing was also done where the different modules were combined and tested as a whole system. It was meant to uncover errors associated with interfacing.

Acceptance testing

Acceptance testing was done where the system was presented to the system users where they checked if all their system requirements and functionalities were met. The changes they recommended were implemented in the prototype until it was refined to their expectations.

The areas addressed during acceptance testing were on:

- Ease of using the system. The users said the system was easy to use as it did not require re training on the issue of using computers. The menus and screens were arranged in a logical way that was consistent too. The interface flowed well and command names were easy to understand too.
- The errors that occurred were easy to recover from.
- Duration it took to key in data and do prediction was within the expected range. The users were impressed that the time it took to key in data and do prediction was within the expectation. It didn't take a long time.

6 Chapter Six: Discussion, Conclusion and Recommendation

6.1. Discussion

The development of the learner analytics dashboard enable student performance to be predicted, probable recommendations for that performance are given and this ensures students can be able to foretell their results and this gives them room for improvement. The students can know the weak areas they should be able to improve on. The lecturers are also able to assist their students as they are able to track on their performances. They will now be able to grade their students on the merit of performers, average performers and poor performers; advice and guide them accordingly. The teachers will also be able to reflect on the impact of their teaching method on the learning behavior and performance of their students, and draw first conclusions about the effectiveness of their teaching. The administration will be able to know the right administrative measures to put in place like student recruitment policies, adjusting course planning and determining hiring needs.

Some of the factors that have been identified to affect student performances can be both internal and external. Internal factors include: unfavorable learning environment; poorly equipped libraries; lack of e-facilities in the library; method of lecture delivery; technology used in classrooms; class attendance; overcrowded lecture rooms therefore students attend fewer sessions; exposure to practical work experience; conditions of lecture rooms; unavailability of recommended texts; late provision of reading/reference materials by lecturers; method of collating and accessing semester results; poor lecturer/student relationship; teacher student ratio where one teacher attends to many students; student learning style – actively/not actively engaged in the learning process; interruption of electricity supply; poor access to internet facilities; incessant strike and closure of school; poor accommodation (hostel) facilities and overcrowded exam time table.

External factors include: extracurricular activities whereby some students give them a priority at the expense of academics; peer influence – positive/negative; family problems; work and financial – some students get employed while still learning so they are not able to balance between class work and job; social and other problems; gender and age differences; level of parental education – high chances are if parents are educated their children will achieve more academically and vice versa; and guidance from parents and lecturers.

The learner analytics dashboard provides an interactive system that will help students; lecturers and administration make informed decisions as concern the student affairs. Depending on individual student performance, recommendations will be made depending on that performance which can then be looked in at detail to be able to assist the affected student.

The system will take in students details and will then be able to display students performance details in modules I, II and III showing the results of each unit in the different modules. It will display performances for the different courses where bar charts and pie charts are used to display number of students in each course, results analysis by grade per module (pass, credit, distinction, refer or absent) and also the same analysis by gender. To predict a student's performance, a user is promoted to key in details of the previous module which then generates results of current module and gives probable recommendation for that performance.

6.2. Conclusion

The study achieved its objectives as discussed below:

6.2.1. Objective one

The following conclusion can be drawn to address the objective: t**o establish the current methods of predicting performance in technical institutions.**

Currently there is no system in place to predict student's performance. The lecturers are only able to tell how a student would perform only if they are able to maintain their academic progress record that shows continuous assessment tests and end of term examination results.

They would also tell a student's performance based on class attendance and behavior in and out of the classroom.

6.2.2. Objective two

The following conclusion can be drawn to address the objective: **to determine the requirements for the dashboard system.**

The requirements needed for the development of the learner analytics dashboard include the student personal details like name, gender and student code/ admission number/index number; the different courses offered in the institution; different units done in the different courses offered in module I, II and III for diploma courses and module I and II for craft courses.

Also required are the factors that affect student performances which include: unfavorable learning environment; poorly equipped libraries; lack of e-facilities in the library; method of lecture delivery; technology used in classrooms; class attendance; overcrowded lecture rooms therefore students attend fewer sessions; exposure to practical work experience; conditions of lecture rooms; unavailability of recommended texts; late provision of reading/reference materials by lecturers; method of collating and accessing semester results; poor lecturer/student relationship; teacher student ratio where one teacher attends to many students; student learning style – actively/not actively engaged in the learning process; interruption of electricity supply; poor access to internet facilities; incessant strike and closure of school; poor accommodation (hostel) facilities and overcrowded exam time table; extracurricular activities whereby some students give them a priority at the expense of academics; peer influence – positive/negative; family problems; work and financial – some students get employed while still learning so they are not able to balance between class work and job; social and other problems; gender and age differences; level of parental education – high chances are if parents are educated their children will achieve more academically and vice versa; and guidance from parents and lecturers.

6.2.3. Objective three

The following conclusion can be drawn to address the objective: **to design a learner analytics dashboard.**

Input and output interfaces were designed and the system users agreed with them.

6.2.4. Objective four

The following conclusion can be drawn to address the objective: **to implement the learner analytics dashboard for use in technical institutions.**

A learner analytics dashboard that can predict student performances in technical institutions has been successfully developed.

6.2.5. Objective five

The following conclusion can be drawn to address the objective: **to test and validate the learner analytics dashboard.**

The learner analytics dashboard was tested and validated using dummy and real data. This is illustrated as shown in appendix VII.

6.3. Recommendation

The study recommends the system to be improved to run on a mobile platform. This would provide convenience of use as students, lecturers and administration are using smart phones.

References

Ayash, M. M. (2016). *Research Methodologies in Computer Science and Information Systems.* [pdf] Available at: http://www.ptcdb.edu.ps/ar/sites/default/files/%D9%88%D8%B1%D9%82%D8%A9%20%D9%85%D9%87%D9%86%D8%AF%20%D8%B9%D9%8A%D8%A7%D8%B4. [Accessed May 3, 2016].

Basu, R., Archer, N., & Mukherjee, B. (2012). *Intelligent decision support in healthcare.* [online] Available at: http://analytics-magazine.org/intelligent-decision-support-in-healthcare/ [Accessedd July 20, 2015.

Bienkowski, M., Feng, M., & Means, B. (2012). *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief.* [pdf] Available at: https://tech.ed.gov/wp-content/uploads/2014/03/edm-la-brief.pdf Accessed April 5, 2016.

Chatti, M., Dyckhoff, A., Schroeder, U., & Thüs, H. (2012). A Reference Model for Learning Analytics. *International Journal of Technology Enhanced Learning* ,[online], p. 5-11. Available at: http://www.thues.com/upload/pdf/2012/CDST12_IJTEL.pdf [Accessed January 10, 2016].

Dennis, A., Wixom, B. H., & Roth, R. M. (2012). *System Analysis and Design.* United States of America: John Wiley & Sons, Inc.

Dey, I. (2005). *Qualitative Data Analysis A user-friendly guide for social scientists.* [pdf] London: Routledge. Available at: http://www.classmatandread.net/class/785wk3/Qualitative_data_analysis.pdf Accessed November 10, 2015].

Dietz-Uhler, B., & Hurn, J. E. (2013). Using Learning Analytics to Predict (and Improve) Student Success: A Faculty Perspective. *Journal of Interactive Online Learning* , [online], Volume 12, Number 1, p. 22. Available at: http://www.ncolr.org/jiol/issues/pdf/12.1.2.pdf [Accessed on April 5, 2016].

DTREG. (2016). *Multilayer Perceptron Neural Networks.* [online] DTREG predictive modelling software . Available at: https://www.dtreg.com/solution/view/21 [Accessed January 7, 2016].

Fakhroutdinov, K. (2014). *Activity Diagrams.* [online]. Availabe at: http://www.uml-diagrams.org/activity-diagrams.html [Accessed June 5, 2016].

Few, S. (2006, January). *Information Dashboard Design*. [ebook] Italy: O'Reilly. Available at: https://www.researchgate.net/publication/220630208_Business_Performance_Management_One_Truth [Accessed August 3, 2015].

Frolick, M. N., & Ariyachandra, T. R. (2014). *Business Performance Management: One Truth.* [online]. Available at: https://www.researchgate.net/publication/220630208_Business_Performance_Management_One_Truth [Accessed January 11, 2016].

GOK. (2013). *Technical and Vocational Education and Training Act No. 29 OF 2013.* [pdf]. Available at: https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/98807/117649/F1763223240/KEN98807.pdf [Accessed August 10, 2015].

Hall, O. P. (2002). *Artificial Intelligence Techniques Enhance Business Forecasts.* [online] Available at: https://gbr.pepperdine.edu/2010/08/artificial-intelligence-techniques-enhance-business-forecasts/ [Accessed June 2, 2016].

Haykin, S. (2005). *Neural Networks A comprehensive foundation second edition.* India: Pearson Education (Singapore) Pte Ltd.

Irick, M. L. (2008). Task-Technology Fit and Information Systems Effectiveness. *Journal of Knowledge Management Practice,[*online*]* Vol. 9, No. 3 . Available at: http://www.tlainc.com/articl165.htm [Accessed August 5, 2015].

Jacobson, I., Spence, I., & Bittner, K. (2011)., *USE-CASE 2.0 The Guide to Succeeding with Use Cases.*[online] Available at: http://www.ivarjacobson.com/download.ashx?id=1282 [Accessed March 18, 2016].

Karimi, A., & Sullivan, E. (2013). *Student Success Dashboard at California State University, Fullerton.* [pdf]. Available at: http://www.fullerton.edu/analyticalstudies/presentations/CSRDE2013_Dash_karimi_sullivan.pdf [Accessed August 4, 2015].

Klipfolio. (2015). *Dashboard Examples and Templates.* [online] Available at: http://www.klipfolio.com/resources/dashboard-examples [Accessed July 27, 2015].

Lenz, G., & Moeller, T. (2008). *Choosing a Software Development Model.* [online] Availabble at: http://flylib.com/books/en/2.634.1.23/1/ [Accessed August 4, 2015].

Long, P. D., & Siemens, G. (2011). *Penetrating the Fog: Analytics in Learning and Education.* [online] Available at: http://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education [Accessed April 12, 2016].

McLeod, S. (2012). *Experimental Method.* [online] Available at: http://www.simplypsychology.org/experimental-method.html [Accessed September 20, 2016].

Minaei-Bidgoli, B., Kashy, D. A., Kortemeyer, G., & Punch, W. F. (2003). *Predicting Student Performance: An Application of Data Mining Methods With the Educational Web-Based System LON-CAPA.* [online] Available at: https://www.researchgate.net/publication/4054565_Predicting_student_performance_an_application_of_data_mining_methods_with_an_educational_Web-based_system [Accessed July 7, 2016].

NationalCouncilforLawReporting. (2013). *Technical and Vocational Education and Training Act No. 29 OF 2013. Nairobi, Kenya.* [pdf]. Available at: https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/98807/117649/F1763223240/KEN98807.pdf [Accessed April 2, 2016].

NewSchoolsVentureFund. (2008). *Performance Dashboards.* [pdf] Available at: http://www.newschools.org/files/PerformanceDashboards.pdf [Accessed August 3, 2015].

Ngerechi, J. (2003). Technical and Vocational Educatiion and Training in Kenya. *Conference on the Reform of Technica And Vocational Education and Training (TVET),* [online], p. 3-4. Available at: http://www.bota.org.bw/sites/default/files/TVET_in_kenya.pdf [Accessed August 20, 2015].

NottinghamTrentUniversity. (2014). *A Short Introduction to the NTU Student Dashboard.* [online] Available at: http://www.ntu.ac.uk/current_students/studying/student_dashboard/index.html [Accessed August 4, 2015].

Parker, R. (2008). *Performance Management Scorecards and Dashboards for IT Operations Data.* [online] Available at: https://www.PerformanceMgmtITOperation [Accessed August 3, 2015].

Ranjan, J. (2009). Business Intelligence: Concepts, Components,Techniques and Benefits. *Journal of Theoretical and Applied Information Technology* ,[online] p. 65. Available at: http://www.jatit.org/volumes/research-papers/Vol9No1/9Vol9No1.pdf [Accessed August 10, 2015].

Roberts, E. (2000). *Neural networks* [online] Available at: http://www-cs-faculty.stanford.edu/~eroberts//courses/soco/projects/2000-01/neural-networks/Architecture/feedforward.html [Accessed December 11, 2015].

Rokach, L., & Maimon, O. (2005). Decision Trees. In *Data Mining and Knowledge Discovery Handbook* (pp. 165-166). New York, USA: Springer.

Rouse, M. (2015). *Business Intelligence (BI).* [online] Available at: http://searchdatamanagement.techtarget.com/definition/business-intelligence [Accessed June 25, 2015].

Rouse, M. (2007). *Software Requirements Specification (SRS).* [online] Available at: http://searchsoftwarequality.techtarget.com/definition/software-requirements-specification [Accessed May 15, 2016].

Rumbaugh, J., Jacobson, I., & Booch, G. (2004). *The Unified Modelling LanguageReference Manual*. [pdf] Available at: https://www.utdallas.edu/~chung/Fujitsu/UML_2.0/Rumbaugh--UML_2.0_Reference_CD.pdf [Accessed August 1, 2016].

Shahir, A. M., Husain, W., & Rashid, N. A. (2015). A Review on Predicting Student's Performance using Data Mining Techniques. *The Third Information Systems International Conference*, [online] p. 414 – 422. Available at: http://www.sciencedirect.com/science/article/pii/S1877050915036182 [Accessed September 15, 2016]

Sherrel, L. (2013). *Evolutionary Prototyping*. [online] Available at: http://link.springer.com/referenceworkentry/10.1007%2F978-1-4020-8265-8_201039 [Accessed April 22, 2016].

Sommerville, I. (2004). *Software Requirements.* [pdf] Available at: http://www.dcs.fmph.uniba.sk/~rjasko/pts/docs/ch6.pdf [Accessed May 11, 2016].

Stanley, G. (2011). *Contemporary Analysis Predictive Analytics.* [online] Available at: http://canworksmart.com/three-types-of-dashboards/ [Accessed August 3, 2015].

Stapel, M., Zheng, Z., & Pinkwart, N. (2016). An Ensemble Method to Predict Student Performance in an Online Math Learning Environment. *Proceedings of the 9th International Conference on Educational Data Mining*, (pp. 231-238). Available at: http://www.educationaldatamining.org/EDM2016/proceedings/paper_62.pdf [Accessed on September 25, 2016]

Swanson, R. A., & Chermack, T. J. (2013). *Theory Building in Applied Disciplines.* [pdf] Available at: https://www.bkconnection.com/static/Theory_Building_EXCERPT.pdf
[Accessed October 2, 2016].

Sweeney, M., Rangwala, H., Lester, J., & Johri, A. (2016, April). *Next-Term Student Performance Prediction:A Recommender Systems Approach.* [pdf] Available at: https://arxiv.org/pdf/1604.01840.pdf [Accessed September 2, 2016].

Tableau. (2016). *Five Dashboards Improving Student Achievement.* [online] Available at: http://www.tableau.com/sites/default/files/media/whitepaper_student_achievement.pdf [Accessed September 15, 2016].

Thai-Nghe, N., Drumond, L., Krohn-Grimberghe, A., & Schmidt-Thieme, L. (2010). Recommender System for Predicting Student Performance. *1st Workshop on Recommender Systems for Technology Enhanced Learning* [Online] p. 2811-2819. Available at: http://ac.els-cdn.com/S1877050910003194/1-s2.0-S1877050910003194-main.pdf?_tid=f2079010-9128-11e6-b4f4-00000aab0f26&acdnat=1476351752_c610337d4fb292a6084675426e33db69 [Accessed September 15, 2016].

Velcu-Laitinen, O., & Yigitbasioglu, O. M. (2012). The Use of Dashboards in Performance Management: Evidence from Sales Managers. *The International Journal of Digital Accounting Research* , [online] Vol. 12, 2012, p. 39 - 58. Available at: http://www.uhu.es/ijdar/10.4192/1577-8517-v12_2.pdf [Accessed September 2, 2016]

Wang, T., & Mitrovic, A. (2002). *Using Neural Networks to Predict Student's Performance.* [online] Available at: https://www.researchgate.net/publication/221319131_Using_Neural_Networks_to_predict_Student's_Performance [Accessed September 22, 2016].

Zainal, Z. (2007). *Case study as a research method.* [pdf] Available at: http://psyking.net/htmlobj-3837/case_study_as_a_research_method.pdf [Accessed September 20, 2016].

Zhang, H. (2004). *The Optimality of Naive Bayes.* [pdf] Available at: http://www.cs.unb.ca/~hzhang/publications/FLAIRS04ZhangH.pdf [Accessed September 15, 2016].

Appendix I: Interview schedule for administration

I. Opening

a. (Establish rapport by shaking hands). My name is ______________ a student of KCA University.

b. (Purpose of the interview) I would like to ask you some questions about the institution.

c. (Motivation) I hope to use this information to design a dashboard system that can aid in predicting students performance.

d. (Time Line) The interview should take about 10 minutes. I hope you will respond to the questions within this time.

e. (Transition):

II Body

Topic – Course details

- How many departments are there in the institutions?
- What courses are offered in this departments?
- Are they diploma, certificate or both?
- What is the entry point of certificate/diploma course?
- How long does it take to complete certificate/diploma course?
- Which are the examining bodies for those examinations?
- How are the students examination records maintained?

III Closing

a. (Summary)

b. (Maintain Rapport) I appreciate the time you took for this interview. Is there anything else you think would be helpful for me to know so that I can successfully model the system?

c. (Action to be taken) Would it be alright to call you later if I have any more questions or to clarify anything that would not be clear? Thank you.

Appendix II: Interview schedule for lecturers

I. Opening

a. (Establish rapport by shaking hands). My name is ______________ a student of KCA University.

b. (Purpose of the interview) I would like to ask you some questions about the institution.

c. (Motivation) I hope to use this information to design a dashboard system that can aid in predicting students performance.

d. (Time Line) The interview should take about 10 minutes. I hope you will respond to the questions within this time.

e. (Transition):

II Body

Topic – Student performance in exams

- Which department are you in?
- What responsibility do you have in the department?
- Which units are you handling?
- What factors affects how individual students perform?
- How often do you give out examinations to your students?
- How do you give feedback to your students?
- How do you maintain the examination records for the students?
- What criteria do you use to rate your students? Does your rating work – how accurate is it?

III Closing

a. (Summary)

b. (Maintain Rapport) I appreciate the time you took for this interview. Is there anything else you think would be helpful for me to know so that I can successfully model the system?

c. (Action to be taken) Would it be alright to call you later if I have any more questions or to clarify anything that would not be clear? Thank you.

Appendix III: Interview schedule for students

I. Opening

a. (Establish rapport by shaking hands). My name is ______________ a student of KCA University.

b. (Purpose of the interview) I would like to ask you some questions about the institution.

c. (Motivation) I hope to use this information to design a dashboard system that can aid in predicting students performance.

d. (Time Line) The interview should take about 10 minutes. I hope you will respond to the questions within this time.

e. (Transition):

II Body

Topic – Student performance in exams

- What grade did you score in KCSE?
- Which department are you in?
- What course are you taking?
- Which module are you in currently?
- How has your performance been for the course work?
- How was your performance in the module examination compared to the course work?
- What factors can you attribute to your performance?
- Do you always get feedback for any internal examinations done?
- How do you maintain your results?

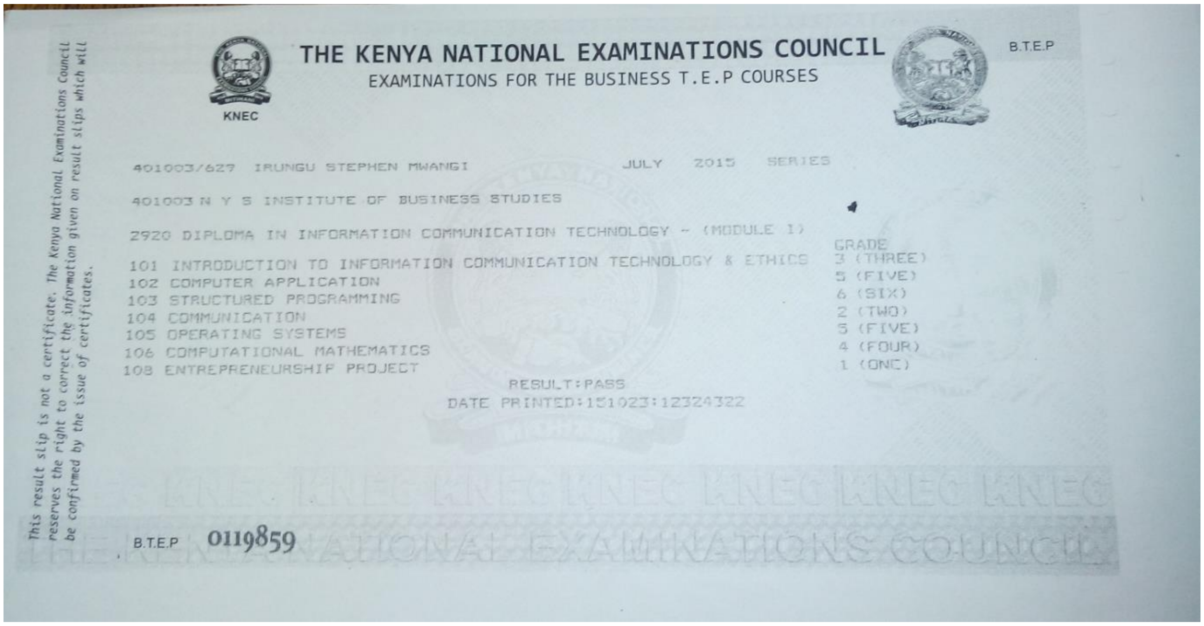
III Closing

a. (Summary)

b. (Maintain Rapport) I appreciate the time you took for this interview. Is there anything else you think would be helpful for me to know so that I can successfully model the system?

c. (Action to be taken) Would it be alright to call you later if I have any more questions or to clarify anything that would not be clear? Thank you.

Appendix IV: Module Result Slips



This result slip is not a certificate. The Kenya National Examinations Council reserves the right to correct the information given on result slips which will be confirmed by the issue of certificates.

KNEC

THE KENYA NATIONAL EXAMINATIONS COUNCIL

EXAMINATIONS FOR THE BUSINESS T.E.P COURSES

B.T.E.P

401003/627 IRUNGU STEPHEN MWANGI JULY 2015 SERIES

401003 N Y S INSTITUTE OF BUSINESS STUDIES

2920 DIPLOMA IN INFORMATION COMMUNICATION TECHNOLOGY - (MODULE I)

	GRADE
101 INTRODUCTION TO INFORMATION COMMUNICATION TECHNOLOGY & ETHICS	3 (THREE)
102 COMPUTER APPLICATION	5 (FIVE)
103 STRUCTURED PROGRAMMING	6 (SIX)
104 COMMUNICATION	2 (TWO)
105 OPERATING SYSTEMS	5 (FIVE)
106 COMPUTATIONAL MATHEMATICS	4 (FOUR)
108 ENTREPRENEURSHIP PROJECT	1 (ONE)

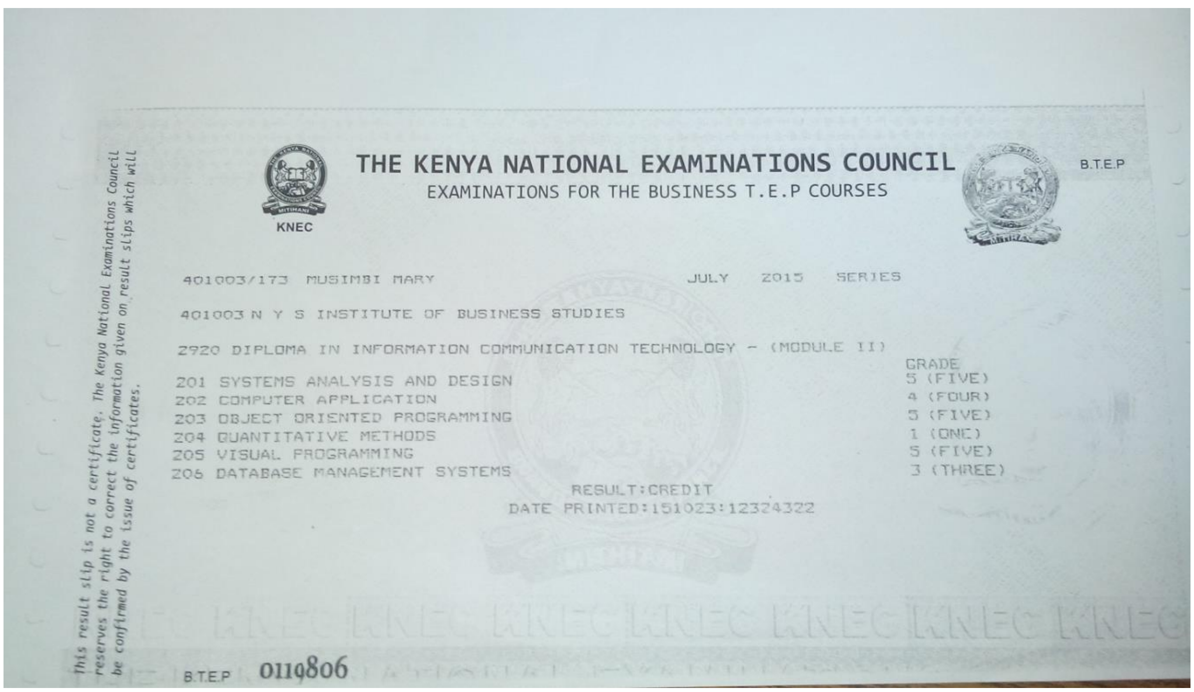
RESULT:PASS

DATE PRINTED:151023:12324322

KNEC KNEC KNEC KNEC KNEC KNEC KNEC

THE KENYA NATIONAL EXAMINATIONS COUNCIL

B.T.E.P 0119859



This result slip is not a certificate. The Kenya National Examinations Council reserves the right to correct the information given on result slips which will be confirmed by the issue of certificates.

KNEC

THE KENYA NATIONAL EXAMINATIONS COUNCIL

EXAMINATIONS FOR THE BUSINESS T.E.P COURSES

B.T.E.P

401003/173 MUSIMBI MARY JULY 2015 SERIES

401003 N Y S INSTITUTE OF BUSINESS STUDIES

2920 DIPLOMA IN INFORMATION COMMUNICATION TECHNOLOGY - (MODULE II)

	GRADE
201 SYSTEMS ANALYSIS AND DESIGN	5 (FIVE)
202 COMPUTER APPLICATION	4 (FOUR)
203 OBJECT ORIENTED PROGRAMMING	5 (FIVE)
204 QUANTITATIVE METHODS	1 (ONE)
205 VISUAL PROGRAMMING	5 (FIVE)
206 DATABASE MANAGEMENT SYSTEMS	3 (THREE)

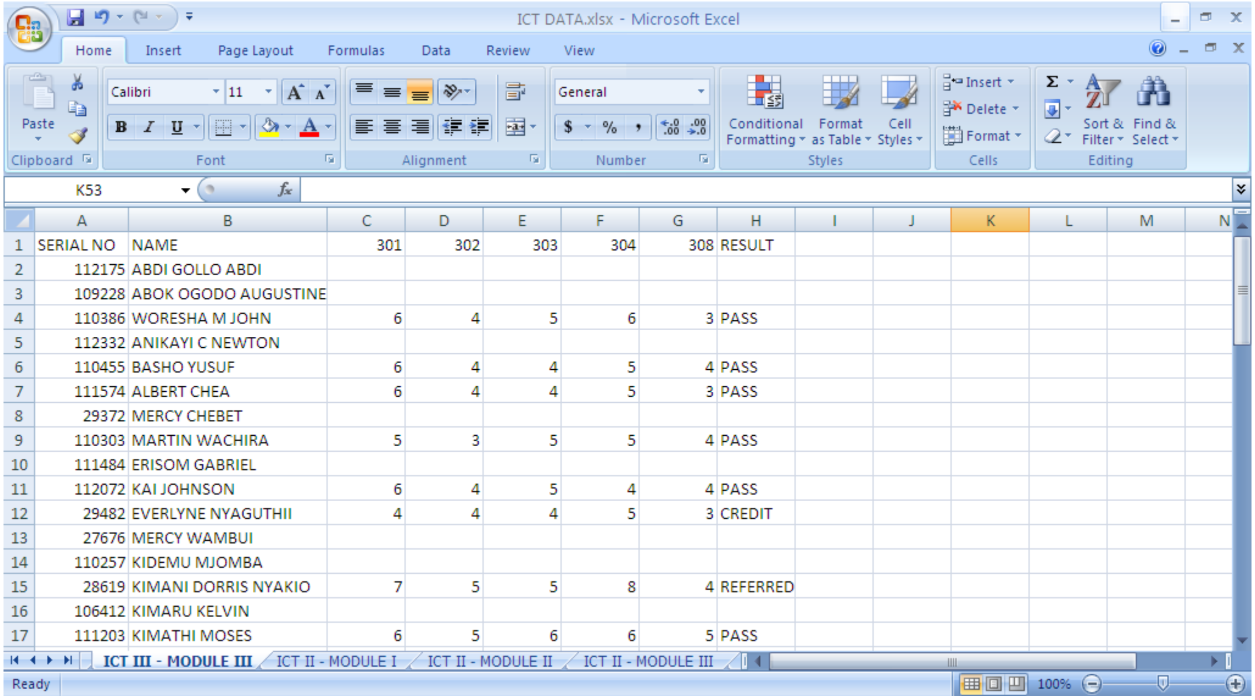
RESULT:CREDIT

DATE PRINTED:151023:12324322

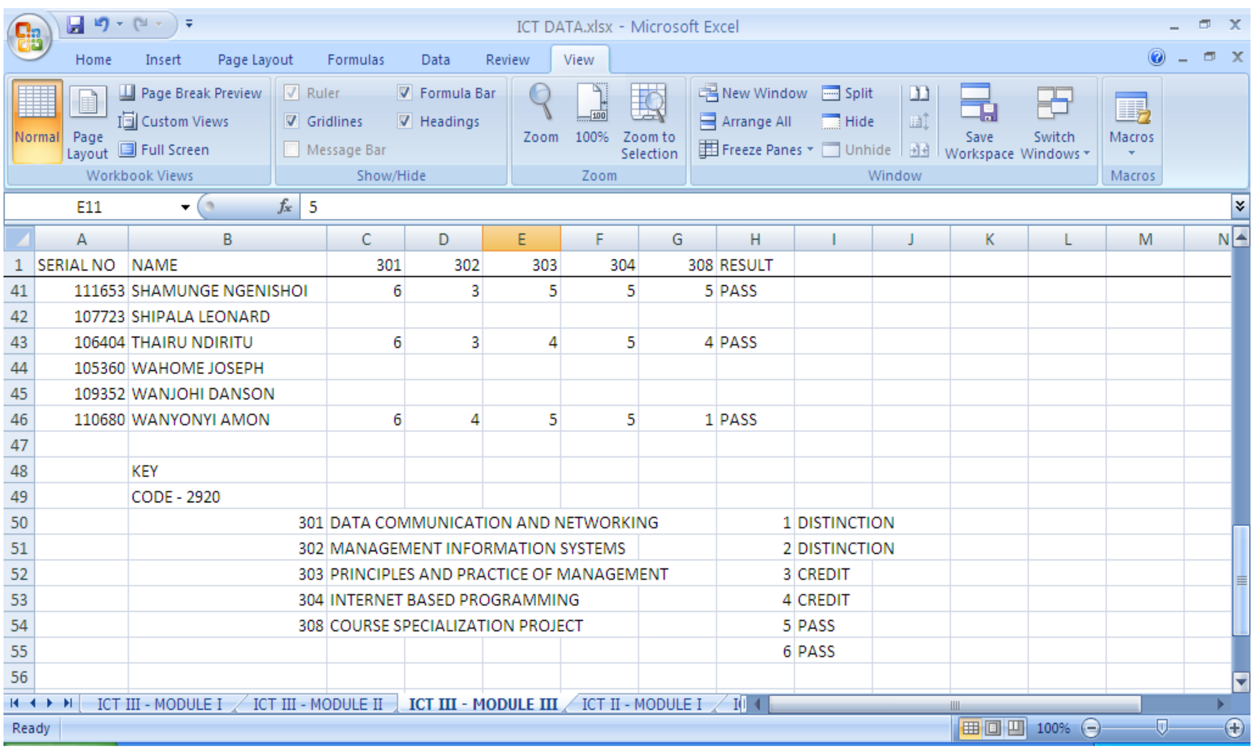
KNEC KNEC KNEC KNEC KNEC KNEC KNEC

B.T.E.P 0119806

Appendix V: Diploma ICT Module III Data

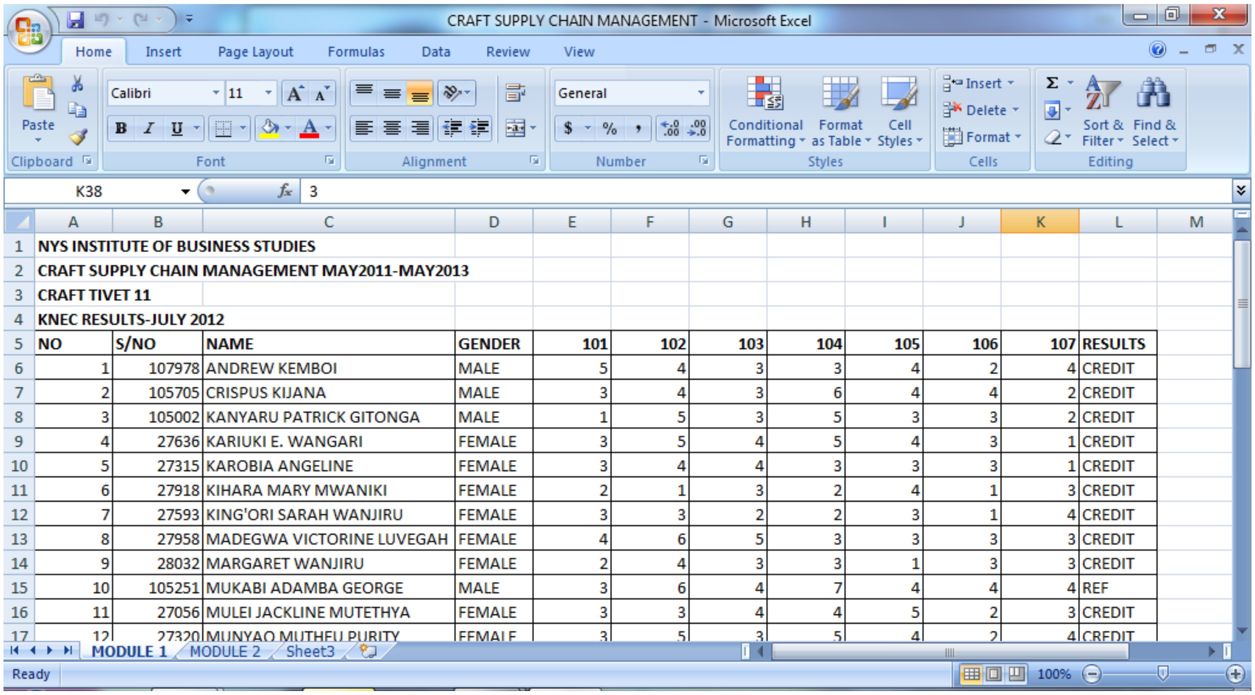


	A	B	C	D	E	F	G	H
1	SERIAL NO	NAME	301	302	303	304	308	RESULT
2	112175	ABDI GOLLO ABDI						
3	109228	ABOK OGODO AUGUSTINE						
4	110386	WORESHA M JOHN	6	4	5	6	3	PASS
5	112332	ANIKAYI C NEWTON						
6	110455	BASHO YUSUF	6	4	4	5	4	PASS
7	111574	ALBERT CHEA	6	4	4	5	3	PASS
8	29372	MERCY CHEBET						
9	110303	MARTIN WACHIRA	5	3	5	5	4	PASS
10	111484	ERISOM GABRIEL						
11	112072	KAI JOHNSON	6	4	5	4	4	PASS
12	29482	EVERLYNE NYAGUTHII	4	4	4	5	3	CREDIT
13	27676	MERCY WAMBUI						
14	110257	KIDEMU MJOMBA						
15	28619	KIMANI DORRIS NYAKIO	7	5	5	8	4	REFERRED
16	106412	KIMARU KELVIN						
17	111203	KIMATHI MOSES	6	5	6	6	5	PASS

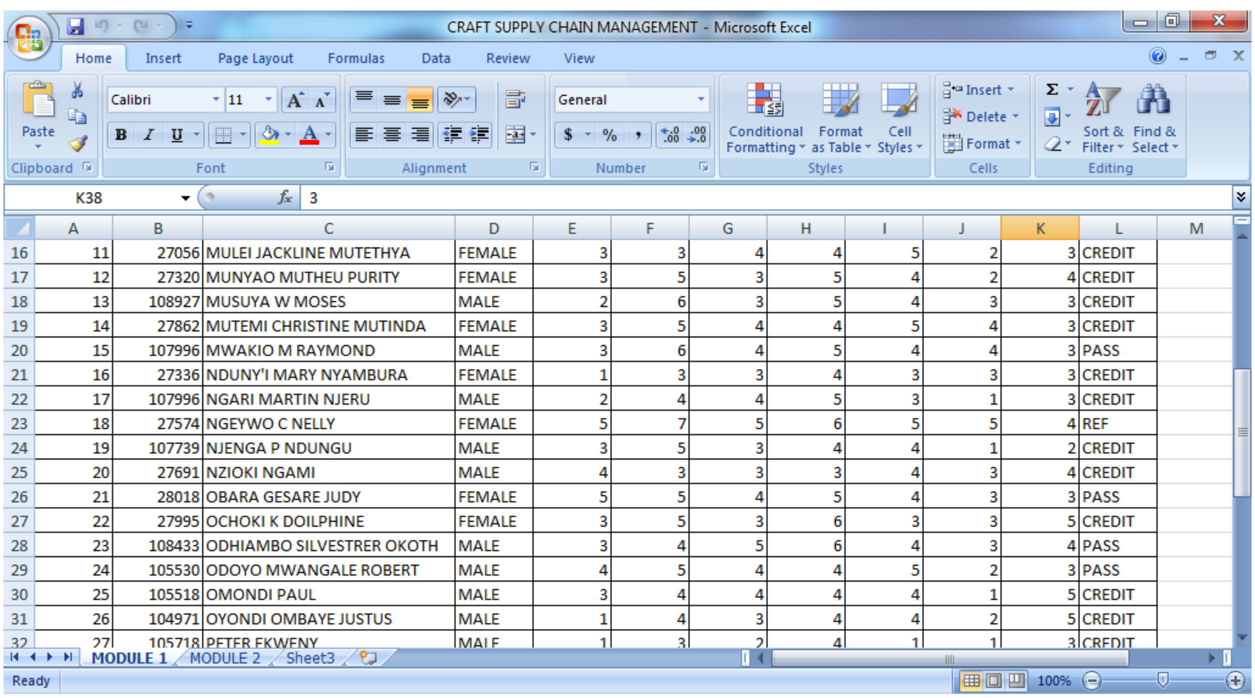


	A	B	C	D	E	F	G	H	I
1	SERIAL NO	NAME	301	302	303	304	308	RESULT	
41	111653	SHAMUNGE NGENISHOI	6	3	5	5	5	PASS	
42	107723	SHIPALA LEONARD							
43	106404	THAIRU NDIRITU	6	3	4	5	4	PASS	
44	105360	WAHOME JOSEPH							
45	109352	WANJOHI DANSON							
46	110680	WANYONYI AMON	6	4	5	5	1	PASS	
47									
48		KEY							
49		CODE - 2920							
50		301	DATA COMMUNICATION AND NETWORKING					1	DISTINCTION
51		302	MANAGEMENT INFORMATION SYSTEMS					2	DISTINCTION
52		303	PRINCIPLES AND PRACTICE OF MANAGEMENT					3	CREDIT
53		304	INTERNET BASED PROGRAMMING					4	CREDIT
54		308	COURSE SPECIALIZATION PROJECT					5	PASS
55								6	PASS
56									

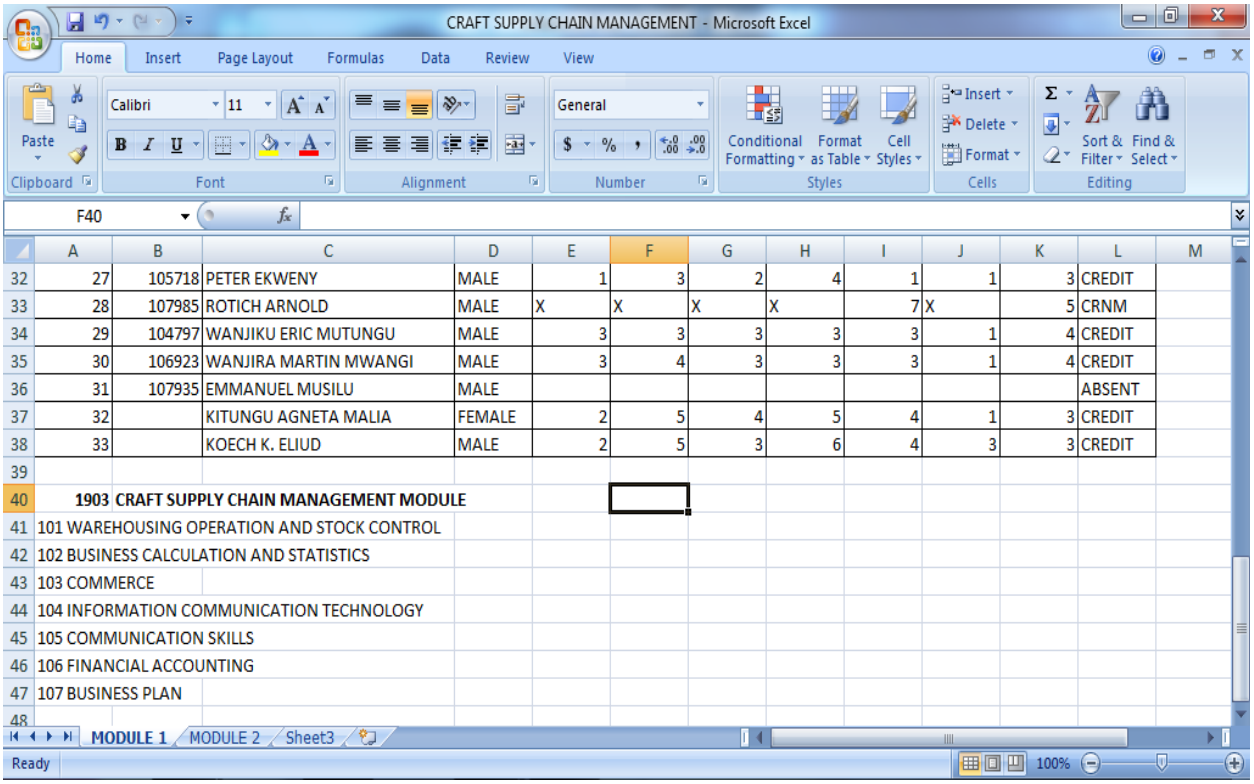
Appendix VI: Craft Supply Chain Management Module I Data



	A	B	C	D	E	F	G	H	I	J	K	L
1	NYS INSTITUTE OF BUSINESS STUDIES											
2	CRAFT SUPPLY CHAIN MANAGEMENT MAY2011-MAY2013											
3	CRAFT TIVET 11											
4	KNEC RESULTS-JULY 2012											
5	NO	S/NO	NAME	GENDER	101	102	103	104	105	106	107	RESULTS
6	1	107978	ANDREW KEMBOI	MALE	5	4	3	3	4	2	4	CREDIT
7	2	105705	CRISPUS KIJANA	MALE	3	4	3	6	4	4	2	CREDIT
8	3	105002	KANYARU PATRICK GITONGA	MALE	1	5	3	5	3	3	2	CREDIT
9	4	27636	KARIUKI E. WANGARI	FEMALE	3	5	4	5	4	3	1	CREDIT
10	5	27315	KAROBIA ANGELINE	FEMALE	3	4	4	3	3	3	1	CREDIT
11	6	27918	KIHARA MARY MWANIKI	FEMALE	2	1	3	2	4	1	3	CREDIT
12	7	27593	KING'ORI SARAH WANJIRU	FEMALE	3	3	2	2	3	1	4	CREDIT
13	8	27958	MADEGWA VICTORINE LUVEGAH	FEMALE	4	6	5	3	3	3	3	CREDIT
14	9	28032	MARGARET WANJIRU	FEMALE	2	4	3	3	1	3	3	CREDIT
15	10	105251	MUKABI ADAMBA GEORGE	MALE	3	6	4	7	4	4	4	REF
16	11	27056	MULEI JACKLINE MUTETHYA	FEMALE	3	3	4	4	5	2	3	CREDIT
17	12	27320	MUNYAO MUTHEU PURITY	FEMALE	3	5	3	5	4	2	4	CREDIT

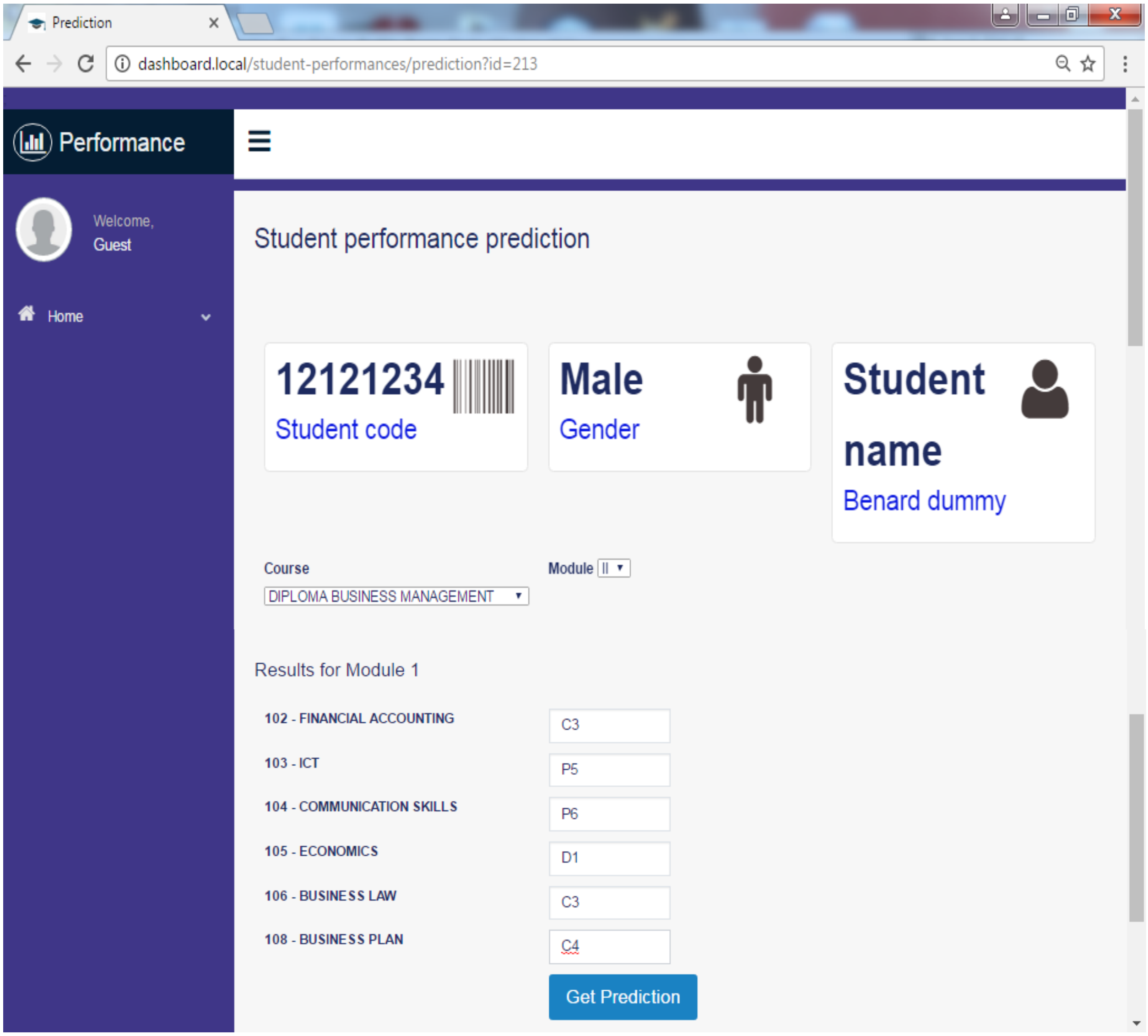


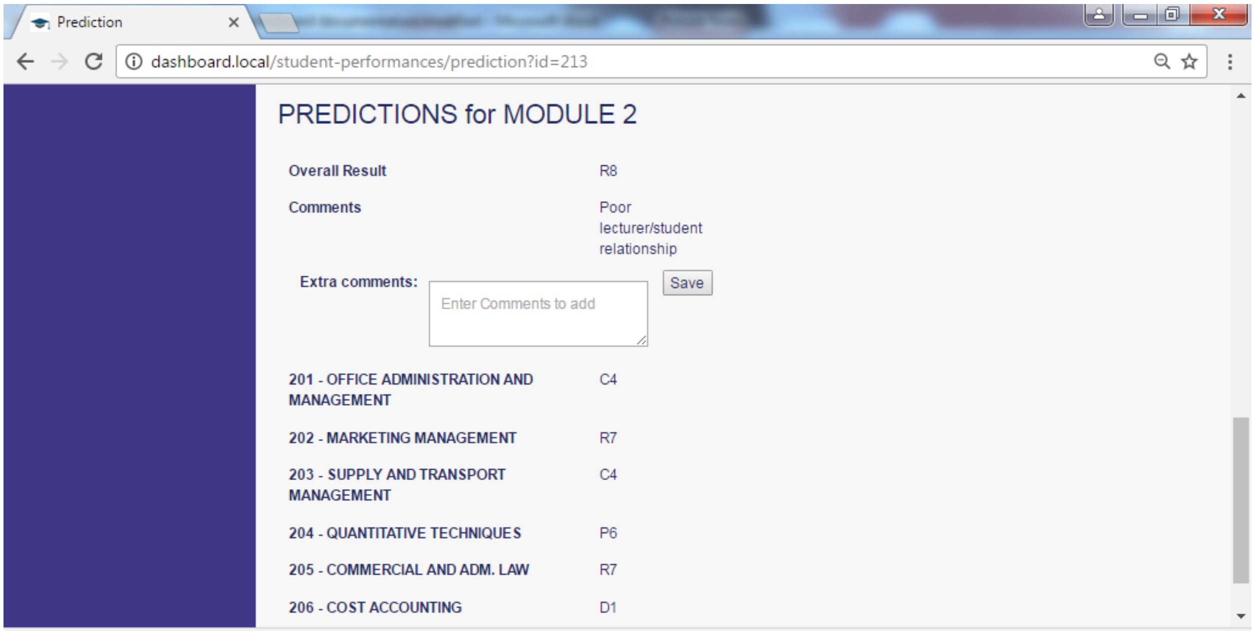
	A	B	C	D	E	F	G	H	I	J	K	L
16	11	27056	MULEI JACKLINE MUTETHYA	FEMALE	3	3	4	4	5	2	3	CREDIT
17	12	27320	MUNYAO MUTHEU PURITY	FEMALE	3	5	3	5	4	2	4	CREDIT
18	13	108927	MUSUYA W MOSES	MALE	2	6	3	5	4	3	3	CREDIT
19	14	27862	MUTEMI CHRISTINE MUTINDA	FEMALE	3	5	4	4	5	4	3	CREDIT
20	15	107996	MWAKIO M RAYMOND	MALE	3	6	4	5	4	4	3	PASS
21	16	27336	NDUNY'I MARY NYAMBURA	FEMALE	1	3	3	4	3	3	3	CREDIT
22	17	107996	NGARI MARTIN NJERU	MALE	2	4	4	5	3	1	3	CREDIT
23	18	27574	NGEYWO C NELLY	FEMALE	5	7	5	6	5	5	4	REF
24	19	107739	NJENGA P NDUNGU	MALE	3	5	3	4	4	1	2	CREDIT
25	20	27691	NZIOKI NGAMI	MALE	4	3	3	3	4	3	4	CREDIT
26	21	28018	OBARA GESARE JUDY	FEMALE	5	5	4	5	4	3	3	PASS
27	22	27995	OCHOKI K DOILPHINE	FEMALE	3	5	3	6	3	3	5	CREDIT
28	23	108433	ODHIAMBO SILVESTRER OKOTH	MALE	3	4	5	6	4	3	4	PASS
29	24	105530	ODOYO MWANGALE ROBERT	MALE	4	5	4	4	5	2	3	PASS
30	25	105518	OMONDI PAUL	MALE	3	4	4	4	4	1	5	CREDIT
31	26	104971	OYONDI OMBAYE JUSTUS	MALE	1	4	3	4	4	2	5	CREDIT
32	27	105718	PETER EKWENY	MALE	1	3	2	4	1	1	3	CREDIT

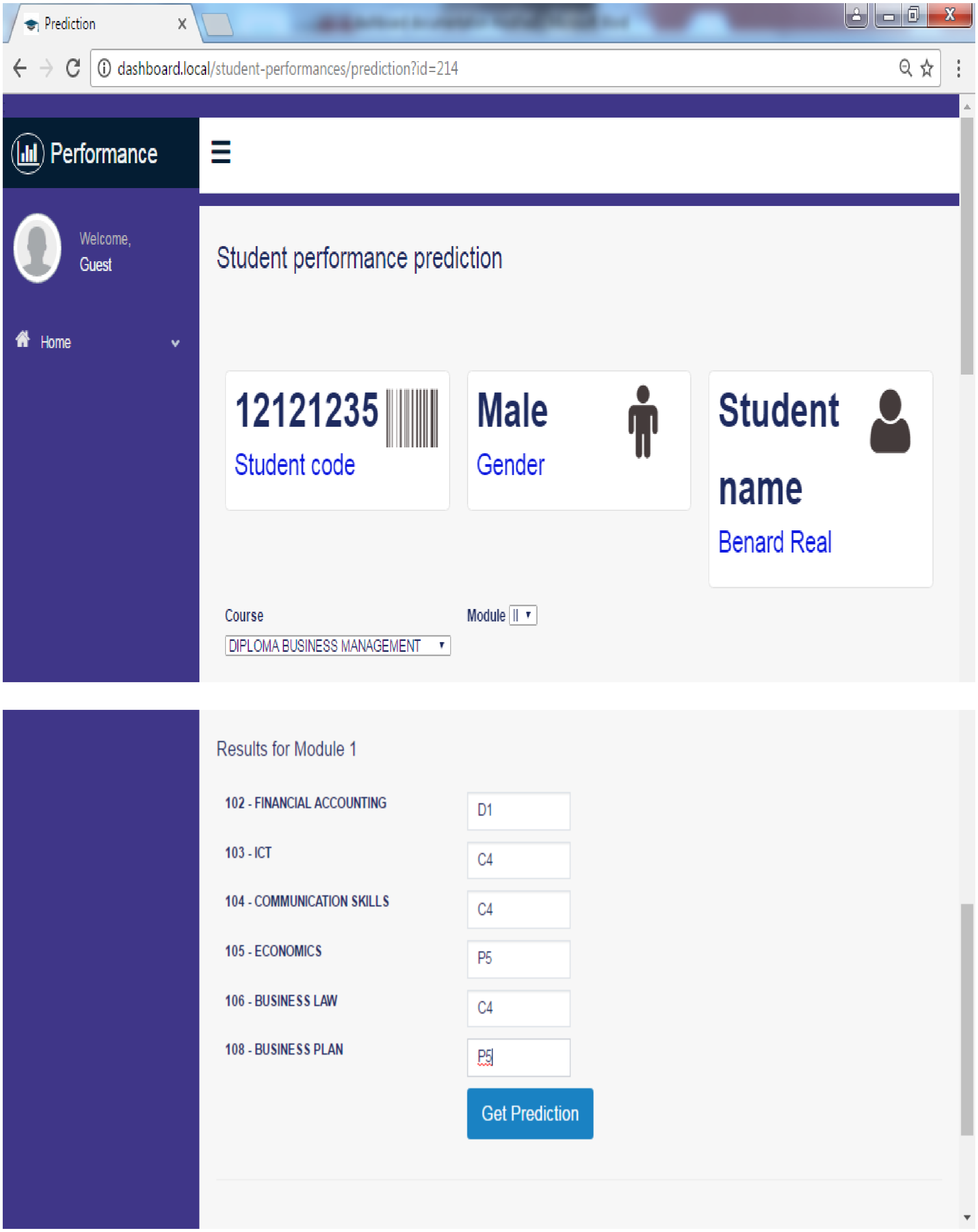


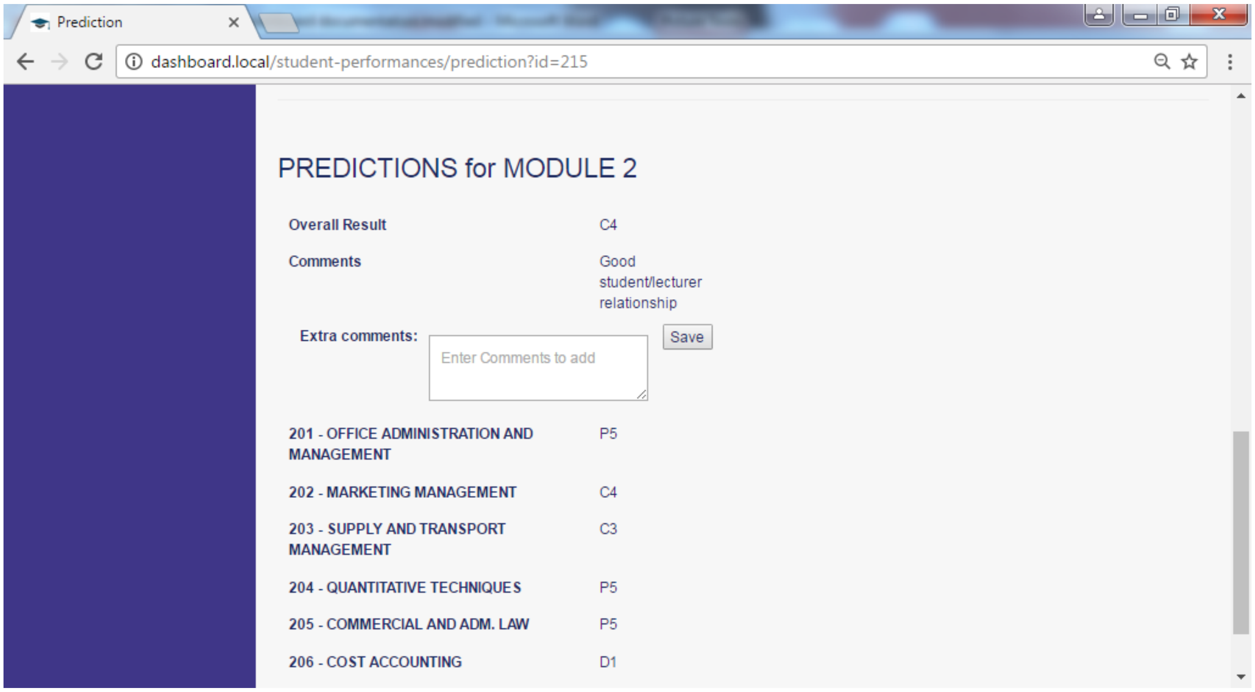
	A	B	C	D	E	F	G	H	I	J	K	L	M
32	27	105718	PETER EKWENY	MALE	1	3	2	4	1	1	3	CREDIT	
33	28	107985	ROTICH ARNOLD	MALE	X	X	X	X	7	X	5	CRNM	
34	29	104797	WANJIKU ERIC MUTUNGU	MALE	3	3	3	3	3	1	4	CREDIT	
35	30	106923	WANJIRA MARTIN MWANGI	MALE	3	4	3	3	3	1	4	CREDIT	
36	31	107935	EMMANUEL MUSILU	MALE								ABSENT	
37	32		KITUNGU AGNETA MALIA	FEMALE	2	5	4	5	4	1	3	CREDIT	
38	33		KOECH K. ELIUD	MALE	2	5	3	6	4	3	3	CREDIT	
39													
40	**1903**	**CRAFT SUPPLY CHAIN MANAGEMENT MODULE**											
41	101 WAREHOUSING OPERATION AND STOCK CONTROL												
42	102 BUSINESS CALCULATION AND STATISTICS												
43	103 COMMERCE												
44	104 INFORMATION COMMUNICATION TECHNOLOGY												
45	105 COMMUNICATION SKILLS												
46	106 FINANCIAL ACCOUNTING												
47	107 BUSINESS PLAN												

Appendix VII: Dummy and Real Data used in System Testing




Prediction
dashboard.local/student-performances/prediction?id=213
PREDICTIONS for MODULE 2
Overall Result
R8
Comments
Poor lecturer/student relationship
Extra comments:
Enter Comments to add
Save
201 - OFFICE ADMINISTRATION AND MANAGEMENT
C4
202 - MARKETING MANAGEMENT
R7
203 - SUPPLY AND TRANSPORT MANAGEMENT
C4
204 - QUANTITATIVE TECHNIQUES
P6
205 - COMMERCIAL AND ADM. LAW
R7
206 - COST ACCOUNTING
D1


Prediction
dashboard.local/student-performances/prediction?id=214
Performance
Welcome,
Guest
Home
Student performance prediction
12121235
Student code
Male
Gender
Student name
Benard Real
Course
DIPLOMA BUSINESS MANAGEMENT
Module II
Results for Module 1
102 - FINANCIAL ACCOUNTING
D1
103 - ICT
C4
104 - COMMUNICATION SKILLS
C4
105 - ECONOMICS
P5
106 - BUSINESS LAW
C4
108 - BUSINESS PLAN
P5
Get Prediction


Prediction
dashboard.local/student-performances/prediction?id=215
PREDICTIONS for MODULE 2
Overall Result
C4
Comments
Good student/lecturer relationship
Extra comments:
Enter Comments to add
Save
201 - OFFICE ADMINISTRATION AND MANAGEMENT
P5
202 - MARKETING MANAGEMENT
C4
203 - SUPPLY AND TRANSPORT MANAGEMENT
C3
204 - QUANTITATIVE TECHNIQUES
P5
205 - COMMERCIAL AND ADM. LAW
P5
206 - COST ACCOUNTING
D1

Appendix VII: Neural Network Code

```
public class Network {


   private String description;


   public void setDescription(String description) {

      this.description = description;

   }


   private final BasicNetwork network;

   private final int inputCount, outputCount;

   private final double input[][];

   private final double ideal_output[][];


   public Network(final int inputCount, final int outputCount, TrainingData trainingData) {

      this.inputCount = inputCount;

      this.outputCount = outputCount;

      network = new BasicNetwork();

      input = trainingData.getInput();

      ideal_output = trainingData.getOutput();

   }


   public void initialize() {

      network.addLayer(new BasicLayer(new ActivationSigmoid(), true, inputCount));
```

```
        network.addLayer(new BasicLayer(new ActivationSigmoid(), true, (inputCount > outputCount ?
inputCount : outputCount) + 1));
        network.addLayer(new BasicLayer(new ActivationSigmoid(), true, outputCount));
        network.getStructure().finalizeStructure();
        network.reset();
        train();
    }

    private void train() {
        NeuralDataSet trainingSet = new BasicNeuralDataSet(input, ideal_output);
        final Train train = new ResilientPropagation(network, trainingSet);
        int epoch = 1;

        do {

            train.iteration();
            if (epoch % 100 == 0) {
                System.out.println("Iteration #" + epoch + " Error:" + train.getError());
            }
            epoch++;
            if (train.getError() <= 0.0001) {
                break;
            } else if (epoch > 100000) {
                break;
```

```
        }
      } while (true);
      double error = train.getError();
      train.finishTraining();
      System.out.println("Finished training '" + description + "' network with " + ideal_output.length + "
records");
      System.out.println("System error: " + error);
//      for (MLDataPair pair : trainingSet) {
//
//          final MLData output = network.compute(pair.getInput());
//          System.out.println(pair.getInput().getData(0)
//                  + "," + pair.getInput().getData(1)
//                  + ", actual=" + output.getData(0)
//                  + ",ideal=" + pair.getIdeal().getData(0));
//
//      }
    }

    public double[] predict(double[][] input) {
      while (input[0].length < inputCount) {
        input = append0_5ToArray(input);
//          System.out.println("Making prediction for " + description + ". inputs " + input[0].length + " vs " +
network.getInputCount());
      }
```

```
        System.out.println("Making prediction for " + description + ". Provided inputs " + input[0].length + " vs Expected inputs" + network.getInputCount());

        for (double[] pair : input) {
            BasicNeuralData data = new BasicNeuralData(pair);
            MLData output = network.compute(data);
            return output.getData();
        }
        return null;
    }

    private double[][] append0_5ToArray(double[][] input) {
        double[] data = input[0];
        double[] required = new double[inputCount];
        for (int i = 0; i < inputCount; i++) {
            if (i < data.length) {
                required[i] = data[i];
            } else {
                required[i] = 0.5;
            }
        }
        return new double[][]{required};
    }
```

```
//    public double predict(double input1, double input2) {
//        Double[] predict = predict(new double[][]{
//            {input1, input2}
//        });
//        return predict[0];
//    }
}
```

Appendix VIII. Code for Training the Neural Network

```
public class Trainer {

    private static Trainer instance;

    public static Trainer getInstance() {
        if (instance == null) {
            instance = new Trainer();
        }
        return instance;
    }

    private Map<Long, Map<Integer, Network>> networks = new HashMap<>();

    public void train() {
        Map<Long, Course_Module> courseData = readData();
        for (Long id : courseData.keySet()) {
            Course_Module courseModule = courseData.get(id);
            for (int i = 1; i < courseModule.results.size(); i++) {
                Double[][] inputs = courseModule.results.get(i);
                Double[][] outputs = courseModule.results.get(i + 1);
                if (inputs.length > outputs.length) {
                    inputs = Arrays.copyOfRange(inputs, 0, outputs.length);
                } else if (outputs.length > inputs.length) {
```

```
            outputs = Arrays.copyOfRange(outputs, 0, inputs.length);
        }
        TrainingData td = new TrainingData();
        td.setInput(inputs);
        td.setOutput(outputs);

        final Network network = new Network(inputs[0].length, outputs[0].length, td);
        Course course = ModelManager.getInstance().findOne(Course.class, courseModule.courseId);
        network.setDescription(course.getName() + " Module " + i);
        Map<Integer, Network> moduleNetwork = networks.get(courseModule.courseId);
        if (moduleNetwork == null) {
            moduleNetwork = new HashMap<>();
        }
        moduleNetwork.put(i, network);
        networks.put(courseModule.courseId, moduleNetwork);
        new Thread(new Runnable() {
            @Override
            public void run() {
                network.initialize();
            }
        }).start();
      }
    }
  }
```

```
    private Map<Long, Course_Module> readData() {

        Query query = ModelManager.getInstance().getEntityManager().createQuery("SELECT d.course,
d.module FROM ModuleUnit d GROUP BY d.course, d.module");
        List resultList = query.getResultList();
        Map<Long, Object> courses = new LinkedHashMap<>();
        Map<Long, Course_Module> courseResults = new HashMap<>();
        for (Object item : resultList) {
            Object[] array = (Object[]) item;
            Course c = (Course) array[0];
            Integer module = (Integer) array[1];
            Map<String, Object> p = new LinkedHashMap<>();
            p.put("module", module);
            p.put("course", c);
            List<CourseModuleResult> results = ModelManager.getInstance().findEntities("SELECT p FROM
CourseModuleResult p WHERE p.module = :module AND p.course = :course ", p);
            System.out.println(results.size() + " results found for " + c.getName() + " module " + module);
            p.clear();
            p.put("results", results);
            List<ModuleUnitResult> unitResults = ModelManager.getInstance().findEntities("SELECT p FROM
ModuleUnitResult p WHERE p.courseModuleResult IN :results ORDER BY p.student, p.courseUnit", p);
            System.out.println("\t" + unitResults.size() + " results found for the units ");
            Map<Long, Map<Long, Double>> studentResults = new LinkedHashMap<>();
```

```
for (ModuleUnitResult unitResult : unitResults) {
    Map<Long, Double> studentScores = studentResults.get(unitResult.getStudent().getId());
    if (studentScores == null) {
        studentScores = new LinkedHashMap<>();
    }
    studentScores.put(unitResult.getCourseUnit().getId(), translateScore(unitResult.getResult()));
    studentResults.put(unitResult.getStudent().getId(), studentScores);
}
System.out.println("\t" + studentResults.size() + " students found for the units ");
LinkedHashMap<Long, Integer> unitIndexes = new LinkedHashMap<>();
//format data
List<Double[]> studentData = new LinkedList<>();
for (long studentID : studentResults.keySet()) {
    Map<Long, Double> unitScores = studentResults.get(studentID);
    Collection<Double> values = unitScores.values();
    boolean include = true;
    for (Double d : values) {
        if (d < 0d) {
            include = false;
            break;
        }
    }
    if (include) {
        Double[] scores = values.toArray(new Double[0]);
```

```
            studentData.add(scores);
          }
        }
        //save results
        Course_Module courseModule = courseResults.get(c.getId());
        if (courseModule == null) {
          courseModule = new Course_Module();
          courseModule.courseCode = c.getCode();
          courseModule.courseId = c.getId();
        }
        courseModule.results.put(module, studentData.toArray(new Double[0][]));
        courseResults.put(c.getId(), courseModule);
        System.out.println("\t" + studentData.size() + " students with valid data found for the units ");

      }
//      ModelManager.getInstance().close();
      return courseResults;
    }

    public double[] predict(long course, int module, double[][] data) {
      Map<Integer, Network> courseNetworks = networks.get(course);
      Network moduleNetwork = courseNetworks.get(module);
      double[] prediction = moduleNetwork.predict(data);
      return prediction;
```

```
}

public boolean hasNextModule(long course, int current) {
   Map<Integer, Network> courseNetworks = networks.get(course);
   Network moduleNetwork = courseNetworks.get(current + 1);
   return moduleNetwork != null;
}

private Double translateScore(String result) {
   if (result == null || !result.matches("(?i)[a-z]+[0-9]+")) {
      //cannot be used to determine outcome
      return -1d;
   }
   Pattern p = Pattern.compile("[0-9]+");
   Matcher matcher = p.matcher(result);
   matcher.find();
   String group = matcher.group();
   Integer i = Integer.parseInt(group);
   return 1d - ((double) i / 8d);
}

public static void main(String[] args) {
   Trainer t = new Trainer();
   t.train();
```

```
    }

    private class Course_Module {

        private long courseId;
        private int courseCode;
        private Map<Integer, Double[][]> results;

        public Course_Module() {
            this.results = new HashMap<>();
        }
    }
}
```