**We need to see your research title**

**CHAPTER ONE**

**1.0 INTRODUCTION**

**1.1 Background of the Study**

In the landscape of higher education, the retention and academic success of students have always been of paramount concern. Student attrition, often referred to as dropout or failure to complete academic programs poses significant challenges to both educational institutions and students themselves. As such, the ability to predict student performance and outcomes is crucial for the improvement of retention and graduation rates and the overall well-being of students. Appropriate career selection is one of the key decisions in students’ lives. A wrongly chosen career sometimes destroy students’ life. The right career selection for a student is always based on the selection of the right educational program. Usually, students of higher education institutions, management of these institutions and parents follow general trends to select or recommend an educational program at the Bachelor and Master Level of studies. This kind of choice is based on the educated guess of the stakeholders and usually does not consider the students' personal, academic and socioeconomic characteristics. Selection of an appropriate educational program needs a prediction system to foresee students’ academic performance prior to taking admission or continuing to classes. However, the development of this intelligent prediction system is a challenging task due to its dependency on various factors, such as personal, socioeconomic, psychological and other factors. Data mining and Machine learning are the key tools which can better perform in this area and can help to predict students’ performance prior to taking admission to Master level programs. With the emergence of Big Data and predictive analytics, there has been a notable surge in research focused on predicting student dropout and academic performance using machine learning techniques.

Student attrition in higher education institutions is a multifaceted issue that has persisted for decades. It involves students disengaging from their academic programs before they can attain their degrees, diplomas, or certificates. This phenomenon is not only a concern for educational institutions but also for funding bodies, government agencies, and the broader society. The financial repercussions of student attrition are considerable, as institutions lose tuition fees and often have to invest more in recruiting new students to maintain enrollment levels (Tinto, 1975). Moreover, dropout rates can have a detrimental effect on a university's reputation and its ability to fulfil its educational mission. On a personal level, academic failure can have a lasting impact on students' well-being and mental health. Those who do not complete their education may face challenges in securing employment and pursuing their career aspirations, leading to personal and societal consequences (Pascarella & Terenzini, 2005).

The advent of Big Data and predictive analytics in recent years has provided a new paradigm for addressing student attrition and academic performance prediction. This technology enables the collection, storage, and analysis of vast amounts of data related to student behaviour, academic records, and personal attributes (Davenport & Harris, 2007). Machine learning, a subset of predictive analytics, has garnered significant attention for its potential to create predictive models and discern patterns that can forecast student behaviours and outcomes (Kotsiantis, Zaharakis, & Pintelas, 2006).

Predictive models in higher education, often referred to as "early warning systems," have gained prominence. These systems analyze data collected during the early stages of a student's academic journey, such as the first semester, to identify students who may be at risk of attrition. Early intervention based on these predictions can lead to improved student outcomes (Arnold, 2010). Predictive analytics can also be harnessed to enhance evidence-based teaching and learning strategies. By predicting academic performance, educators can tailor their teaching methods and support mechanisms to address individual students' needs, ultimately improving academic outcomes (Siemens & Long, 2011).

In considering Machine Learning Technique, Various machine learning techniques, including decision trees, k-nearest neighbours (KNN), and logistic regression, have proven effective in predicting academic performance and student attrition. These models consider historical data, personal attributes, and course-related information to make informed predictions about students' future success or failure (Romero & Ventura, 2010).

In response to the challenges posed by student attrition and academic failure, many educational institutions, including Baze University, have implemented intervention programs. These programs aim to provide targeted support to students at risk of failing to graduate. Identifying these at-risk students early is crucial for delivering effective support and increasing their chances of academic success (Bean & Metzner, 1985).

Additionally, accurately predicting student performance in advance can help educational institutions enhance the quality of education and academic results. By reducing failure rates, universities can bolster student success and contribute to a more prosperous and educated society (Buckley & Seppälä, 2011).

In conclusion, the application of predictive models and machine learning techniques has emerged as a transformative approach to address the challenges of student attrition and academic performance prediction in higher education. This research at Baze University seeks to leverage the potential of predictive analytics to identify patterns indicative of students' academic performance, ultimately contributing to more effective teaching, learning, and support systems within the institution.

**1.2 Statement of the Study**

In higher education, the challenges of student attrition and the need to predict and improve academic performance are pressing issues that demand attention. At Baze University and many other institutions globally, the struggle to retain students and enhance their academic outcomes is not just a financial concern but also a matter of student well-being and educational effectiveness. The emergence of Big Data and predictive analytics has opened the door to predictive modelling, offering the potential to address these challenges with data-driven precision.

Student attrition and academic failure pose significant challenges to academic institutions, funding bodies, and students. Predicting student performance and identifying at-risk students can help improve retention and graduation rates. However, existing predictive models are based on similar sets of macro-level data (e.g., institutional data, academic performance) or micro-level data (e.g., click-stream data), and Meso-level data on students' daily experience and engagement with both other students and the university itself are largely absent from this research. Data is the lifeblood of predictive analytics, yet many educational institutions, including Baze University, still face challenges in data collection, integration, and quality (Romero & Ventura, 2010). Incomplete, inconsistent, or inaccessible data can hinder the development and deployment of effective predictive models

The use of predictive analytics and machine learning in higher education has shown promise in predicting student dropout and academic performance. However, these models often rely on cross-sectional predictions and may not capture the dynamic nature of student performance and engagement.

Baze University, like many other academic institutions, faces challenges in predicting and improving students' academic performance. Developing predictive models that can identify patterns indicating students' academic performance can help the university in implementing targeted interventions and support systems to improve student outcomes. While predictive models are being increasingly explored in higher education (Arnold, 2010), they often lack the specific institutional context and domain knowledge needed to accurately forecast academic performance. Generic models may not fully capture the unique factors contributing to student success at Baze University.

Predictive models can identify students at risk of attrition or academic failure, the translation of these predictions into effective interventions remains a significant problem. Identifying at-risk students is only the first step; subsequent support mechanisms and interventions must be in place to address their unique needs (Bean & Metzner, 1985). The use of predictive models in education raises ethical questions, particularly regarding data privacy and fairness (Siemens & Long, 2011). Ensuring that the use of predictive analytics aligns with ethical standards and safeguards the rights and interests of students is a complex issue to navigate. Assessing the actual impact of predictive models on academic outcomes and student retention is essential. Understanding whether these models translate into tangible improvements in student success is a critical problem (Davenport & Harris, 2007).

In light of these problems, the research at Baze University seeks to address these challenges and develop context-specific predictive models that can identify patterns indicative of students' academic performance accurately, thereby contributing to the institution's efforts to enhance student success and retention.

**1.3 Aim and Objectives of this study**

The aim of the study "Predictive Models Analysis to Identify Patterns that indicate Students' Academic Performance at Baze University" is to develop predictive models using machine learning techniques to identify patterns indicating students' academic performance at Baze University. The specific objectives of this study are:

1. Develop predictive models using machine learning techniques to identify patterns indicating students' academic performance at Baze University.
2. Evaluate the performance of the predictive models using appropriate validation approaches.
3. Identify the key factors and variables that contribute to students' academic performance.
4. Provide recommendations for the implementation of targeted interventions and support systems based on the findings of the predictive models.

**1. 4 Research Questions to the Study**

The research questions that will guide this study are:

1. What are the key factors and variables that contribute to students' academic performance at Baze University?
2. How can predictive models be developed using machine learning techniques to identify patterns indicating students' academic performance?
3. What is the performance of the predictive models in terms of accuracy, precision, recall, and F1-score?
4. What are the implications of the findings for the implementation of targeted interventions and support systems at Baze University?

**1.5 Significance of the Study**

Predictive Models Analysis in the context of student academic performance at Baze University holds immense significance for various stakeholders, including the academic institution, students, and the broader educational community. This research, situated within the scope of higher education, predictive analytics, and academic success, aims to address several critical issues and offers substantial value.

One of these critical issues is Improved Retention and Graduation Rates. Enhancing student retention and graduation rates is a primary concern for Baze University, as it is for many institutions worldwide. By developing predictive models tailored to the university's unique context, this study provides the potential for early identification of students at risk of attrition or academic failure. Early intervention strategies can then be implemented to provide support and guidance to these students, ultimately increasing the likelihood of their successful progression through their academic programs. Improved retention and graduation rates signify a more successful and thriving institution (Pascarella & Terenzini, 2005).

Secondly, Enhanced Student Well-being. This study's outcomes have direct implications for the well-being of students. By identifying those at risk of academic failure early, universities can provide targeted support to help students overcome academic challenges. Preventing academic failure not only contributes to students' academic well-being but also has positive effects on their mental and emotional health, reducing the stress and anxiety associated with the risk of dropout (Tinto, 1975).

Thirdly, Evidence-Based Teaching and Learning. The development of predictive models allows for evidence-based decision-making in teaching and learning strategies. By using data-driven insights, educators can adjust their teaching methods and curricular approaches to cater to the diverse needs and capabilities of students (Siemens & Long, 2011). This not only improves the academic outcomes but also creates a more engaging and effective learning environment.

Meanwhile, Tailored Student Support Mechanisms is also another critical issue. The predictive models developed in this study pave the way for the creation of targeted student support systems. Instead of employing a one-size-fits-all approach, universities can customize support mechanisms, such as academic advising, tutoring, or counseling, to address the unique challenges faced by individual students. This, in turn, enhances the overall learning experience for students (Bean & Metzner, 1985).

Lastly, Ethical Data Use. The ethical considerations inherent in the use of predictive analytics are of paramount importance. This study offers an opportunity to explore and address the ethical aspects of data privacy and fairness in education. By navigating these ethical dilemmas, the research contributes to the responsible and ethical use of data in higher education (Davenport & Harris, 2007).

**1.6 Scope of the Study**

The scope of the study is to develop predictive models using machine learning techniques to identify patterns indicating students' academic performance at Baze University. The study will focus on identifying patterns indicating students' academic performance at the University using machine learning techniques to develop predictive models and evaluating the performance of the predictive models using appropriate validation approaches. The study will also identify the key factors and variables that contribute to student's academic performance and provide recommendations for the implementation of targeted interventions and support systems based on the findings of the predictive models. The study will be limited to Baze University and will not be generalizable to other academic institutions.

**REFERENCES**

Arnold, K. E. (2010). Signals: Applying Academic Analytics. Educause Quarterly, 33(1), 1-3.

Bean, J. P., & Metzner, B. S. (1985). A Conceptual Model of Nontraditional Undergraduate Student Attrition. Review of Educational Research, 55(4), 485-540.

Davenport, T. H., & Harris, J. (2007). Competing on Analytics: The New Science of Winning. Harvard Business Review Press.

Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2006). Machine Learning: A Review of Classification and Combining Techniques. Artificial Intelligence Review, 26(3), 159-190.

Pascarella, E. T., & Terenzini, P. T. (2005). How College Affects Students: Volume 2: A Third Decade of Research. Jossey-Bass.

Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(6), 601-618.

Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. Educause Review, 46(5), 30-32.

Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. Review of Educational Research, 45(1), 89-125.

Buckley, P., & Seppälä, T. (2011). The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment. Harvard Business Press.