PREDICTING THE SUCCESS OF TVET LEARNERS IN A HIGHER EDUCATION ENGINEERING PROGRAMME

Rodney Alan Stops

Submitted in fulfilment of the requirements for the degree of Master of Engineering in the Department of Chemical Engineering, Faculty of Engineering, and the Built Environment at Durban University of Technology

FINAL SUBMISSION

Professor Darren Lortan
Supervisor
30th October 2021

Professor Suresh Ramsuroop
Co-Supervisor
30th October 2021

DURBAN, SOUTH AFRICA

October 2021

Abstract

Where there is a will, there is a way!

William Hazlitt, 1822

South Africa has experienced radical political and social change since 1994. Recognised for their role in the transformation process, universities have been and remain at the forefront of this change. While learners recognise that higher education is paramount to changing their socio-economic condition, the massification of education along with new and advanced curricula has presented ongoing challenges for both universities and learners. Coming from diverse backgrounds and dealing with a variety of academic choices, learners encounter many challenges to entering Higher Education (HE). University programmes accept learners with varying competency levels and needs but are required to ensure that graduates meet standards that are acceptable to both industry and the academia. In the quest for universities to improve the delivery of educationally sound and industrially relevant programmes, ongoing research is being conducted and new and innovative ways have had to be developed to solve the problems associated with larger numbers of underprepared learners. An emerging method being employed in HE is the use of Data Analytics and Education Data Mining (EDM) techniques to derive solutions to assist institutions in maximising retention, and through-put rates.

Durban University of Technology (DUT) has, since 1994, accepted learners into the Report 151 National Diploma from Technical Vocational Education and Training (TVET) Colleges. These learners, in respect to the Articulation Policy for the Post-School Education and Training (PSET) system of South Africa in terms of Section 8(2)(b) of the NQF Act, 2008 (Act 67 of 2008), are among those referred to as articulating learners. The perception among DUT staff involved with the teaching of these learners, is that they are as able to cope with the complexity and quality of engineering programmes as those learners entering the institution directly after completing their school leaving Senior Certificate/National Senior Certificate (SC/NSC).

As no previous formal tracking, analysis or research has been conducted to determine the success or failure of learners articulating from TVET Colleges into DUT in general or into DUT's

Department of Electrical Power Engineering specifically, this research utilised Educational Data Mining and Inferential Statistics on an engineering learner dataset, to determine hidden patterns and relationships. Specifically, those relationships that promote progression, correlation, and selection were investigated. The Waikato Environment for Knowledge Analysis (WEKA) was employed to do the EDM, and a tool was developed to assist with the selection of learners entering the department, as well as ranking those entrance requirements that affect course content and throughput and various statistical methods were employed to conduct a retrospective longitudinal study. The Knowledge Discovery in Databases Process is used to work with 5 years of learner data.

Both from the perspective of the progression of learners from semester 1 to semester 2 and the cohort throughput analysis, the results showed that there were no statistically significant differences in the performance of learners articulating from TVET colleges into the National Diploma: Electrical Engineering at DUT and their counterparts from high school admitted directly in the said qualification. The findings are that learners from a TVET College articulating into an HEI qualification, specifically the National Diploma: Electrical Engineering at DUT, complete the course in similar rates and in similar proportions to those learners admitted directly from school. The statistical analysis indicates that 77.6% of TVET N4 learners are promoted to semester 2, compared to 70.0% of SC/NSC learners and the EDM prediction tool developed for TVET N4 learners, the IBK classification tool resulted in a 77.61% accuracy, while the ANN classification tool returned an accuracy of 77.56% for the SC/NSC learners.

Declaration

I, Rodney Alan Stops, declare that this thesis is my own unaided work, except where due acknowledgements is made to others. This thesis is being submitted for the Master of Engineering Degree to The Department of Chemical Engineering at The Durban University of Technology and has not been submitted previously for any other degree or examination.

Rodney Alan Stops

Date: 28th October 2021

Dedication

Firstly and foremost, I dedicate this Master's Qualification to my wife, Amanda, who has been my rock and sounding board during the entire process. She has had to listen to explanations ranging from education to neural networks and has always listened with an attentive ear and given honest, fair, logical responses. She has been there when the powers that be, have thrown everything they could at us, but we have stayed strong. Without her by my side I doubt I would have been able to finish. Thanks for picking up the slack.

Together with Amanda, I dedicate this thesis to my small family, Keagan and Madison, the love of my life. Keagan, your fighting spirit to overcome CP, amazes me every single day. Madison, your loving kindness keeps us all better people. I do it all for you!

Finally, I dedicate it to my brother, Dudley, who was diagnosed with stage 4 cancer, while I was finishing this thesis. It reminds me that life is short, and you need to live it to the fullest. Let us get back to living life!

I LOVE YOU ALL!

Acknowledgement

I sincerely thank the following people for their contribution.

Professor Darren Lortan and Professor Suresh Ramsuroop for agreeing to be my supervisors. I cannot thank you enough. I always felt that I needed supervisors that I admired and a topic that I was passionate about. I managed to find that in this Master's Thesis and in my supervisors. Thank You.

Professor David Day, although neural networks remain a mystery to you, I thank you for the many, many hours of discussions we had while walking, refreshing the mind and engaging on that topic and others. Your insights always make interesting points to consider. I thank you for your friendship and support during these tough times. Thanks Bud!

Dr Gillian Reid, I cannot thank you enough for the enormous contribution you gave to this thesis and the on-going support to keep me on track. A late' is always on the house. Thank You!!

Contents

LIST OF ACRONYMS IV

LISTS OF TABLES VII

LIST OF FIGURES IX

CHAPTER 1 - INTRODUCTION AND BACKGROUND 1

1.1 INTRODUCTION 1

1.2 INTRODUCTION TO THE DEPARTMENT 2

1.3 BACKGROUND INFORMATION 5

1.4 PURPOSE OF THE STUDY 7

1.5 HYPOTHESIS AND AIM OF STUDY 8

1.6 STRUCTURE OF THE DISSERTATION 10

CHAPTER 2 - LITERATURE REVIEW 11

2.1 INTRODUCTION 11

2.2 HISTORY OF HIGHER EDUCATION IN SOUTH AFRICA 11

2.2.1 SOUTH AFRICA, A HISTORICAL OVERVIEW PRE-1994 11

2.2.2 SOUTH AFRICA, A HISTORICAL OVERVIEW POST-1994 12

2.3 ARTICULATION 16

2.3.1 DUT BASELINE ARTICULATION STUDY 2017 19

2.3.2 THE THIRD NQF CONFERENCE 22

2.6 CONCLUSION 30

CHAPTER 3 - METHODOLOGY AND METHODS 32

3.1 INTRODUCTION 32

3.2 ETHICAL APPROVAL 32

3.3 METHODOLOGY 32

3.4 BIG DATA 33
3.4.1 STATISTICS 34
3.4.2 EDUCATIONAL DATA MINING 36
3.5 DATA MINING TECHNIQUES 39
3.5.1 LINEAR REGRESSION 40
3.5.2 LOGISTIC REGRESSION 41
3.5.3 NAÏVE BAYES 43
3.5.4 DECISION TREES 44
3.5.5 SUPPORT VECTOR MACHINE 47
3.5.6 ARTIFICIAL NEURAL NETWORKS 49
3.6 PREVIOUS EDUCATIONAL DATA MINING RESEARCH 61
3.7 KNOWLEDGE DISCOVERY PROCESS (KDD) 62
3.7.1 PRE-KNOWLEDGE 62
3.7.2 SELECTION 63
3.7.3 DATA CLEANING AND PROCESSING 64
3.7.4 DATA REDUCTION 65
3.7.5 DATA MINING TYPE 66
3.7.6 DATA MINING ALGORITHMS 66
3.7.7 ACTUAL DATA MINING 67
3.7.8 INTERPRETATION 67
3.7.9 CONSOLIDATION 68
3.8 RULES FOR PROGRESSION 68
3.9 CONCLUSION 70

CHAPTER 4 - RESULTS 71

4.1 INTRODUCTION 71
4.2 STATISTICAL ANALYSIS 71
4.2.1 THE SAMPLE 71
4.2.2 DATA ANALYSIS 71
4.2.3 COMPARISONS OF SCORES BETWEEN THE GROUPS 73
4.2.4 SCHOOL SUBJECT COMPARISON 74
4.2.5 UNIVERSITY SUBJECT COMPARISON 76

4.2.6 CROSS TABULATIONS 78
4.2.7 CORRELATIONS 79
4.2.8 BINARY LOGISTIC REGRESSION MODEL 83
4.3 EDUCATIONAL DATA MINING 86
4.3.1 SC/NSC SUBJECT INPUT EVALUATOR 86
4.3.2 TVET SUBJECT INPUT EVALUATOR 88
4.3.3 EVALUATOR RANKING RESULTS COMPARISON 89
4.4 TVET N4 EDUCATIONAL DATA MINING 90
4.4.1 SUMMARY TVET N4 SUCCESS PREDICTOR 91
4.4.2 TVET N4 IBK ALTERNATE CLASSIFIER 93
4.4.3 TVET N4 IBK ALTERNATE CLASSIFIER SUMMARY 93
4.5 SC/NSC EDUCATIONAL DATA MINING 95
4.5.1 SUMMARY SC/NSC SUCCESS PREDICTOR 95
4.6 SC/NSC MATHEMATICS VS TVET N4 MATHEMATICS SYLLABI REVIEW 97
4.7 SC/NSC PHYSICAL SCIENCE VS TVET N4 ENGINEERING SCIENCE SYLLABI REVIEW 99
4.8 TRACKED COHORT OF FIRST-TIME REGISTERING LEARNERS 101

CHAPTER 5 - CONCLUSION 104

5.1 INTRODUCTION 104
5.2 SUMMARY 104
5.3 FURTHER RESEARCH 107
5.4 CONCLUDING STATEMENT 108
ANNEXURE 1 – TVET N4 RESULTS 2013 –2017 117
ANNEXURE 2 – SC/ NSC RESULTS 2013 - 2017 119
ANNEXURE 3 – WEKA TOOL [105] 128
ANNEXURE 4 – WEKA INPUT EVALUATOR SC/NSC 129
ANNEXURE 5 – WEKA INPUT EVALUATOR TVET N4 130
ANNEXURE 6 – WEKA TVET N4 CLASSIFICATION MODELS 131
ANNEXURE 7 – WEKA IBK CLASSIFIER 132
ANNEXURE 8 – WEKA SC/NSC CLASSIFIER [104] 133

List of Acronyms

ANC	-	African National Congress
ANN	-	Artificial Neural Networks
BEng Tech	-	Bachelor of Engineering Technology Degree
BEng Tech Hon	-	Bachelor of Engineering Honours Degree
BTech	-	Bachelor of Technology Degree
CAT	-	Credit Accumulation and Transfer
CAO	-	Central Application Office
CHE	-	Centre for Higher Education
CNE	-	Christian National Education
DEPE	-	Department of Electrical Power Engineering
DOE	-	Department of Education
DT	-	Decision Tree
DHET	-	Department of Higher Education and Training
DUT	-	Durban University of Technology
ECSA	-	Engineering Council of South Africa
EDM	-	Educational Data Mining
FET	-	Further Education and Training
GETCA	-	General Education and Training Certificate for Adults
GFETQSF	-	General and Further Education and Training Sub-Framework
HE	-	Higher Education

List of Acronyms (Cont.)

HEI	-	Higher Education Institution
HEMIS	-	Higher Education Management Information System
HEQSF	-	Higher Education Qualifications Sub-Framework
IBk	-	Instance Base Learner (nearest K)
KDD	-	Knowledge Discovery in Databases
LG	-	Logistic Regression
LMT	-	Logistic Model Trees
LR	-	Linear Regression
MIS	-	Management Information System
ML	-	Machine Learning
MLS	-	ML Sultan Technikon
MOU	-	Memorandum of Understanding
NB	-	Naïve Bayes
NCV	-	National Certificate Vocational
NDip	-	National Diploma
NLRD	-	National Learners Record Database
NSC	-	National Senior Certificate
NQF	-	National Qualifications Framework
OQSF	-	Occupational Qualifications Sub-Framework

List of Acronyms (Cont.)

OHSAct	-	Occupational Health and Safety Act
PQM	-	Programme Qualification Mix
PSET	-	Post-School Education and Training
QC	-	Qualifications Council
RF	-	Random Forest
RPL	-	Recognition for Prior Learning
SAQA	-	South African Qualifications Authority
SC	-	Senior Certificate
STEM	-	Science. Technology, Engineering and Mathematics
SVM	-	Support Vector Machines
TIMSS	-	Trends in International Maths and Science Study
TN	-	Technikon Natal
TVET	-	Technical Vocational Education and Training
UoT	-	University of Technology
WEKA	-	Waikato Environment for Knowledge Analysis

Lists of Tables

Table 1.1: First Year Entering Learners 7
Table 3.1: Previous Educational Data Mining Research 61
Table 3.2: Symbol / Numeric Value Conversion [8] 65
Table 3.3: Semester 1 Subjects / Subject Codes and FTE Credit Value 69
Table 4.1: English A and English B Results Comparison 72
Table 4.2: Results Table 73
Table 4.3: Significance Test 75
Table 4.4: Significance Test University Subjects 77
Table 4.5: Chi-Squared Test 78
Table 4.6: TVET N4 Bivariate Correlation 79
Table 4.7: SC/NSC Bivariate Correlation 81
Table 4.8: Binary Logistic Regression Model Output Summary 83
Table 4.9: Classification Table 83
Table 4.10: Variable in the Equation 84
Table 4.11: Summary of SC/NSC Attribute Evaluation 87
Table 4.12: Summary of TVET N4 Attribute Evaluation 88
Table 4.13: TVET N4 and SC/NSC Input Results Comparison 89
Table 4.14: TVET N4 Data Mining Accuracy Table 91
Table 4.15: TVET N4 Data Mining Error Table 91
Table 4.16: TVET N4 Data Mining Statistics Table 91
Table 4.17: TVET N4 Data Mining Accuracy Table (Alternate) 93
Table 4.18: TVET N4 Data Mining Error Table (Alternate) 93
Table 4.19: TVET N4 Data Mining Statistics Table (Alternate) 93
Table 4.20: SC/NSC Data Mining Accuracy Table 95
Table 4.21: SC/NSC Data Mining Error Table 95
Table 4.22: SC/NSC Data Mining Statistics Table 96
Table 4.23: TVET N4 vs SC/NSC Mathematics Syllabi 97
Table 4.24: TVET N4 Mathematics vs Semester 1 Mathematics Success Rate 98
Table 4.25: TVET N4 Engineering Science vs SC/NSC Physical Science Syllabi 100
Table 4.26: Cohort Study of 3 Year Programmes First-time Entering Students for 2013 101

Table 4.27: Cohort Study of 3 Year Programmes First-time Entering Students for 2014 102
Table 4.28: Cohort Study of 3 Year Programmes First-time Entering Students for 2015 102
Table 4.29: Cohort Study of 3 Year Programmes First-time Entering Students for 2016 103
Table 4.30: Cohort Study of 3 Year Programmes First-time Entering Students for 2017 103

List of Figures

Figure 2.1: Professional Designated Engineers on the NLRD [20] 24
Figure 2.2: Age Profile of Engineering Professionals [20] 25
Figure 2.3: NQF Paths - Possible vs Actual [20] 26
Figure 2.4: NQF Possible Pathways for Technologists [20] 27
Figure 2.5: Engineering Learning Pathways [20] 28
Figure 2.6: Engineering Qualification Graphical View [53] 29
Figure 3.1: Normal Distribution Plot [57] 35
Figure 3.2: Knowledge Discovery in Databases Process (KDD) [63] 38
Figure 3.3: Classification Accuracy by Technique [65] 39
Figure 3.4: Sigmoid Function Graph [67] 42
Figure 3.5: Decision Tree Illustration [69] 44
Figure 3.6: Entropy Outline [69] 45
Figure 3.7: SVM Hyperplane [71] 47
Figure 3.8: SVM Analysis [71] 48
Figure 3.9: The Brain Cell [76] 50
Figure 3.10: Mathematical Representation of a Neuron [77] 51
Figure 3.11: Three Layer Network [75] 52
Figure 3.12: Artificial Neuron Representation [77] 53
Figure 3.13: Nonlinear Model of a Neuron [79] 54
Figure 3.14: The Sigmoid Function [77] 56
Figure 3.15: Two-Layer Perceptron [77] 58
Figure 4.1: Entrance Subjects for TVET N4 & SC/NSC (Matric) 74
Figure 4.2: University Subject Comparison 76

Chapter 1 - Introduction and Background

1.1 Introduction

This chapter introduces the background to the reason for conducting this study as well as the background to the Department of Electrical Power Engineering (DEPE) at Durban University of Technology (DUT), in which the study takes place. It also contextualises the purpose for the study and the reasons for extrapolating the results to other departments within DUT's Faculty of Engineering and the Built Environment.

South Africa is currently experiencing an extremely high unemployment rate of 32.6% [1]. In the age category 15-34, the unemployment rate is a staggering 46.3%. In contrast there is a shortage of qualified professionals in the engineering and science fields. A common problem, found presently across the world, is that of a dissipating interest in Science, Technology, Engineering and Mathematics (STEM) subjects at secondary school level. This has translated into a decline in the number of learners[1] exiting secondary education who meet the STEM requirements to enter engineering fields. This could be attributed to the low standard of mathematics and science education in schools [2]. Higher Education institutions are subsequently not able to meet the required demand for these graduates whose role in maintaining and growing the South African Economy is vital. A study in 2010 [3] indicated that there were fewer than 2500 black South African learners satisfying all entrance requirements for admission to any undergraduate programme in South Africa. The Sasol Inzalo Foundation reports that fewer than 10% of learners who enroll for primary education will complete their secondary education, and less than 20% of those choosing to study a STEM discipline. Of these learners only 50% are predicted to graduate [4]. In order to respond to the paucity of STEM graduates and the mismatch between graduate attributes and industry expectations, universities must not only address certain shortcomings on the higher education level, but also adopt and nurture schools to address the inadequate performance of learners in the STEM subjects [5]. University initiatives might include additional classes to school learners, practical projects or minimally assisting learners that have the interest and required subjects but lack

[1] The term learner is used exclusively instead of student, unless specifically used in a quote or rule.

the necessary outcomes [2]. One initiative to improve this situation is the Joint Initiative on Priority Skills Acquisition (JIPSA) Secretariat [6] which is currently engaged with the Department of Education to increase the pool of SC/NSC learners with quality Mathematics and Physical Science passes. They accomplish this through upgrading the skills of existing Mathematics and Science Teachers.

1.2 Introduction to the Department

The Department of Electrical Power Engineering (DEPE) is, at time of the writing, one of ten departments in the Faculty of Engineering and the Built Environment at DUT. As with other Departments at DUT, it developed from the merger between Technikon Natal (TN) and ML Sultan Technikon (MLS) in 2001. It currently offers both the Bachelor of Engineering Technology Degree (BEng Tech) and Bachelor of Engineering Technology Honours Degree (BEng Tech Hon) in Electrical Power Engineering and has phased out the National Diploma (NDip) and is phasing out the BTech Degree (BTech) in Electrical Engineering.

The DEPE has an annual intake of 120 learners. Learners accepted into the programmes must, at minimum, have a school leaving certificate - the Senior Certificate (SC) obtained prior to 2008, or the National Senior Certificate (NSC) - introduced post-2008 and quality controlled by Umalusi [7]. In addition, learners are required to have met both minimum university and departmental entrance requirements [8, 9].

The limit on the annual intake of 120 first-time entering learners is influenced by industrial demand [10]. These numbers also reflect DUT's compliance with the requirement in the National Plan for Higher Education, Department of Education (DOE, 2001) for Higher Education Institutions (HEI) [11]. The aim of these enrolment figures is to increase the participation rate of learners in science, engineering, and technology. The DEPE traditionally admitted learners directly from school, who were in possession of a Senior Certificate or National Senior Certificate with exemption and the minimum institutional requirements [8, 9]. In addition, since the mid-nineties, the DEPE has accepted learners articulating from the Technical Vocational Education and Training (TVET) or Further Education and Training

Colleges (FET) as they were previously known, into the programme. Articulation, briefly, is a little-known policy at present. The Articulation Policy for the Post-School Education and Training (PSET) enables learners to achieve their aim of entering a higher education programme of their choice. It also provides HE departments an alternative means to meet enrolment targets [12] and sustains the pipeline supplying industry with sufficient graduates. This academic rehabilitation or articulation stream developed historically from the Department's close ties to the TVET and previous FET colleges. The close ties were entrenched in the evolution of the Technikons within the education environment [12, 13] and the nature of the staff complement at the time; Technikons [14] having grown organically from Technical Colleges with most lecturers previously being artisans or technicians, and employed as higher education lecturers for their technical and practical experience in industry. Technikons, in more detail, were *technological* in character having been created for the purpose of teaching the "application of knowledge" [11]. The primary function of Technikon programmes was therefore occupational and centered around training individuals for careers [11, 13]. As a result of their long-held function of providing industry with a vocationally educated and occupationally skilled workforce, academics working in Technikons had developed partnerships with industry and maintained ties with the Technical Vocational Education and Training (TVET) or Further Education and Training Colleges (FET).

The DEPE had previously also offered the subjects required for the National Government Ticket; a qualification closely linked to the National N Diploma. The National N Diploma is awarded to learners on completion of their N4, N5 and N6 certificates at a TVET/FET College, followed by 18 months of relevant practical application in places of work [15]. Learners with a N6 and relevant managerial experience may choose to apply to write the "Government Ticket". Learners from a Technikon/University of Technology (UOT) or traditional university may all apply to write the Government Ticket Exam [16]. The Department of Labour inspects all subjects that learners have completed before admitting candidates to write the two exams [17]. The Government Ticket compromises two examinations, namely the Occupational Health and Safety Act (OHSAct) and a theoretical exam (Plant Engineering Exam) that covers all previous engineering studies. The candidate becomes a Certified Engineer on successful completion and registration with the Engineering Council of South Africa (ECSA) [17].

In order to broaden access to the Department of Electrical Power Engineering programmes as required by DUT, an alternative access route was established for learners who had not met the required SC or NCS admission requirements or who had not satisfied the minimum entrance requirements for Mathematics and / or Physical Science in the Senior Certificate or National Senior Certificate. As stated previously, this approach was also adopted for the DEPE to remain compliant with the National Plan for Higher Education (DOE, 2001) [6]. In terms of this alternative access pathway, learners who had registered for N4 Mathematics, N4 Engineering Science, N4 Electrotechnics and N4 Industrial Electronics at a TVET/FET college and achieved a pass mark of 50% in each of these courses were admitted to the National Diploma: Electrical Engineering. This practice was adopted by several other departments in the Faculty of Engineering at the legacy Technikon Natal (TN), now DUT.

Historically, N4 Mathematics, N4 Engineering Science, N4 Electrotechnics and N4 Industrial Electronics were offered by the TVET/FET colleges on a trimester-basis. This enabled learners who had failed to meet departmental entrance requirements at the first annual registration and who went to a TVET/FET college to complete the subjects in time for the second semester registration at TN / DUT - usually in July of each year. This was possible because the Diploma was offered on a semester basis, with new learners starting every 6 months. The practice of accepting TVET/FET learners continued as most lecturing staff from all the departments within the faculty held the view at that time, that these learners had supplemented their SC/NCS results and in the sense contemplated in the National CAT Policy (Clause 15a and d). In addition, their performance was deemed to be equivalent to those who obtained their SC/NSC on the first attempt. [18] This pathway or articulation mechanism allowed previously unsuccessful candidates a chance to enter the system quicker – in that they did not have to spend a year repeating failed subjects or the entire SC/NSC subject compliment.

Unfortunately, no active tracking of these cohorts of articulating learners were conducted over the years. Consequently, there exists no data on the performance of these learners. In particular their performance could not be pitted against their counterparts admitted into the National Diploma directly from High School. The sustainability of this alternative pathway remained contingent upon positive perceptions, that at time of writing, was showing signs of waning interest, fueled by a rise in negative anecdotal perceptions about the quality of such learners.

This research, traversing the fields of Engineering, Mathematics and Education, intends to determine whether statistically significant differences exist between the performance of cohorts of these articulating learners and their SC/NSC counterparts. The study intends to quantify the utility of articulating learners and develop a predictive tool to assist with learner selection for this articulating pathway. This will be achieved by providing evidence of how articulated TVET learner cohorts have performed in previous Electrical Engineering qualifications by comparing them to other entrance systems. In addition, this study plans to identify those inputs that might contribute to the success and retention of Electrical Engineering learners in the currently offered BEng Tech programme.

1.3 Background Information

The Department of Electrical Power Engineering at DUT, as per the Programme Qualification Mix (PQM) submitted to the Department of Higher Education and Training (DHET), only currently registers learners for the BEng Tech Degree in Electrical Power Engineering. The Department has phased out the National Diploma and introducing the BEng Tech Degree. The first cohort of successful applicants for the degree were registered in 2018.

The current 3-year degree programme in Electrical Power Engineering, as mentioned previously, is a development from the previously offered three-year National Diploma: Electrical Engineering and 1-year BTech Electrical Power Engineering that is, offered part-time over 2-years. The entrance requirements for the diploma were a minimum of a SC/NSC with a pass (4) for Mathematics, Physical Science and English. An additional entrance requirement for the degree is that the combined marks for Mathematics and Physical Science must equal or exceed 120 points. The change from the diploma to the degree, as supported and promoted by industry, necessitated this adjustment to reflect the degree levels and status. This change has exacerbated the department's difficulty in recruiting learners with the prerequisite subjects and marks. A trend generally being experienced in the STEM sector [19-21].

The department presently has the human resources, facilities, and equipment to accept 120 new learners into the programme each year, but this target is not always achieved. This is due in part to large numbers of applicants being in possession of a SC/NSC with exemption but falling short in key subjects, - English, Mathematics or Physical Science. This has the potential to prevent the department from meeting the number of graduates required by industry as well as failing to meet the funding formulae requirements [22], introduced in 2004. These requirements regulate enrolment, increase retention, throughput, and success rates of learners.

Annual applications for the Department of Electrical Power Engineering's degree and previous diploma are generally in excess of 2500, with only 120 available places. Despite the number of annual applicants, most learners, unfortunately, do not meet the minimum requirements. The number of first-time entry learners who have met the minimum requirements from 2018 to 2021 are shown in Table 1.1. The shortcoming is addressed by considering applicants from TVET Colleges intending to articulate into university. This invariably means that articulating learners join the programmes only after the start of the academic year. In the previously offered National Diploma, learners who were accepted from the TVET/FET Colleges appeared to perform satisfactorily, if not better than the learners straight out of school. This perception, as mentioned previously, has, unfortunately never been validated through formal evidence-based research. Although the Department has specified minimum entrance requirements, the Department does select learners strictly on merit, when numbers exceed the Departments requirements. The Department, like the other Departments at DUT uses the Central Applications Office (CAO) to process and accept learners. The Department does have an application cut-off date, but invariably has had to wait until after lectures have begun before walk-in candidates find their way to the Department to attempt to fill the 120 available places.

Table 1.1: First Year Entering Learners

BEng Tech: Power Engineering	
Year	**No of First Year Learners**
2018	97
2019	116
2020	112
2021	70

1.4 Purpose of the Study

As mentioned previously, there has been a prevailing perception that with respect to Mathematics and Physical Science at higher education level, articulating TVET/FET learners have coped as well as their first-time accepted counterparts - the SC/NSC learners. In January 2018, the BEng Tech Degree qualification was introduced by the DUT's Department of Electrical Power Engineering, and for the first time in twenty years the acceptance of TVET/FET college learners was queried and debated. This resulted in no TVET/FET learners being accepted into the first year of the degree. Other departments within the Faculty of Engineering and The Built Environment continued to accept articulating learners into the degree. This decision was later reversed during 2018 and TVET learners were again allowed to register in DUT's Department of Electrical Power Engineering Programme from January 2019.

As the success rates of articulating TVET/FET learners remained undetermined, and in the light of the DEPE's previous decisions, it became necessary to measure the performance of these learners rather than rely on anecdotal perceptions. This study is being conducted to establish the effectiveness of articulation, and more specifically to establish a system to predict the success of TVET learners in higher education engineering programmes. Tracking is required to determine the performance of these learners - in particular, to determine if there is a statistically significant difference between articulated learners and learners admitted with only an SC or NSC.

This study aims to advance our knowledge of supplementary articulation educational routes to inform both policy and practice. It intends to explore the reliability of a predictable mechanism with respect to the future performance of a learner, considering the numerous factors that influence failure/success. In particular, the research plans to whether TVET Mathematics and Engineering Science subjects have an impact on learner performance in the university engineering programme under study.

1.5 Hypothesis and Aim of Study

The hypothesis being tested in this study is that, there is no statistically significant difference in the performance of SC/NSC entrance learners and N4 TVET learners, both in terms of their progression from semester 1 into semester 2 of the qualification, and in terms of their overall throughput rates.

The delineation of this hypothesis is both retrospective and prospective. For aims 1, 2, 3, 6, 7 and 8, the entrance data, first semester results, and the throughput results are used for comparison. For aims 4 and 5, the entrance data are utilised to design a machine learning tool to predict the future success of applicants.

The aims of this research are:

1 Conduct a longitudinal study on the performance of DUT learners articulating from a TVET college, from semester 1 into semester 2 of the National Diploma: Electrical Engineering programme.

2 Conduct a comparative study on the performance from semester 1 into semester 2, between the aforementioned learners and those learners who were admitted into the National Diploma: Electrical Engineering programme at DUT directly having achieved a SC or NSC.

3 Review and rank, using Educational Data Mining (EDM), for both articulating TVET and SC/NSC learners, the subjects affecting progression from semester 1 into semester 2.

4 Review and analyse, using the results from the TVET subjects as input variables, methods of Educational Data Mining (EDM) that can be used to select learners from TVET Colleges based on progression from semester 1 into semester 2.

5 Review and analyse methods of Educational Data Mining (EDM) that can be used to select learners from SC/NSC, based on progression from semester 1 into semester 2 using the results from the SC/NSC subjects as input variables.

6 Conduct a basic qualitative comparative analysis of the Science and Mathematics (STEM) subjects from FET and Umalusi curricula.

7 Conduct a longitudinal analysis of throughput of articulating TVET learners and SC/NSC learners, in the National Diploma: Electrical Engineering programme over five separate cohorts as per the DHET reporting standard.

8 Track the aforementioned cohorts of learners registering for the first time in the National Diploma: Electrical Engineering programme at DUT and completing in minimum time, minimum time + 1 year and minimum time + 2 years.

Although this research is primarily focused on the performance of a particular learner cohort from one qualification, the background to and the appreciation of the reasons for, opinions of, and motivations for an articulation route to the HE education sector must be understood. A detailed literature review is undertaken, in the next chapter, to ensure that the qualitative areas of the research will be located within an identified lacuna in the literature review. To achieve the quantitative results of this study readily available data from DUT's management information system (MIS) will be obtained and analysed and used to develop predictive models to determine the success of a cohort of learners. This model should thereafter enable the prediction of future learner success before entering DUT, with possible generalisation to other HEIs.

1.6 Structure of the Dissertation

Chapter 1 - Introduction

Chapter 1 introduced the background to the problem. It discussed the purpose and aim of the study. It outlined the basic methodology that will be conducted to fulfill the requirements of the research. It also introduced articulation in relation to the DEPE.

Chapter 2 - Literature Review

Chapter 2 is a detailed literature review of education within South Africa, concentrating on articulation. The literature review describes the status of articulation and the need for articulation in the South African context.

Chapter 3 - Methodology

Chapter 3 starts by introducing Data mining and education data mining. As part of the methodology artificial neural networks are described and reviewed as a method for educational data mining and a means of analysing the problem. The methodology followed is described in detail to ensure understanding and possible further study.

Chapter 4 - Results

Chapter 4 elaborates in detail the results of the analysis.

Chapter 5 - Conclusion

Chapter 5 concludes this research while identifying possible future research.

Chapter 2 - Literature Review

2.1 Introduction

The purpose of the first chapter was to provide a brief historical background to the set of circumstances that motivated the Department of Electrical Power Engineering at DUT to adopt a little-known policy, the Articulation Policy for the Post-School Education and Training (PSET) system of South Africa in terms of Section 8(2)(b) of the NQF Act, 2008 (Act 67 of 2008), referred to as articulation which: (i) enables learners to achieve their aim of entering a higher education programme of their choice; (ii) provides an alternative means to meet enrolment targets [12]; (iii) ensures that levels of participation in STEM subjects are maintained [23]; and (iv) sustains the pipeline supplying industry with sufficient graduates. Chapter 2 begins with a brief outline of the educational history of South Africa and unfurls the present post-school education and training education environment, with special attention being paid to articulation within the sector.

2.2 History of Higher Education in South Africa

2.2.1 South Africa, A Historical Overview Pre-1994

The Union of South Africa was a country that operated from 1910 with several education systems that were primarily initiated by British churches and missionaries [14, 24, 25]. A somewhat segregated country [15], at the time, it became more segregated along racial lines in 1948 [26], when the National Party achieved political power. From 1961, as the Republic of South Africa, basic education was provided under a binary system, at provincial level, for Black, Coloured and Indian populations, through the Department of Education and Culture. Higher education institutions for Black, Coloured and Indian learners were few in numbers, had limited qualification offerings and were underfunded [13]. White English and Afrikaans learners had access to several separate well-funded institutions with many offerings provided at a national level [27, 28] by the Department of National Education. The Nationalist

Government separated higher education institution into Technikons, Colleges and Universities [12, 29]. Technikons provided training for technologists and technicians and awarded diplomas and certificates; Colleges provided training in nursing, police, agriculture, and teacher education, awarding certificates and diplomas. Technical Colleges or Further Education and Training Colleges (FET) also awarded certificates and diplomas, namely, agriculture, civil, computer, electrical, information and mechanical related industries for grades 10 through to 12 [14]. Universities provided general and professional education and awarded degrees at undergraduate and postgraduate levels [12, 29, 30].

2.2.2 South Africa, A Historical Overview Post-1994

As early as 1991 research revealed that the education system in South Africa was recognised by many to be unjust and unfair and had been operating under severe financial constraints [31]. Following the historic 1994 elections the African National Congress (ANC) led government began the process of altering the old order to create a more equitable society in the Republic [28, 32, 33]. This included the restructuring of the education system as it was recognised that personal education accounts for an improvement in health, a decrease in infant mortality, positively impacts on labour productivity, returns higher wages and reduces poverty – aspects that had negatively affected the Black, Indian and Coloured populations as historically the education system had been restrictive [19] and limited the prospects of learners of colour reaching and achieving a higher education [3].

"The Constitution of South Africa, 1996 – Chapter 2: Bill of Rights, Section 7" [34] defines the rights for every South African. It enshrines the rights of all people in South Africa and affirms the democratic values of human dignity, equality, and freedom. It further states that everyone has the right to freedom of conscience, religion, thought, belief, and opinion. These statements alone, in essence, require that the previous education system be completely changed as its epistemology was based on indoctrination and control. This is confirmed in the very clear statement of The Constitution of South Africa– Chapter 2: Bill of Rights [34].

In Section 29 [35] in which it stated that everyone has the right to basic education, including basic adult education and to further education which, through reasonable measures, must be made progressively available and accessible.

The post-1994 South African education landscape is governed by a legislative framework called the National Qualification Framework (NQF), ACT 67 of 2008 [36]. This Act applies to all educational programmes that lead to qualifications or part qualifications offered in the Republic of South Africa. Specifically, the objectives of the NQF are to facilitate access to, and mobility across and progression within, education, training, and career paths. This framework is sub-divided into three distinct sub-frameworks: General and Further Education and Training Sub-Framework (GFETQSF), Higher Education Qualifications Sub-Framework (HEQSF), and the Occupational Qualifications Sub-Framework (OQSF). Whilst the NQF promotes the concepts of progression and articulation, the realisation of these principles between the three sub-frameworks has been limited. This is due to a lack of interaction caused by insufficient data on the potential success of articulation and progression between the three frameworks. No research could be found to the success of articulation and progression between various types of institutions within the sub-frameworks [15].

Higher education, post-1994, further underwent change in relation to the demographics in former white higher educational institutions. Despite the massification of learners into the higher education system, and a change in the demographic membership of engineering professional bodies [19], there has, however, been a reduction in throughput rates in higher education, and what Bundy (2006) [23] from Cooper & Subotzky (2003) [37] referred to as "skewed revolution" – a situation that has previously disadvantaged learners enrol in humanities rather that in business, technology and science degrees. As recognised by the Ministries of Labour, Trade and Industry, the economy and infrastructure of South Africa requires engineering professionals to aid in the development of the country [38]. The critical shortage of skilled and qualified professional engineers, is an encumbrance to the growth of the economy [39]. In South Africa there are approximately 315 registered engineers per million population. Japan, in comparison, has 3300; while the United Kingdom and Norway have 3220 and 8190 engineers per million of their respective populations [40]. This very basic statistic is comparable to the relative health of the economy of a country indicating that South Africa is drastically short of engineers. Trends also show that the uptake of qualifications for commerce far exceeds that of engineering [3]. The Trends in International Maths and Science Study (TIMSS 2015) [2] assessment for Grade 8 learners indicated that South African came last for science, and second to last for mathematics in 2015 [41]. This set of results most probably reflects learners, with a SC or NSC, leaving the school system with very weak mathematics and physical science marks – subjects that constitute part of the entrance criteria for engineering programmes at South African HEIs. Failure to meet these requirements means applicants are denied admission to HE engineering qualifications. [8] However, the same learners, if enabled to articulate into a HE through an alternate route such as enrolling in an engineering N4 Certificate to obtain the relevant knowledge and necessary marks, would be given another opportunity to achieve the entrance requirements.

Following extensive consultation with several stakeholders in 1996, a report titled 'A Framework for Transformation' by the Council on Higher Education [42], identified three central attributes for a future, transformed, unitary higher education system [13, 23]. Along with proposals in the Green Paper of 1996 [43] and the Education White paper 3 of 1997 [44], the National Education Policy Act [45] and confirmed in the National Plan on Higher Education (NPHE) [11] a number of broad proposals regarding the transformation of the higher education sector was set out. It was the third proposal which spoke to "increased access and higher participation rates", that is of importance to this study, and the framework and policies from which, the Articulation Policy for Post-School Education and Training System [46, 47] of South Africa developed.

The National Qualification Framework (NQF), ACT 67 of 2008 [36], one could argue, is the legislative foundation of the principle of articulation. The NQF mentions that South African Qualifications Authority (SAQA) and the Qualifications Councils (QC) must seek to achieve the objective of the NQF by developing, fostering and maintaining an integrated and transparent national framework for the recognition of learning achievements [36]. All spheres of education should therefore work together to enable the mobility of learners across institutions and qualifications within the PSET landscape. As mentioned previously, the NQF

[2] TIMSS data have been collected from students at grades 4 and 8 every 4 years since 1995, with the United States participating in every administration of TIMSS. TIMSS Advanced studies the achievement in advanced mathematics and physics of students in their final year of secondary school. It has been conducted in 1995, 2008, and 2015, with the United States participating in 1995 and 2015.

incorporates the three frameworks of post-school education – the General and Further Education and Training [48], contemplated in the GENFETQA Act; Higher Education, contemplated in the Higher Education Act [49] and Trades and Occupations, contemplated in the Skills Development Act [50]. It is possible and should therefore be imperative that systems and practices be developed to assist learners transitioning between institutions and / or qualifications.

The Higher Education Qualifications Framework (HEQF), set out as a policy in terms of Section 3 of the Higher Education Act, 1997 (Act 101 of 1997) [49] should:

- "Facilitate the education of graduates who will contribute to the social, cultural and economic development of South Africa and participate successfully in the global economy and knowledge society;"
- "Be simple, clear, easy to understand and user-friendly for the higher education system and its clients;"
- "Facilitate qualification articulation across the higher education system and assist learners to identify potential progression routes, particularly in the context of lifelong learning;"

The HEQF, which clearly pronounces articulation across higher education qualifications as an imperative, incorporates credit accumulation and transfer (CAT) [18, 49]. This is a process whereby a learner's achievements are recognised and may contribute to further learning - even if a qualification is not achieved. The policy describes how at least 50% of credits obtained in one qualification may be transferred to another.

2.3 Articulation

As described previously, many learners exit the South African basic schooling / education system in possession of a SC or NSC, but without the necessary entrance requirements to be admitted into HE. Given the myriad reasons for these learners not having been able to obtain the minimum requirements necessary for admittance to many HE programmes as well as the dire need for South Africa to produce graduates in STEM fields, leaners should be provided with simpler more streamlined means to obtain a second chance to be admitted to a HEI without having to spend at least one year repeating their entire Grade 12 suite of subjects. Articulation is a means by which this may be realised.

Articulation is defined in the Articulation Policy for the Post-School Education and Training system of South Africa in terms of Section 8(2)(b) of the NQF Act, 2008 (Act 67 of 2008) [47]. The Articulation Policy establishes the overarching conceptual structure, principles and policy statements to support the implementation of credible approaches to articulation within the South African PSET system. The Policy summarises articulation for South African Higher Education as:

> “A well-articulated system is one in which there are linkages between its different parts: there should be no silos, no dead ends. If a student completes a course at one institution and having gained the relevant knowledge and skills at the necessary levels, their achievements must be recognized by other institutions if the knowledge gained is sufficient to allow epistemological access to programme(s) that the student wants to enter. Learners must be supported in their individual learning and work pathways” [47].

The Articulation Policy [47, 49] underpins the principles of the constitutional right to education through references to mobility, progression within education and training, and establishing career paths for learners. In terms of the Policy [47], articulation should occur “between the three qualifications sub-frameworks, across all levels of the NQF and learning institutions. A common core curriculum [47], at NQF 4 level must underpin the further development and design of the National Senior Certificate, the National Certificate Vocational (NCV), the National Senior Certificate for Adults (NASCA) and other exit qualifications at NQF level 4”.

As such all institutions, at all levels, must work together to ensure that learning pathways are developed and that they are most practically and logically linked to facilitate a smooth learning pathway for learners wanting to further their levels of education.

The scope of the Articulation Policy [47] clearly encompasses Basic Education, TVET and HE, therefore encompassing all spheres of this research. The Articulation Policy defines articulation in 3 ways; as either *systemic* articulation – that which is based on Government legislation or various other official elements aligned to and supportive of, learning and work pathways; or *specific* articulation – that which is based on agreements between institutions within the education system, such as Memoranda of Understanding (MOU) and thirdly pathways followed by individuals as they progress and they would be supported by institutions [47].

In the absence of any systemic articulation policies other than the National Policy itself, articulation is therefore specific and relies on the addressing of 'boundary-crossing practices' [15]. Learners encounter boundary zones between differing elements of learning pathways and will need to utilize 'boundary-crossing practices' in navigating their learning pathways [15, 51]. These practices or support reduce the gap between learning pathways-related policy development and implementation' strengthening specific pathways and enhancing the opportunities to progress along these pathways' the quality of education and training' Flexible Learning and Teaching Provision' appropriate career advice and other various types of support required [51].

Blom [52], describes articulation as being deceptively simple but does admit that deliberate and considered effort has to take place to enable learners' progression along practical and logical learning and career pathways. Blom [52], proffers a description of 'Articulation as a Stairwell' as a/another method of articulation. This method allows second-chance learners, who did not meet the requirements for entry into a programme on their first attempt, a second chance. Such an example would be enabling learners from the Basic Education system another opportunity to obtain their desired Higher Education programmes entrance requirements by completing relevant, and outstanding subjects at a TVET college. The Department of Electrical Power Engineering at DUT developed an 'Articulation as a Stairwell' initiative, in terms of which learners who did not meet all the admission requirements, were directed to TVET colleges to register for those subjects' preventing admission to their first-choice qualification. Upon completion of these subjects, learners articulate into the National Diploma: Electrical Engineering at DUT. This approach was predicated upon the critical shortage of learners in possession of the required Mathematics and Physical Science results to be admitted to the Department. As this shortage is experienced nationally by higher education science and engineering programme [3], learners from TVET colleges constitute a potential source of learners for HE programmes in STEM.

The following principles are outlined in the Articulation Policy [47]:

1. Systemic Articulation is based on national policy and legislation while Specific Articulation is local agreements within the education and training system.
2. Articulation must assist education in meeting the needs of the South African economy. Addressing skills shortages and pathways must be established for learners to migrate through the system.
3. Articulation must address continued and lifelong learning. Education should be efficient for the citizens of South African to reach their preferred educational destination.
4. Articulation must ensure redress, equity, and inclusiveness. At no point should any individual not have the right to education, and it should go as far as to assist those that need it most.
5. Articulation ensures programmatic rather than institutional solutions. All spheres of education should be assisting learners to migrate between institutions.
6. Articulation must be straight forward and ensure accessibility. Too often the system is designed to be complicated and to exclude rather than being simple and inclusive.
7. Articulation must promote learning outcomes. All similar looking outcomes from formal or non-formal systems should be equitable.
8. Articulation routes, policies and principles should be predesigned. Educators in the know should trouble shoot and design articulation routes within the educational frameworks to assist learners.
9. Articulation must be credible and fair. All practices should be credible, relevant, reliable, fair and transparent.

The implementation of articulation clearly requires a commitment at multiple levels and collaboration from the following stakeholders: Council for Higher Education (CHE); Umalusi, the Quality assurance body for the NSC; and Quality Council for Trades and Occupations

(QCTO). In addition, other government departments, employers, educational and training Institutions, skill development providers and learners need to develop, foster and maintain the aims and objectives of an integrated and transparent education system in line with the NQF.

Similar articulation policies and frameworks exist internationally. In the Australian Qualifications Framework (AQF) [53], there is a pathway policy that has a very similar purpose to the South African Articulation Policy, that is to maximise the credit learners may gain for learning already undertaken. This policy applies to learner progression into and between AQF qualifications. It recognises that there are multiple pathways for learners and these pathways may by formal and non-formal. The policy emphasises the need to support pathway development in qualification design. The United Kingdom have a long tradition of institutional autonomy and although flexible learning pathways are supported in principle, policy dictates what is allowed rather than what is required. Credit transfer relies of arrangements between institutions and learners transfers are limited, mainly due to a potential loss in funding [54] .

2.3.1 DUT Baseline Articulation Study 2017

The SAQA-DUT Baseline report [15], resulted from a research partnership between SAQA and DUT, was designed to provide extensive information and evidence to support the implementation of the DHET's Draft Articulation Policy and SAQA's NQF policy suite, as listed by SAQA's objectives listed in the NQF document [36], Section 11 which are:

(a) "advance the objectives of the NQF"
(b) "oversee the further development and implementation of the NQF and"
(c) "co-ordinate the sub frameworks."

It is therefore SAQA' s mandate and duties to fulfil the further development, organisation and governance of the NQF. Regarding to the functions as seen in Section 13, specifically 13 (1)(k) in the NQF, with respect to research.

(i) “Conduct or commission investigations on issues of importance to the development and implementation of the NQF, including periodic studies of the impact of the NQF on South African education, training, and employment; and

(ii) “Publish the finding of the investigations referred to in subparagraph (i)”.

Before the long-term research partnership between SAQA and DUT[3] could proceed, it was realised that a baseline study was required. This resulted in the National Articulation Baseline Study Report (NABS) (SAQA 2017), [15] which investigated articulation between Technical and Vocational Education and Training (TVET) Colleges and Higher Education Institutions (HEIs).

The Project sought to:

i. “Identify, analyse and document good practice models and relationships for learners transitioning between (a) TVET Colleges and HEIs, (b) between UoTs and traditional universities, and (c) between Colleges or HEIs, into workplaces:”
ii. “Identify and explore the potential for developing collaborative models in three types of articulation scenarios: ‘developed’, ‘emerging’ and ‘latent’, in order to take the good practice models and relationships to scale' and”
iii. “Identify the nature of activities and support that institutions, staff, and learners need for successful transitioning.”

The in-depth study of the 50 public TVET Colleges and 26 public HEIs was conducted to determine their understanding of articulation, descriptions of existing articulation arrangements, and challenges resulting from articulation. It also included an investigation into existing successes and enablers in practice and was aimed at determining the extent to which accurate tracking of learners’ movements into, through and out of the institutions were taking place.

[3] “Developing understanding of the enablers of student transitioning between TVET Colleges and Higher Education Institutions and beyond”

The following is a summary of the finding of the Baseline study:

1. There is generally a high awareness of articulation, but less understanding of the flexible support required by learners.
2. Little understanding of all types of Articulation and enabling aspects thereof.
3. Most institutions, TVET & HEI, reported that articulation was occurring, although mainly informal and failure to adhere to formal agreements resulted in challenges.
4. Lack of tracking of learners by TVET Colleges and Universities and the required guidelines thereof.
5. Varied levels of Articulation across institutions.
6. Varied perceptions on the success of Articulation.
7. Several enablers to Articulation were identified.

The above findings of the study resulted in 7 recommendations:

1. Successful articulation models need to be further investigated and documented.
2. Research conferences should be convened to promote system wide sharing and learning.
3. Articulation champions/offices need to be established within institutions to drive articulation.
4. The development of collaborative relationships at both provincial and national level should be explored.
5. Greater systemic articulation should be facilitated.
6. Successful transition points should be replicated and where possible upscaled.
7. Systematic reporting requirements and guidelines for institutions to track and report on articulation practices should be developed.

It is specifically points 4 “Lack of tracking of learners by TVET Colleges and Universities and the required guidelines thereof.” and 6 “Varied perceptions on the success of Articulation.” of the findings and point 7 “Systematic reporting requirements and guidelines for institutions to track and report on articulation practices should be developed.” of the recommendations which align with the purpose of this was study.

2.3.2 The Third NQF Conference

Building on the outcomes, findings and recommendations of the SAQA-DUT Research partnership and other previous studies, namely the SAQA-Rhodes University Partnership on Learning Pathways, SAQA-University of Western Cape into Recognition of Prior Learning (RPL), SAQA-University of Western Cape research into Lifelong Learning and Flexible Provision and SAQA-University of Kwazulu-Natal research into TVET College lecturers, various conferences have been convened to ensure the momentum is maintained. One such conference was SAQA's Third NQF Conference [20], titled "Strengthening Learning-and-Work Pathways in Community Development, Early Childhood Development and Engineering" reinforced the need to deepen the culture of articulation within the country. The conference attended by the major role players in education, namely Department of Basic Education (DBE), SAQA, CHE, QCTO and Umalusi, Council for Quality Assurance in General and Further Education and Training was convened to reflect on: Community Development, Early Childhood Development and of interest to this investigation, Engineering. The main objectives of the conference were to further enhance existing articulated learning and work, pathways, and obtain commitment from key role-players for implementation.

Dr Hersheela Narsee, at the official opening of the conference [20], on behalf of the Minister of Education Science and Innovation, Dr Blade Nziimande, observed that it is through reflection at such events that policy may be addressed to remove the constraints to articulation. She challenged participants to ensure that the National Plan for PSET be improved for articulation to broaden access, efficiency, success, quality, and equity.

Ms Trudi van Wyk, Chief Director: Social Inclusion, Equity, Access, Quality, Open Learning and Career Development Services (DHET) intimated that "the NQF and articulation are about learner access and progression". She advocated for open learning, the removal of barriers and access to RPL and learner support, arguing that educators need to focus on implementation and the simplification of the system and provide guidance to learners.

Dr Amani Saidi, Director of Quality Assurance and Promotion in the Executive Office of the Centre for Higher Education (CHE), presented on articulation in HE. The points that were highlighted were:

- Articulation enables flexible lifelong learning with no dead-end qualifications.
- Integration and Parity of esteem should be promoted across all institutions in the NQF sub-frameworks.
- Articulation requires promotion, facilitation, and support.

Dr Saidi specifically pointed out that HEIs give preference to learners with NSC over those with occupational qualifications. He intimated that the CHE is seeking ways to identify and eliminate 'dead ends', working with HEIs to develop and implement new progressive intuitional policies for access, RPL, CAT and articulation. Dr Saidi highlighted four areas of difficulty regarding articulation that have been identified by the CHE:

- Articulation Gap – Learners from NSC / SC / NCV battle to transition into HE due to low knowledge, skill, literacies, or social attributes.
- Research – Limited research exploring articulation barriers.
- Incentives – Lack of incentives for HEI to engage in articulation practices.
- Monitoring – Lack of articulation monitoring within the PSET system.

Mr Marco Macfarlane, Director for Research and Analysis at QCTO asserted that articulating artisans into engineering qualifications was still a challenge and that discussion were underway between the QCTO and the CHE regarding the alignment of Occupational Qualifications to meet the admission to HEQSF qualifications. Further research and developments in these areas were taking place, so that artisans could articulate between the OQSF and the HEQSF.

Ms Omotola Akindolani, Deputy Director: Data Analysis, Reports and Publications (SAQA), presented a report on the analysis of learning pathways of specific interest in engineering from data obtained from the National Learners Record Database (NLRD). She informed the conference that the NLRD is the information system of the NQF, having data for 19.5 million people; 11 000 currently registered qualifications; 106 recognised professional bodies and 385 professional designations. She explained that 'learning pathways' allow learners to move vertically, diagonally and horizontally through NQF levels giving learners recognition for part of full completion of qualifications or part-qualifications.

Ms Akindolani presented the following data analysis for ECSA registered Engineers, Technologists and Technicians:

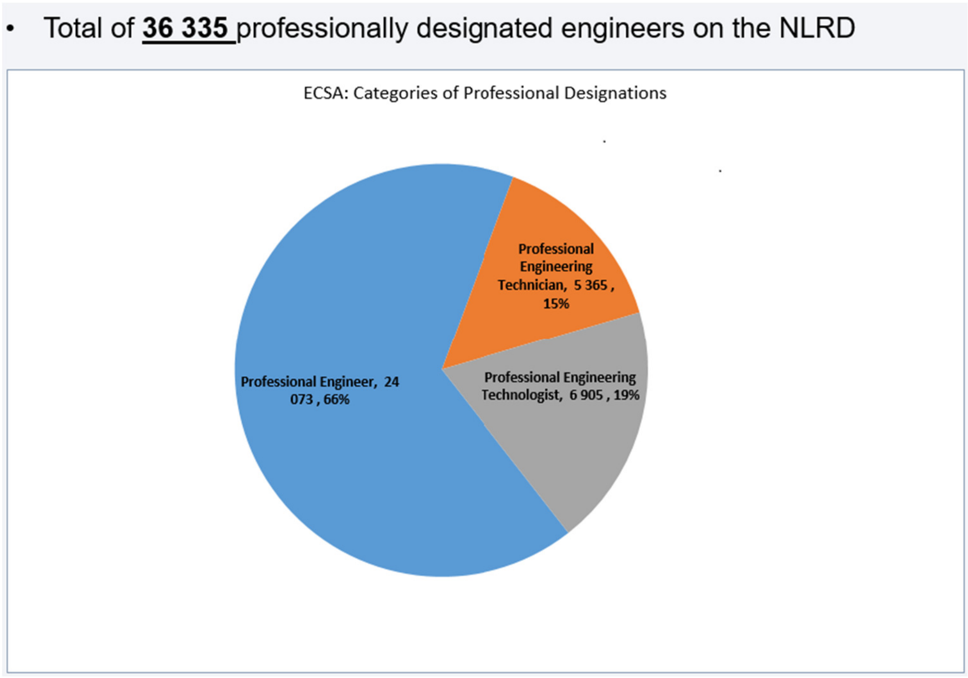


Figure 2.1: Professional Designated Engineers on the NLRD [20]

Figure 2.1 reflects the drastic shortage of registered engineers [20], as highlighted by Lawless (2005) [40], with a total of only 36335 professionally designated engineers on the NLRD.

The numbers of registered technicians and technologists are relatively low compared to professional engineers as seen in Figure 2.2 [20].

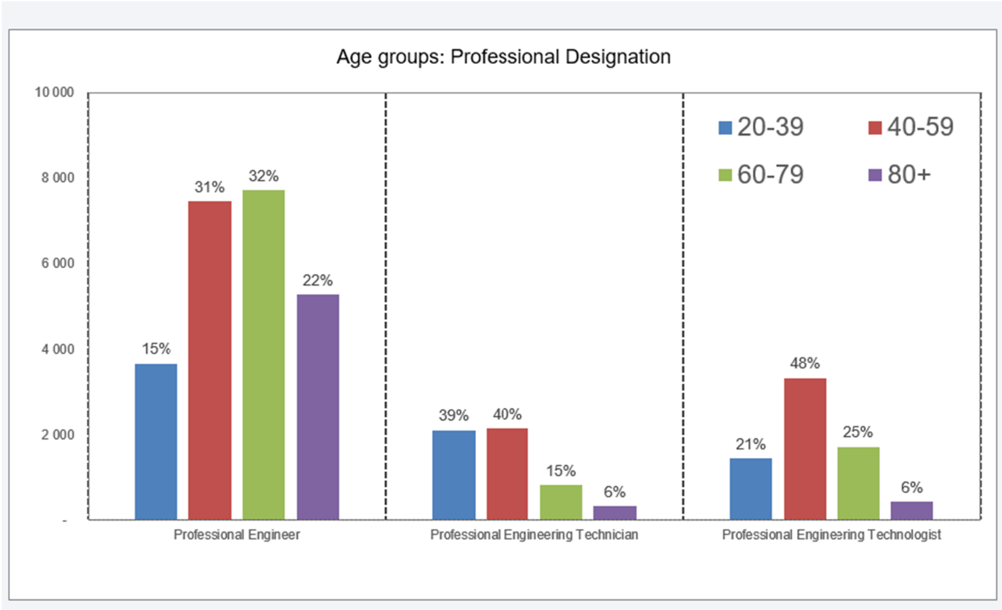


Figure 2.2: Age Profile of Engineering Professionals [20]

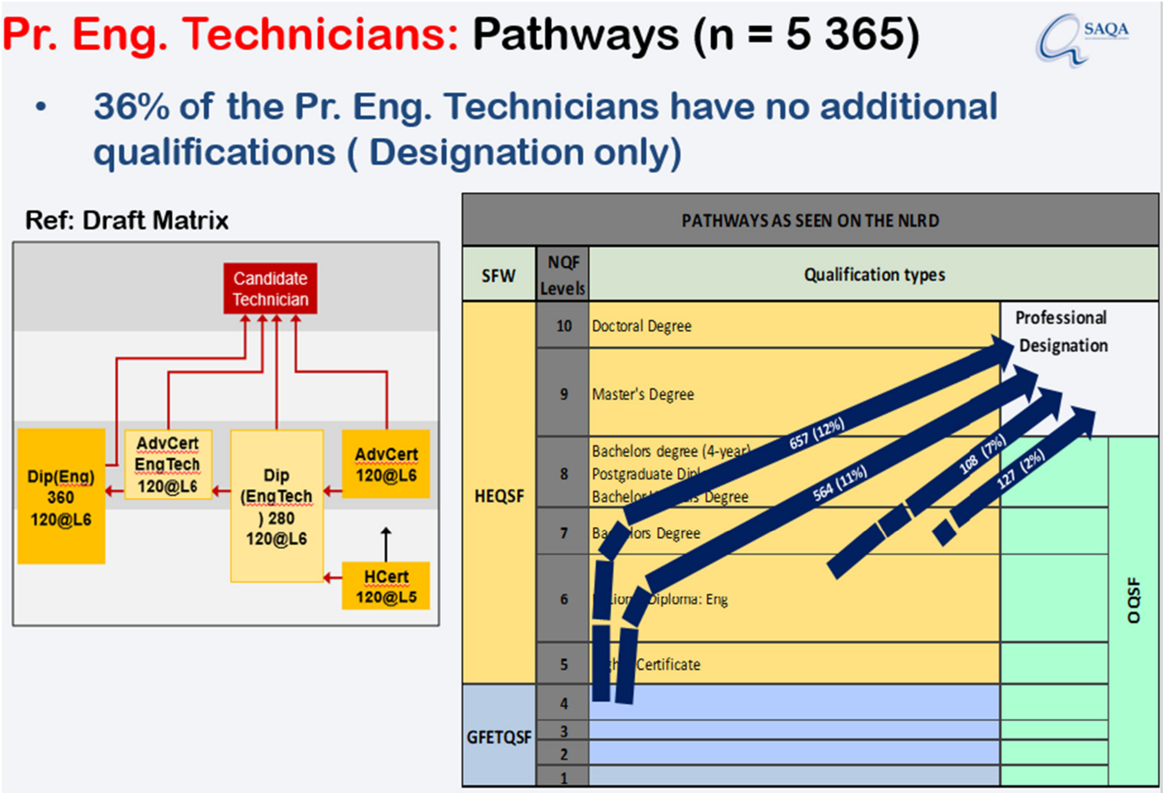


Figure 2.3: NQF Paths - Possible vs Actual [20]

Figure 2.3 shows possible pathways shown on the left and actual pathways shown on the right for technicians [20].

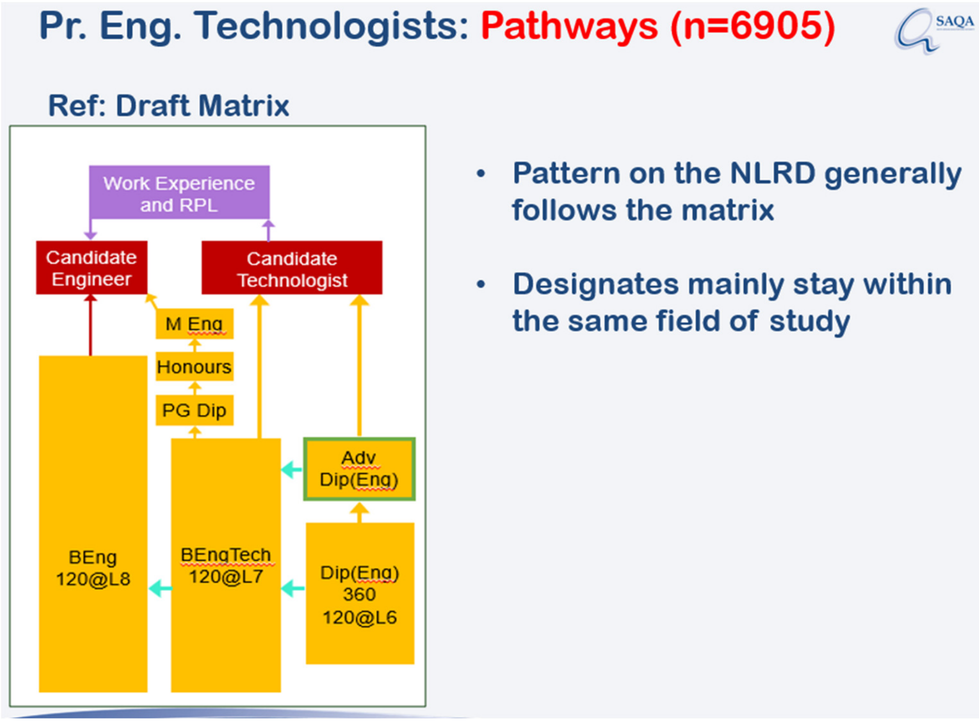


Figure 2.4: NQF Possible Pathways for Technologists [20]

The NLRD shows that individuals with professional designations have generally followed qualifications pathways outlined / required by professional bodies, as seen in Figure 2.4 [20].

Figure 2.5, shows the graphical layout being develop by SAQA for the various articulation methods [20].

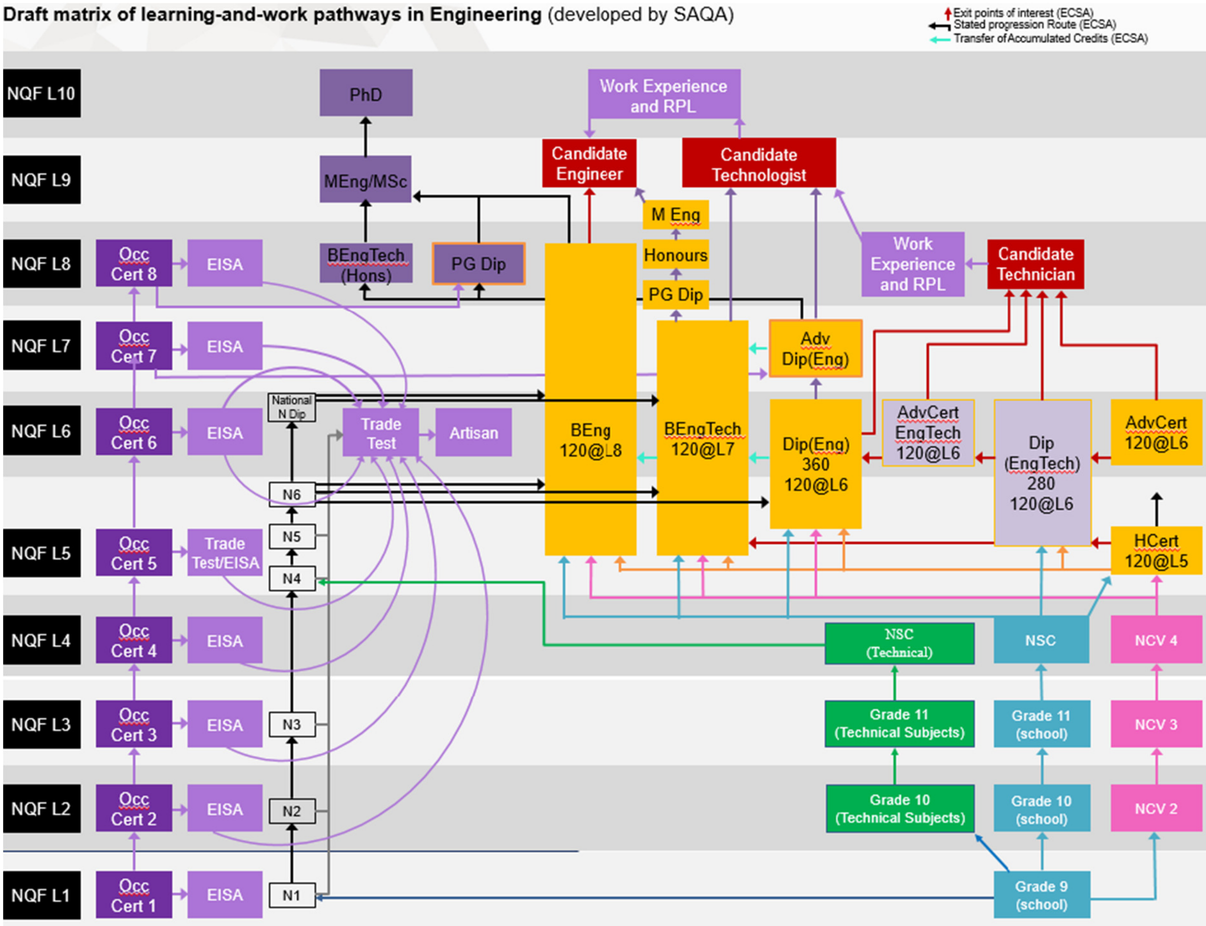


Figure 2.5: Engineering Learning Pathways [20]

Mr. Jones Moloisane, representing the Engineering Council of South Africa (ECSA) provided an overview of the engineering sector. He reported that since 2000 it has been compulsory for all Engineers, Technologists, Technicians and Artisans to be registered with ECSA. He noted that although ECSA was short-staffed for optimal RPL processing, it did conduct RPL for candidates with experience without formal qualifications. He presented the pathways map, shown in Figure 2.6 below [55], developed by ECSA in collaboration with SAQA, which runs NCV2 all the way to PhD.

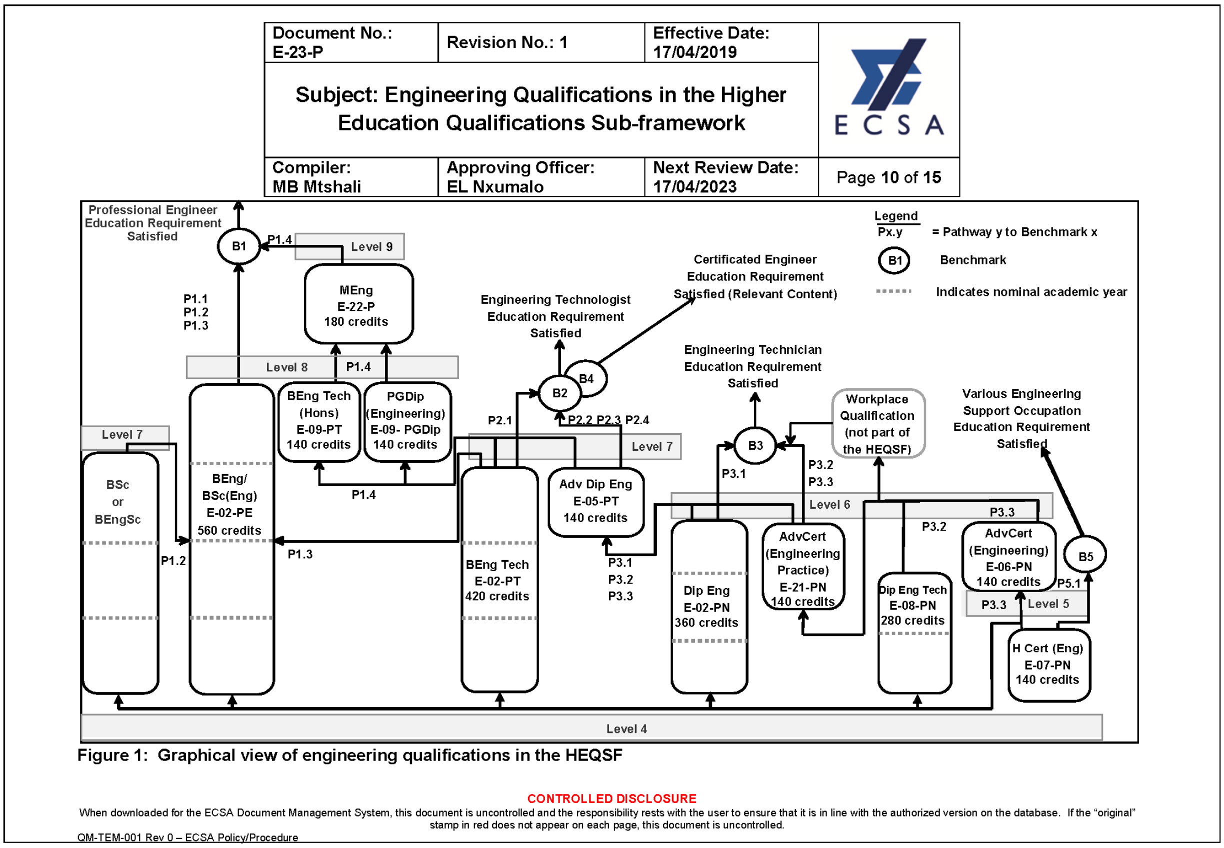


Figure 2.6: Engineering Qualification Graphical View [55]

From the Engineering sector articulation sector plan, it was clear that the role of industry was essential and that research into the role articulation was required. In certain Engineering fields the technician to technologist articulation route already worked well, while in the BEng, articulation did not work well and ECSA is currently working on that standard. The Engineering Professions Act (Act 46 of 2000) already caters for RPL [56]. The sector plan highlighted the need for articulation to be embedded into the education system rather than be left to individual decision-making and that articualtion between TVET and HE required attention. Mr Moloisane asseverated that the National Development Plan (NDP) shapes the engineering agenda in terms of infrastructure and a statistical analysis was underway to determine the requirements for the future. He reported that the number of applicants from high school with passes in Mathematics and Science was shrinking and needed to be addressed. Analysis is required using NLRD, STATS-SA, Labour Force Surveys and Sector Skills planning to determine skills requirements.

On the way forward the following points on articulation points were highlighted:

- Resources for articulation need to be allocated.
- Colloborative relationships for curriculum alignment and advocacy of learning-and-work pathways should be developed.
- Formal articulation agreements should be implemented and sustained.
- Admission criteria need to be more inclusive.
- Flexible learner progression should be enabled.
- Further research in articulation should be undertaken.
- Good Practice should be documented and reported.

This dissertation embodies some of the points above .

2.6 Conclusion

This literature review highlighted the economic and social importance of higher education for the development of South Africa and its people. The right to education is ensconced in the South African Constitution [35] and enacted by the National Qualifications Framework [36]. The Constitution further underscores individuals' access to these rights and places fiduciary responsibility at the door of educational institutions to assist individuals with their educational aspirations. Articulation is a critical route for some learners to achieve their educational goals, and a route that all stakeholders should promote to assist learners to meet their individual needs, ambitions, and educational aspirations.

At time of writing, no quantitative data existed on the performance of learners from TVET colleges in an engineering programme at DUT. No memorandum of understanding (MOU) – or specific articulation agreement as per Blom's definition, or collaboration existed between DUT and TVET Colleges. All that is being utilised and has been utilized, from 1994, is the Department of Electrical Power Engineering entrance requirements captured in the Departmental Handbook, Rule: 4 (b)(ii) "National N4 Certificate - N4 Certificate with passes at 50% in four (4) relevant subjects including Mathematics and Electro-techniques, or an equivalent SAQA NQF Level 4 qualification, as well as compliance with the English language requirements as stated in the General Rules" at DUT, that allows articulating learners acceptance into the BEng Tech Power Engineering Degree (previously the National Diploma: Electrical Engineering).

Recommendation 7 of the National Articulation Baseline Survey Report [15] identified the need for institutions to track and report their articulation practices. It further recommended that systems be upgraded to include information relating to articulation.

This evidence-based research study aims to further investigate and advance the knowledge of education policy to improve policy implementation and practice. In doing so, this research intends to search for a predictable mechanism with respect to the future performances of transitioning learners whilst taking into consideration the numerous factors that influence learner performance. It also plans to identify inputs that best contribute to learner academic success and retention. In particular the extent to which good performance in TVET College N4 Engineering Mathematics, N4 Engineering Science, N4 Electrotechnics and N4 Industrial Electronics subjects correlates with learner performance in an Engineering qualification at DUT. This research therefore intends to quantify the success of articulating learners and to develop a predictive tool to assist with selection of articulating learners from TVET colleges.

Chapter 3 - Methodology and Methods

3.1 Introduction

This chapter begins with a description of the methodology used in this study. It is also serves as an extension of the literature review in Chapter Two.in which Big Data, Data Mining, and interferential statistical analysis; methods that are relevant to the quantitative methodology used in this study, are described. It explicates, specifically Educational Data Mining techniques, that may be used to predict the success of TVET learners that have articulated into the National Diploma: Electrical Engineering at DUT. This is followed by an explanation and techniques of how the data was selected, captured, and analysed.

3.2 Ethical Approval

The research proposal for this study was reviewed by the DUT Faculty Research Committee for Engineering and The Built Environment. Approval was given on the 12th June 2020, with ethical clearance category 1.

3.3 Methodology

Research is a process in which data is collected, analysed and interpreted for the understanding, prediction and control of an educational or psychological phenomenon or to empower individuals in such contexts [57, 58]. The theoretical framework, separate from the theory, refers to the paradigm [59, 60], that in which the study is positioned, and provides direction for the intention, motivation and expectations of the researcher [61]. Positivism, one such paradigm, referred to as the 'scientific method' or 'science research', is 'based on a rationalistic, empiricist philosophy in which cause effects outcome [62]. Positivists test a theory "through observation or measurement, in order to control forces that surround us" [58]. Replaced by Postpositivism after World War II, "Postpositivists work from the assumption that any piece of research is influenced by a number of well-developed theories apart from, and as well as, the one which is being tested [63].” Postpositivist and Positivist research is most commonly aligned with quantitative methods of data collection and analysis.

Research literature uses quantitative and qualitative as two distinct terms, but often loosely uses method and methodology interchangeably [64]. O’Leary (2004) defines the terms as

> “adjectives for types of data and their corresponding modes of analysis, i.e. qualitative – data represented through words, pictures, or icons analysed thematic exploration' and quantitative data – data that is represented though numbers and analysed using statistics [58].”

This definition prefers that these terms are linked to collection methods, analysis and reporting, rather than the theoretical approach to the research. This studies has been conducted using academic and quantitative literacies [65]. Using the above definition, the author will apply the quantitative postpositivist method to this research, as all data, obtained from DUT’s MIS is in the form of numbers. It will be analysed using statistics and educational data mining.

3.4 Big Data

The term big data was coined by O’Reilly Media in 2005 [66]. Big data is an all-inclusive expression used to describe the volume of data generated by organizations in society. Big data is described in terms of volume (quantity), velocity (speed at which it is generated) and variety (mixture of variables) [67]. There are three primary sources of big data, namely social data (general media, social websites, and public domain), machine data (industrial equipment, sensors, smart devices, cameras, satellites, games, and cell phones) and transactional data (daily financial transactions). Essentially all information that can be measured and stored is classified as big data. Big Data, in terms of Educational Institutions is not limited to biographical information at registration; it includes all inputs, such as test and examination results, time taken to use online learning material, library use, online learning platforms and extracurricular activities. It can even be used to monitor the types of food purchased from food outlets on campus to evaluate and predict the nutritional status of users.

3.4.1 Statistics

The longitudinal study is conducted using inferential statistical analysis and techniques. The analysis is done using SPSS software version 27.0. All techniques and methods are part of that software. A few of those techniques are discussed for clarity and understanding.

Arithmetic Mean (Mean)

In statistics arithmetic mean (or simply mean) is the average value of a set of data. This would be the sum of the set of data divided by the number of data samples. In this study, the data applies to learners' results [68].

Standard Deviation

Standard deviation (SD) is a measure of the amount of variation or 'dispersion' in the data set, as seen in Figure 3.1. A low SD indicates that the values are close to the mean, while a high SD indicates that the data is spread over a wider range. Figure 3.1 shows a normal distribution plot or bell curve [68].

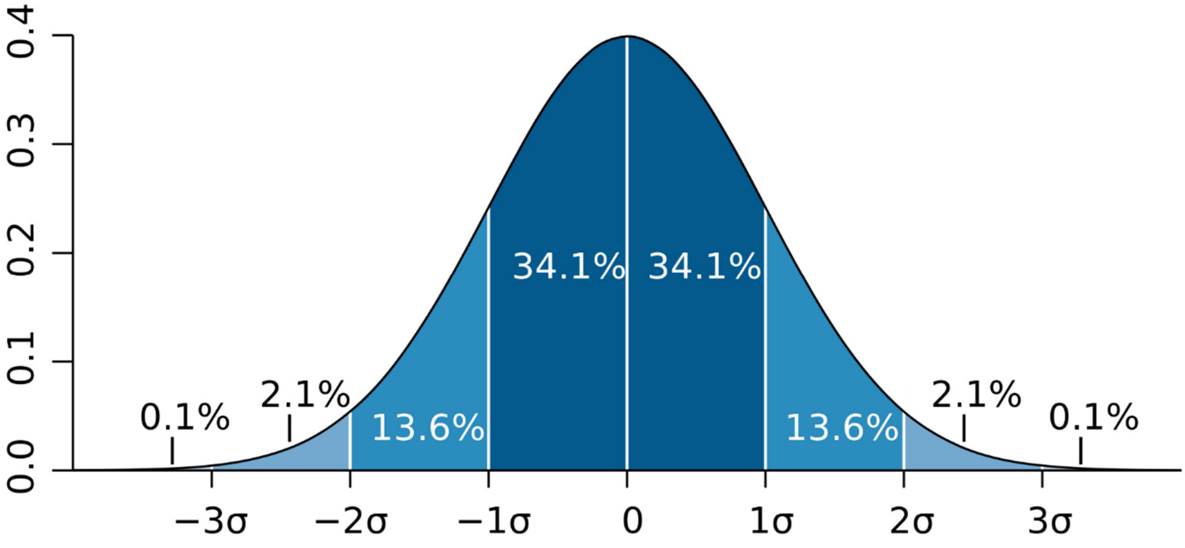


Figure 3.1: Normal Distribution Plot [68]

Pearson Correlation

Statistically Pearson Correlation Coefficient or bivariate correlation is a measure of linear correlation between two sets of data.

It is the ratio between the covariance of two variables and the product of their standard deviations. It is essentially a normalised measurement of the covariance, such that the result always has a value between −1 and 1. As with covariance itself, the measure can only reflect a linear correlation of variables and ignores many other types of relationship or correlation. Preferably Pearson correlation coefficient should be significantly greater than 0, but less than 1 (as 1 would represent an unrealistically perfect correlation).

Correlation Coefficient Interpretation • Exactly –1. A perfect negative linear relationship • –0.70. A strong negative linear relationship • –0.50. A moderate negative relationship • –0.30. A weak negative linear relationship • 0. No linear relationship • All negative relationships are inverse. • +0.30. A weak positive linear relationship • +0.50. A moderate positive relationship • +0.70. A strong positive linear relationship • Exactly +1. A perfect positive linear relationship • All positive relationships are directly proportional

Analysis of cross-classifications using Crosstabs

The crosstabulation table is the basic technique for examining the relationship between two categorical variables. The Crosstabs procedure tests the independence and measures of association and agreement for nominal and ordinal data.

Mann-Whitney Test

The Mann–Whitney U test is a nonparametric test of the null hypothesis that, for randomly selected values *X* and *Y* from two populations, the probability of *X* being greater than *Y* is equal to the probability of *Y* being greater than *X*. The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed [68].

3.4.2 Educational Data Mining

Data mining refers to the finding of patterns and structures in big data sets [69]. Pattern detection and model building are two aspects of data mining. Pattern detection groups similar data together, while modeling requires the building of mathematical simulations. Statistical modeling and data modeling are very similar. Size of data sets, however, creates additional problems for statistical modeling. The size of big data is irrelevant to pattern detection [70]. Due to the data being filtered and reduced to smaller sections. Although statistics have been used to examine some cases of data mining most pattern detection is done using computer algorithms [69]. Educational data mining is therefore the application of data mining using academic data to influence decisions for both the educational institution and learners. This data is not usually only vast, but varied and normal systems battle to find the relationships in the available data.

Educational data, like all forms of data, are rapidly expanding due to the nature of the digital world. There is a demand for current, credible measure of education quality [71]. Learners on an education campus today are consistently using ‘digital platforms’ [72] to communicate, transact and be educated. All data gathered during an individual’s time as a learner, can be analyzed for a better understanding of their behaviour. This can then be used to create the best possible learning environment for future learner cohorts. The relatively new field of educational data mining, attempts to determine the factors affecting learner retention, pass rates, and ultimately throughput rates and therefore graduate success [73]. In a study titled “Predicting Engineering Student Success Using Machine Learning”, Taodzera, et al [73] evaluated learner performance using five Education Data Mining (EDM) learning algorithms. They managed to achieve an accuracy of 65% in the study. This study clearly indicates that enormous amounts of educational, ‘big data’, has not only been stored for several years for record-keeping purposes, but it also now has the potential to be extracted and used to discovery of new knowledge [73].

One process of extracting and using stored big data is referred to as Knowledge Discovery in Databases (KDD) [74]. This process has nine discrete steps to extract knowledge from large data bases and solve a data mining problem. These steps, as shown in below Figure 3.2 [74], are Pre-Knowledge; Selection; Data cleaning and processing; Data reduction; Data Mining; Data Mining algorithms; Actual Data Mining; Interpretation and Consolidation. This system/process may be used for educational data mining, as a subset of Data Mining. The KDD process is further elaborated in Section 3.4. when working with the data obtained from the DUT MIS system to ensure completeness and accuracy [75]. The KDD techniques are used in this research to gain a better understanding of patterns and trends in higher education; more specifically it is used to predict the chance of success of TVET learners with passes of at least 50% in Engineering Mathematics, Engineering Science, Electrotechnics and Industrial Electronics, entering the Department of Electrical Power Engineering at DUT, via an articulation arrangement.

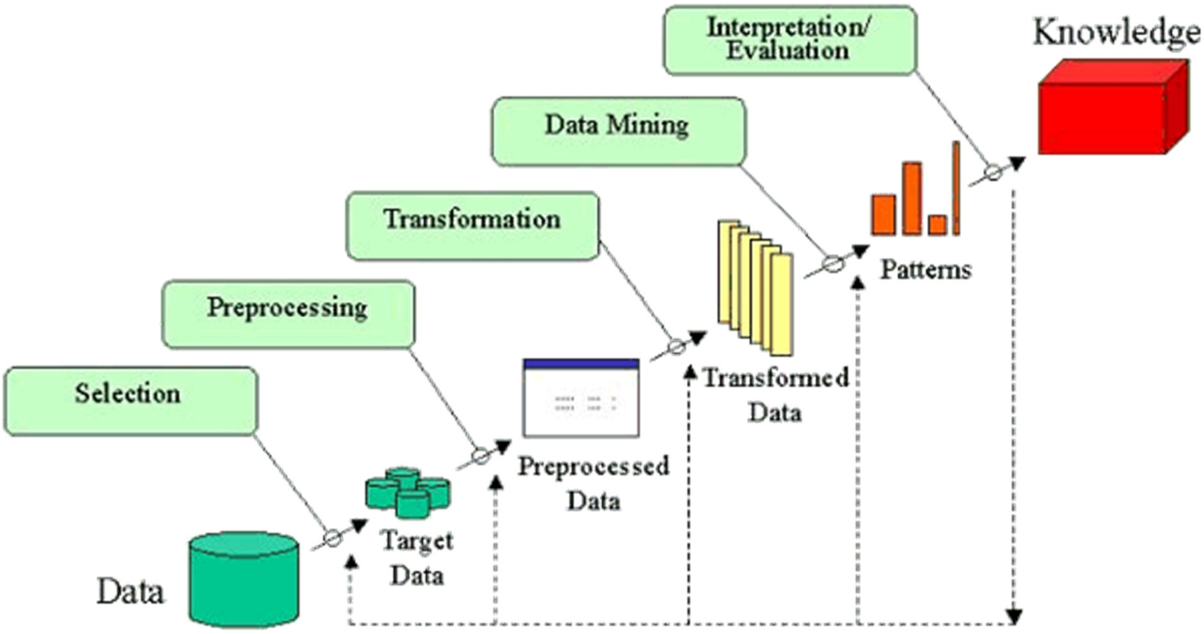


Figure 3.2: Knowledge Discovery in Databases Process (KDD) [74]

Shahari, et al [76], in the review of different data mining techniques namely; Artificial Neural; Networks (ANN), Decision Tree (DT), Support Vector Machine (SVM), and Naïve Bayes (NB) determined that Artificial Neural Networks had the highest prediction accuracy - that of 98%, when used to predict learner performance, as seen in Figure 3.3 [76]. It is referred to as an adaptable system as it can capture non-linear systems between dependent and independent variables. The review reported that ANN worked well on admission (biographical) data, students' attitude to self-learning and academic performance [76]. In the remainder of this chapter a review of a few of the relevant data mining techniques, their strengths and weaknesses will be presented to gain a better understanding of their application to Educational Data Mining.

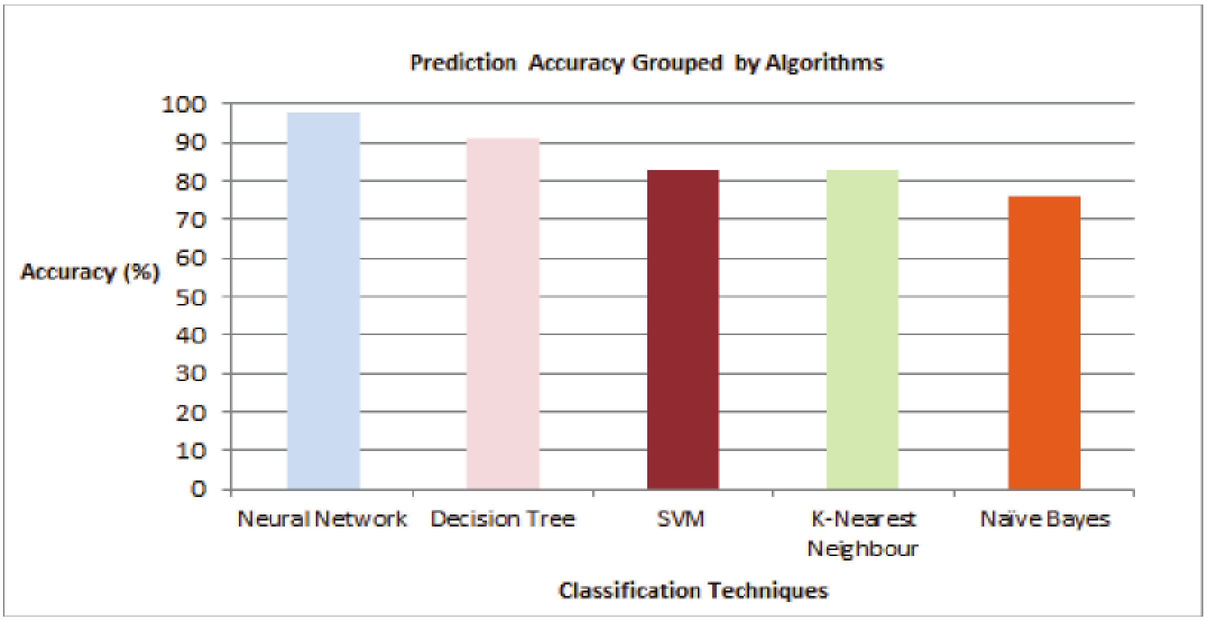


Figure 3.3: Classification Accuracy by Technique [76]

3.5 Data Mining Techniques

All data mining methods (classification methods) have distinct advantages and disadvantages, dependent on the data that is presented. Data mining is an experimental science. In this section we outline a few of the classification methods/algorithms used in data mining that may be used on Educational Data and that have been used previously by researchers in the field of educational data mining (EDM). Comment on the advantages and disadvantages of each are also briefly presented.

3.5.1 Linear Regression

Regression is a common statistical and data mining model, used primarily when the attributes are numeric. It models the relationships between one or more input attributes and an output variable. The ideal relationship is linear and continuous.

Linear regression has a straight-line relationship, where the unknown variable y and the known variable *x* are given by the equation:

$$y = w_0 + w_1 x_1 \qquad \text{(Equation 3.1)}$$

where w_0 the y-intercept and w_1 the slope, are regression coefficients. Generally, the data will result in the plotted points not all falling on the straight line. An error variable is added to the equation to account for the fault. The equation thus becomes:

$$y = w_0 + w_1 x_1 + \alpha \qquad \text{(Equation 3.2)}$$

where α is the error, a random variable that accounts for the inaccuracy of the model to generalize properly. Usually more than one variable is used, which gives the equation:

$$y = w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_k x_k + \alpha \qquad \text{(Equation 3.3)}$$

where y is the unknown variable to be predicted and $x_1, x_2, \ldots, x_k$ are different attributes of the dataset and $w_1, w_2, \ldots, w_k$ are the weights. This is referred to as multilinear regression.

The summed equation:

$$y_1 = w_0x_0 + w_1x_{11} + w_2x_{21} + \cdots + w_kx_{k1} = \sum_{j=0}^{n} w_j\, x_{j1}$$

(Equation 3.4)

where $w_0, w_1 \dots . w_k$ are the weights, y_1 the first example of the class and attributes x_{01}, x_{11} $\dots .\ x_{k1}$, with the second subscript denoting that it is the first example, and an extra attribute x_0 which is always 1, which is the predicted value for the first example.

Linear regression is implemented using the least squares method. The method of least squares is used to estimate these parameters, y_i. That is, the estimates are such that the sum of the squares of the differences between the observed and estimated class values are minimised by minimising the following sum:

$$\sum_{j=1}^{n} (\, x_i - w_j a_{ji})^2$$

(Equation 3.5)

Since it is simple and easy to use, it is implemented in other prediction models. It is more reliable when all models are linear and suffers from the drawback that regression is sensitive to data outliers [73, 77].

3.5.2 Logistic Regression

Logistic regression (LR) [78] is a similar machine learning algorithm to linear regression. It is also a predictive analysis algorithm based on the concept of probability. The logistic regression model is based on the ‘Sigmoid Function’ or ‘logistic function’ instead of the linear function used in linear regression. The logistic regression hypothesis tends to limit the cost function between 0 and 1. The cost function is expressed as a difference between the predicted value and the actual value. Therefore linear functions fail to represent values outside these limits. The logistic regression hypothesis expectation is given by.

$$0 \leq h_{\Theta}(x) \leq 1 \qquad \text{(Equation 3.6)}$$

The Sigmoid function, seen in Figure 3.4 [78] is used to predict probabilities.

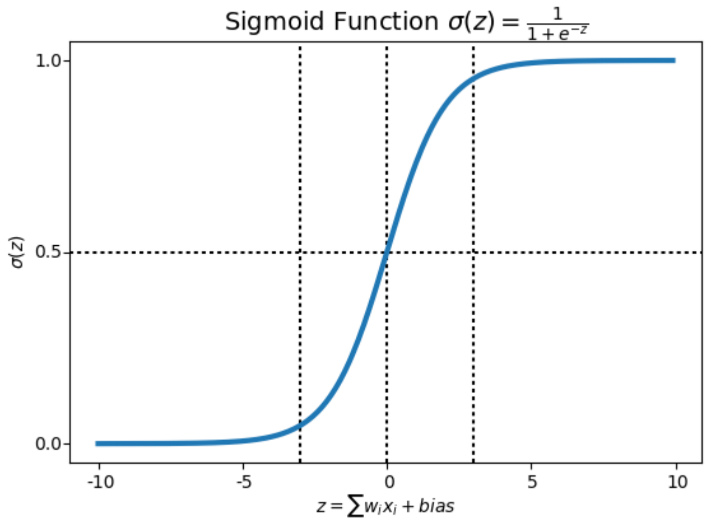


Figure 3.4: Sigmoid Function Graph [78]

The formula of the Sigmoid function is given by the equation,

$$f(x) = \frac{1}{1 + e^{-(x)}} \qquad \text{(Equation 3.7)}$$

Logistic regression therefore returns a set of outputs above or below the threshold value based on the probability when passed through the prediction function and returns a value between 0 and 1.

Logistic regression is easy to implement, simple to understand and widely accepted as a scientific tool. It is not as effective with small amounts of data, especially if there are missing variables or the data is complex. The model is also susceptible to overfitting. Overfitting can be described as when the algorithm fits too closely to the training data and cannot generalise well on new data once implemented.

3.5.3 Naïve Bayes

Bayes' theorem is named after Thomas Bayes [79] , a nonconformist English clergyman who pioneered probability theory during the 18th century. This classifier naively assumes that all attributes are equally important *a priori* and statistically independent, that is knowing the value of one attribute says nothing about the value of another (if the class is known). Bayes' rule states that the probability of event H given evidence E,

$$Pr[H|E] = \frac{Pr[E|H]\,Pr[H]}{Pr[E]} \qquad \text{(Equation 3.8)}$$

- *Pr* $[H]$ is the priori probability of H – probability before evidence is seen.
- *Pr* $[H \mid E]$ is the posteriori probability of H – probability of event after evidence is seen.

The assumption is naïve in that the evidence splits into parts that are independent, as shown by the following equation:

$$\Pr[H|E] = \frac{\Pr[E_1|H]\ \Pr[E_2|H] \ldots \Pr[E_n|H]\ \Pr[H]}{\Pr[E]} \qquad \text{(Equation 3.9)}$$

Naïve Bayes is preferred as a classifier as it is easy to implement, produces good results and is computationally efficient. It handles missing or unknown data and is implementable on small data sets. Similar or redundant attributes will however cause problems for the classifier.

3.5.4 Decision Trees

Decision Trees (DT) are supervised machine learning algorithms that progressively divide a set of data into smaller groups in a top-down approach until each group can be classified by a single attribute. The process resembles how a tree splits from the tree trunk into each branch getting smaller until finally reaching a leaf, as seen in Figure 3.5 [80]. This classifier starts at the root node and splits, finding the optimal solution for the classification of the data presented.

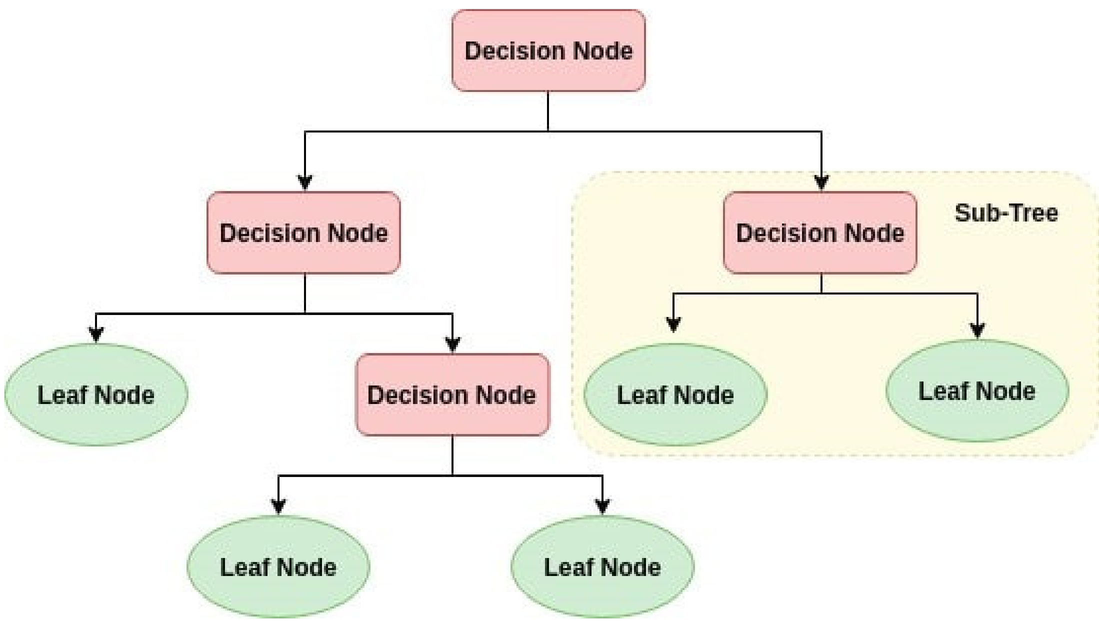


Figure 3.5: Decision Tree Illustration [80]

The Decision Tree (DT) uses a system of information gain to determine priorities at each node. Entropy is a measure of the conformity of the data at each node. Information gain is the decrease in entropy as seen in Figure 3.6 [80]. Information gain calculates the difference between entropy before the node is split and the average entropy after. The entropy will be zero if all the data at the node are alike and one if the data is split equally.

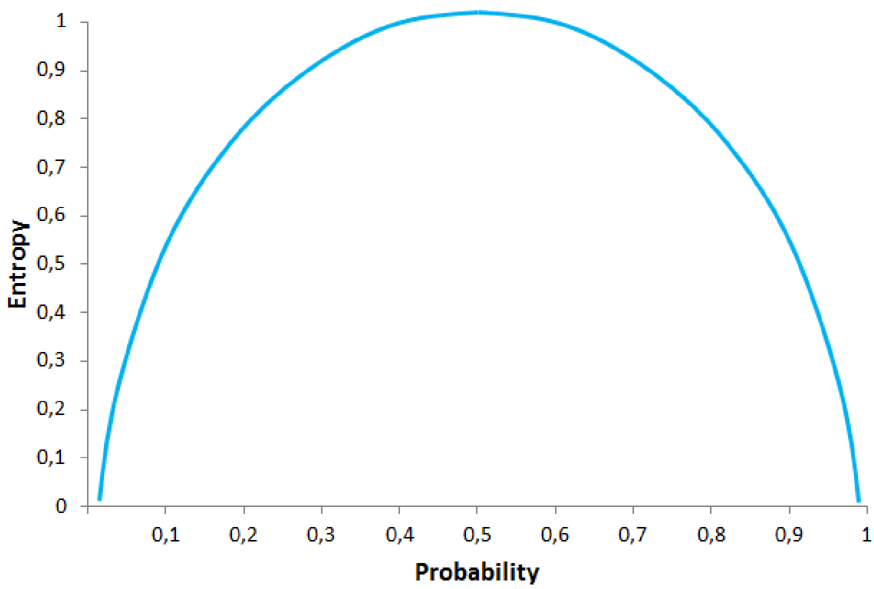


Figure 3.6: Entropy Outline [80]

The equation for the curve above is given by:

$$Entropy = -0.5log_2 0.5 - 0.5log_2 0.5 \qquad \text{(Equation 3.10)}$$

The curve above illustrates a decision at a node between two choices. If the amount of data for each set is equal the probability will be 0.5 and the entropy 1, and as one set increases in number the entropy will decrease.

The C4.5 algorithm, an example of a DT algorithm, designed by Ross Quinlan in 1993, based on his previous ID3 algorithm uses information gain to select the best attribute at each node using a single pass algorithm derived from binomial confidence limits [76]. It deleted the limitation that features must be categorical by dynamically defining a discrete attribute (based on numerical variables) that partitions the continuous attribute value into a discrete set of intervals. C4.5 converts the out of the ID3 algorithm into sets of if-then rules. The accuracy of each defining rule is then evaluated to resolve the order in which they should be applied. Pruning, the removal of a node, is done by removing a rule's precondition if the accuracy improves without it.

Decision Trees are easy to implement, interpret and visualize. They require less pre-processing by the user and handle missing data. The decision tree has no assumptions about the distribution because of the non-parametric nature of the algorithm [81]. They are sensitive to noisy data, and therefore are suspectable to overfitting. Noisy data can be defined as information that is meaningless, corrupt, or basically not able to be understood by computers. This information might be due to spelling mistakes, abbreviations or out of text. Overfitting can be described as when the algorithm fits too closely to the training data and cannot generalise well on new data once implemented. Also, DTs are sensitive to small changes in data, which might result in a change in the whole tree.

3.5.5 Support Vector Machine

The Support Vector Machine Algorithm (SVM) is used as a regression or classification classifier [82]. This algorithm performs classifications by finding the hyperplane which maximises the distance between the two attributes, where the vectors defining the hyperplane are the support vectors, as seen in Figure 3.7 [82].

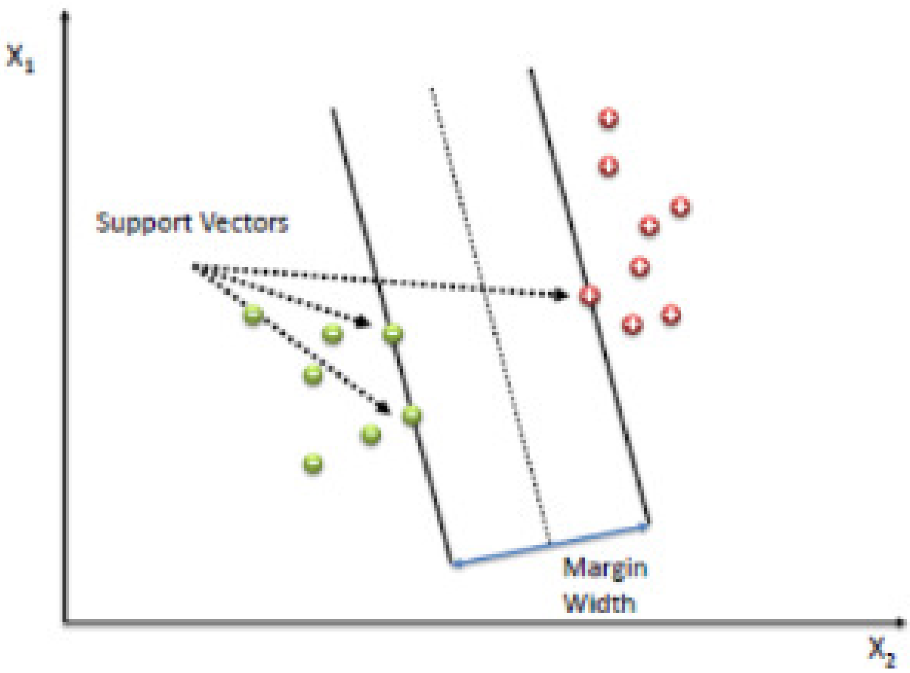


Figure 3.7: SVM Hyperplane [82]

The optimal hyperplane is defined where the margin is maximised, and any misclassifications introduce a penalty term. To define the optimal hyperplane, the width of the margin *w* and where *b* is an offset and *b/||w||* is the distance from the distance from the hyperplane to the origin, as seen in Figure 3.8 needs to be maximised [82]. The solution for *w* and *b* is found by solving the objective function using Quadratic Programming.

$$min\frac{1}{2}\|w\|^2 \qquad \text{(Equation 3.11)}$$

$$\textit{such that } y_i(w.x_i + b) \geq 1, \forall x_i$$

The ideal SVM analysis produces a hyperplane completely separating the vectors into two non-over-lapping attributes, however this might not always be possible. In this case SVM determines the hyperplane, maximising the margin and minimising the misclassifications.

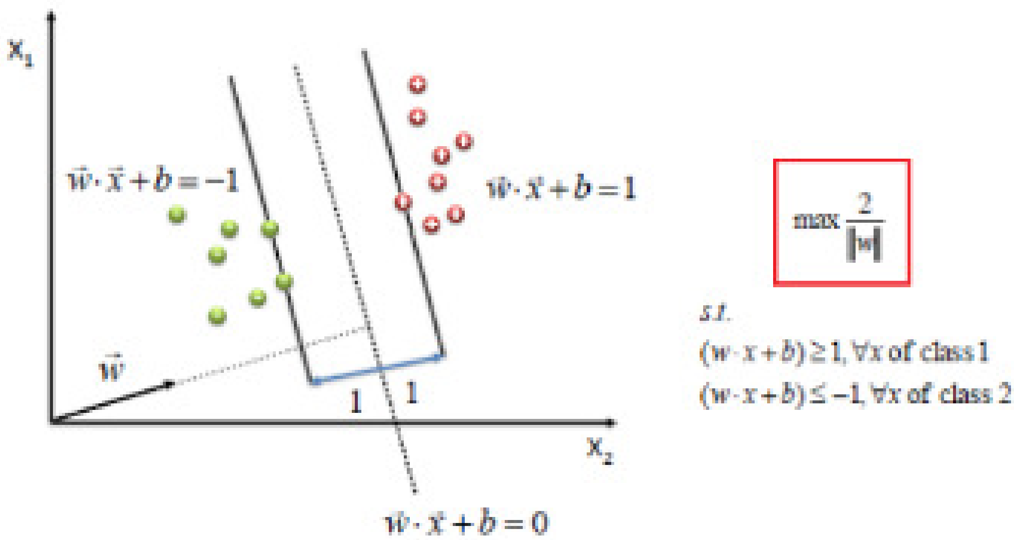


Figure 3.8: SVM Analysis [82]

The SVM classification algorithm has high prediction accuracy in classification tasks and can handle complex functions, while being memory efficient. It is less reliable with training time on large datasets and noisy input data.

3.5.6 Artificial Neural Networks

The historical account of computing is fascinating and requires some unpacking [83]. The goal of Babbage's original difference machine was to calculate error free mathematical tables. The evolution of computing machines and the associated technology in the 19^{th} and 20^{th} century can only be described as exponential. This development was driven by business / industry and World War 2. One early computer, the ENIAC (Electronic Numerical Integrator and Computer) was used to calculate ballistic trajectories. What is known as common problems today, this computer had too little storage, too many tubes (computer power) and took too long to re-programme. Several years later in 1952, the Remmington UNIVAC computer was able to use statistical analysis to predict Eisenhower's landslide victory in the US presidential elections. As computer speeds, storage ability and size continued to increase so too did the applications. The increase in technical ability of computers will allow the development of neural network in the future.

As described above, computers can now store vast amounts of information and can perform repetitive tasks extremely quickly. They can, however, only perform as instructed and execute commands in a serial or linear order. Artificial Neural Networks (ANN) are designed to behave like the interconnected network structure of nerves found in the human brain – albeit through the use of mathematical algorithms [84]. These algorithms consist of an array of interconnected nodes, as does the human brain, that exchange information. As a result, these nodes learn, in a manner similar to the development of the human brain. It is particularly useful in pattern matching and learning, such as visual image classification, speech recognition and medical diagnosis. This makes ANN a useful tool when trying to solve problems that are difficult or impossible to solve using traditional computational and statistical methods [85].

The interest in the research of neural networks has grown in the last few years [86]. The investigation of brain-style computation or parallel distributed-processing systems can be compared to that of the traditional serial computer. The first machine called Perceptron [87], was built by Frank Rosenblatt, an American psychologist in 1958, in an attempt to sense, recognize, remember and respond like the human mind. Developed upon previous work by

Warren McCulloch and Walter Pitts in the 1940's. Perceptron behaved like a human brain neuron that receives signals from dendrites or nerve cells. As seen in the Figure 3.9 [87], in a human neuron the Cell Body processes all incoming stimuli from the human interaction and triggers signals to other neurons via the axon. A brain would consist of tens of billions of such cells and tens of trillions of such connections, thus having an estimated storage capacity of 10^{14} interconnects and a processing speed of 10^{16} interconnects per second [85].

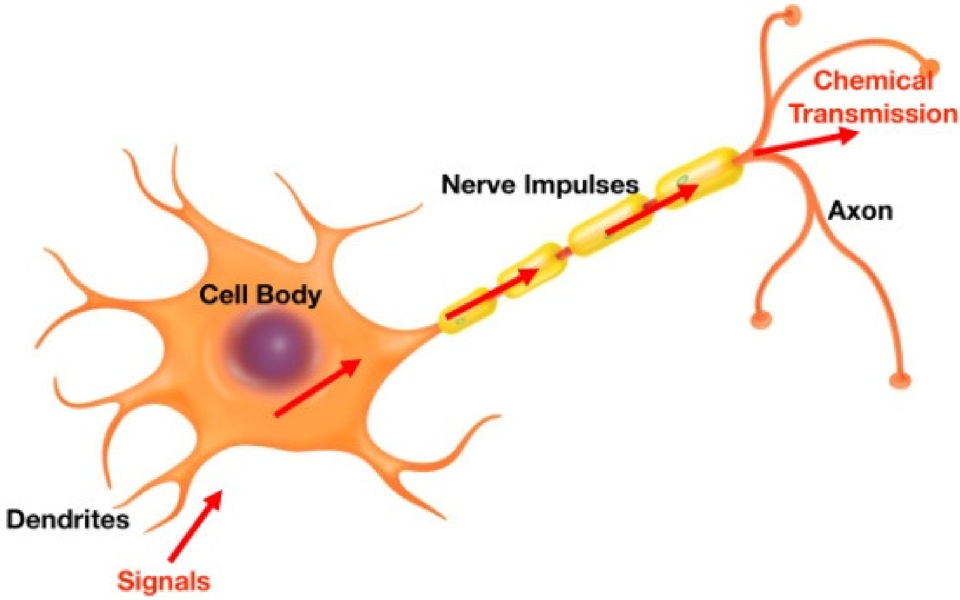


Figure 3.9: The Brain Cell [87]

This structure of the artificial neural network or neuron is now depicted mathematically in Figure 3.10 [88]. This neuron has several data inputs (blue circles) or possible axon from previous neurons. Each data input has a specific weighting (importance - could be negative or positive). Within the neuron these weights will be summed and have a predetermined bias added (green circle). That result will then have a mathematical step function (purple rectangle) applied - resulting in an output. The structure of this typical processing unit is an ANN, also called an artificial neuron or perceptron [89].

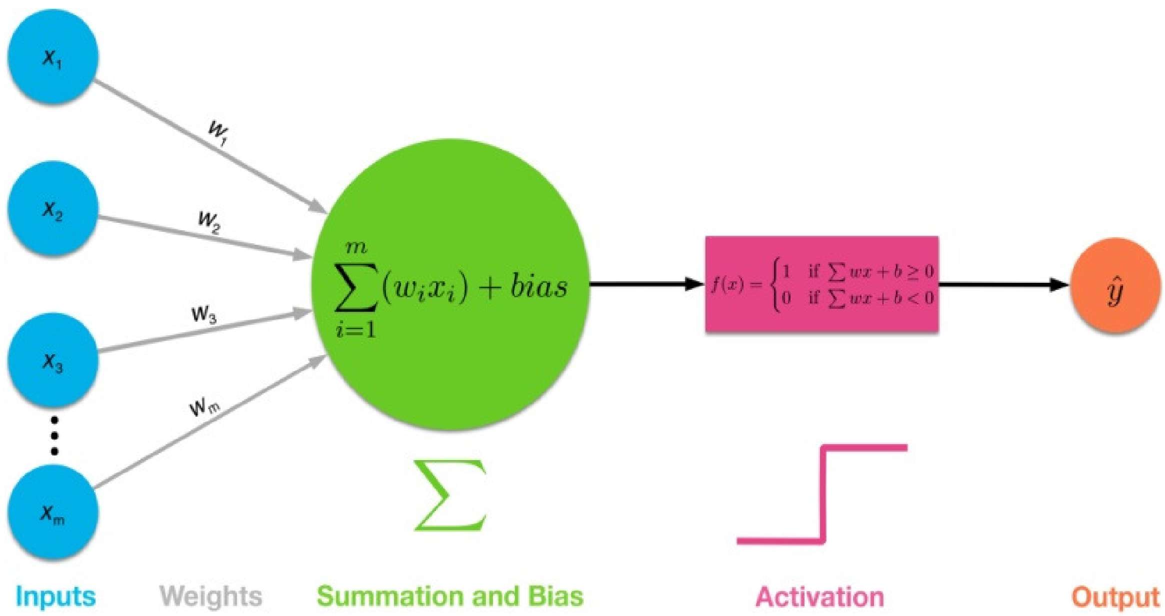


Figure 3.10: Mathematical Representation of a Neuron [88]

These simple single neurons can now be arranged into an artificial neural network (ANN), as multi-layer devices. By selecting differential or non-linear functions with a back propagation learning algorithm training can be achieved. In other words, by grouping the individual neurons into a parallel system with feedback between the output and input, the network may be taught. The first layer is the input layer, and the second layer is usually called the hidden layer and there can be more than one depending on the complexity of the ANN, as depicted in Figure 3.11 [86].

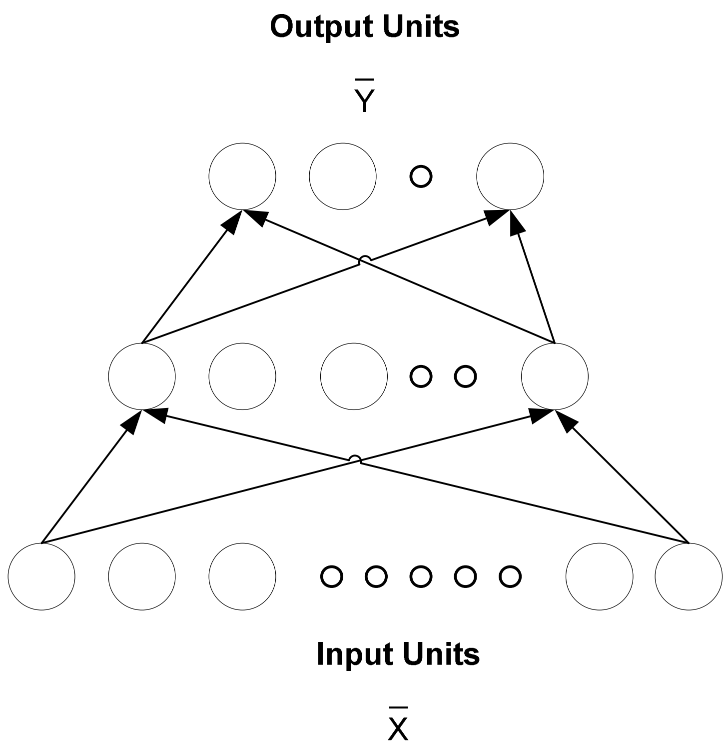


Figure 3.11: Three Layer Network [86]

The multi-layer image expands on the simple ANN, connecting each input neuron to every second layer neuron and every second layer neuron is connected to every output layer neuron. Each connection has a corresponding weight (w). The weights from each input are summed and manipulated by the threshold function before becoming an output. The values of the weights (w_1) are obtained by the process of training the network. In supervised learning, the approach adopted in this study, the input layer is presented with inputs and targets. Optimisation of the neural network involves making choices on the size and structure of the overall neural network [89]. Back propagation has become the single most popular method to train artificial neural networks. Back propagation is a mathematical algorithm used to calculate the error between the target and output and by adjusting various inputs, bringing the outputs closer to the target, thereby training the network.

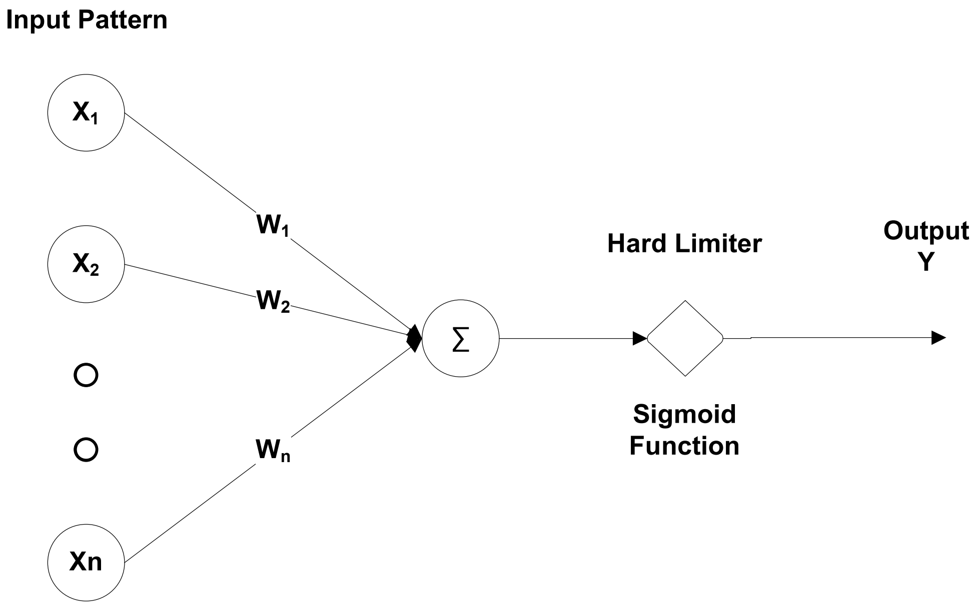


Figure 3.12: Artificial Neuron Representation [88]

The basic perceptron shown in Figure 3.12 [88] is defined by the following equation:

$$Y = f_h\left(\sum_{i=0}^{n} w_i x_i\right) \quad \text{(Equation 3.12)}$$

Early neural networks used an activation function *(f_h)* that was hard limiting, which gave an output of +1 whenever $\sum w_i x_i$ was greater than the threshold value. Learning became, teaching the network the value of w_i until the neuron produced the output Y that was required.

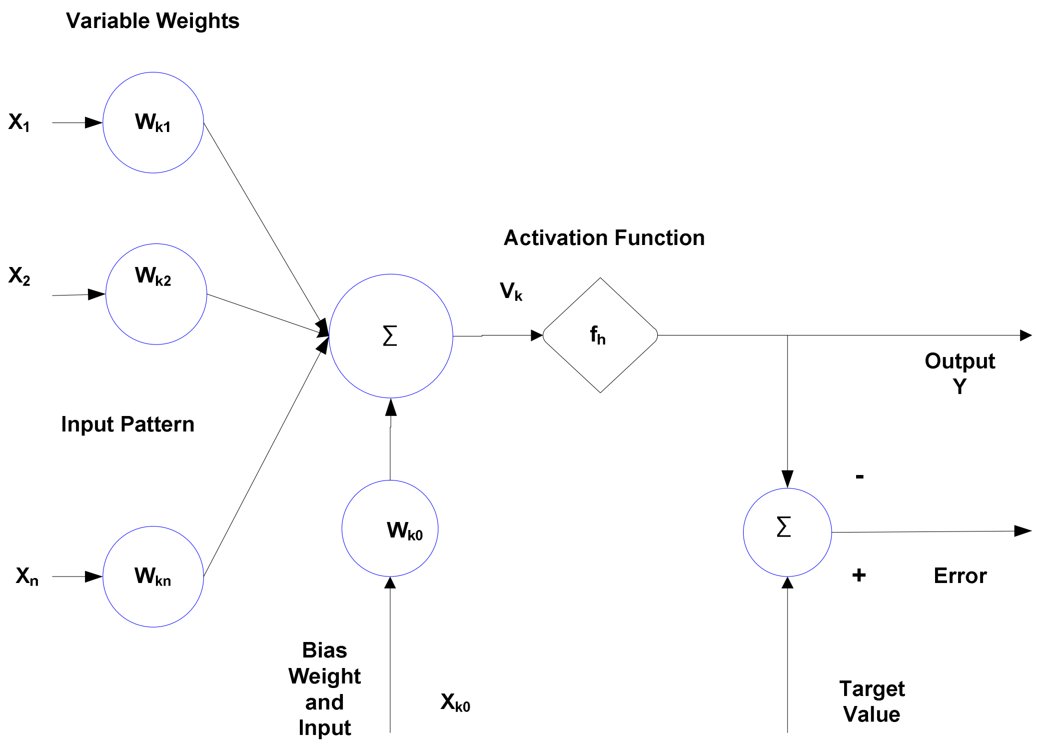


Figure 3.13: Nonlinear Model of a Neuron [90]

An updated nonlinear model of a neuron is shown in Figure 3.13 [90]. The added value is an externally applied bias w_0. This bias increases or decreases the input to the activation function [89, 91]. The equation representing the N_{th} neuron now becomes:

$$Y = f_h\left(\sum_{i=0}^{n} w_i x_i + w_0\right) \qquad \text{(Equation 3.13)}$$

Where x_0 = +1 or -1, and here w_0 is the bias weight.

The early activation functions were step or threshold functions, of the form:

$$f(h) = \begin{cases} 1, & v \geq 0 \\ 0, & v \leq 0 \end{cases} \qquad \text{(Equation 3.14)}$$

This was known as the McCulloch-Pitts model [90, 92]. The activation function can also be a piecewise linear function,

$$f(h) = \begin{cases} 1, & v \geq 0.5 \\ v + 0.5, & -0.5 < v \leq 0.5 \\ 0, & v \leq -0.5 \end{cases} \qquad \text{(Equation 3.15)}$$

However, modern multi-layer neural networks make use of a sigmoid activation function,

$$f(h) = \frac{1}{1 + exp(-av)} \qquad \text{(Equation 3.16)}$$

Or the hyperbolic functions [92],

$$f(h) = tanh\left(\frac{v}{2}\right) = \frac{1 - exp(-v)}{1 + exp(-v)} \qquad \text{(Equation 3.17)}$$

The sigmoid function, Figure 3.14, being asymptotic can limit the output of the neuron. However, what makes it the most used is the fact that the function is differentiable which is important for the minimisation of the squared error criteria as the basis for multi-layer learning algorithms [91].

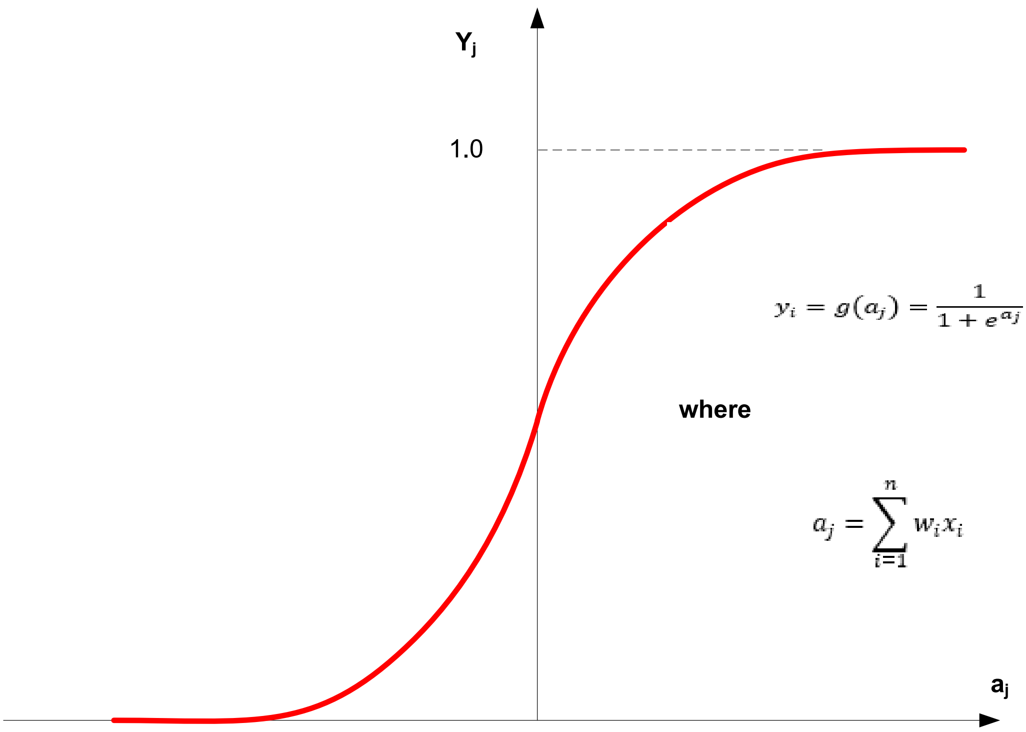


Figure 3.14: The Sigmoid Function [88]

With reference to Figure 3.14 [88] the error function can be defined as the difference between the neuron output and the target value. It is more convenient to square the error to eliminate any negative values. Therefore, the sum-of-squares error function can be used as follows:

$$E = \frac{1}{2}(y^p - t^p)^2 \qquad \text{(Equation 3.18)}$$

Where y^p is the multi-layer network output for an input p and t^p is the target value [89]. The error function E is minimised by differentiation with respect to every weight w_i in the network.

An individual weight is given by:

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i} \qquad \text{(Equation 3.19)}$$

To achieve this differentiation, a continuous differentiable function such as a sigmoid function shown in Figure 3.14 must be used; specifically:

$$y_i = g(a_j) = \frac{1}{1 + e^{a_j}} \qquad \text{(Equation 3.20)}$$

where:

$$a_j = \sum_{i=1}^{n} w_i x_i \qquad \text{(Equation 3.21)}$$

A two-layer perceptron, as depicted in Figure 3.15 [88] is now considered.

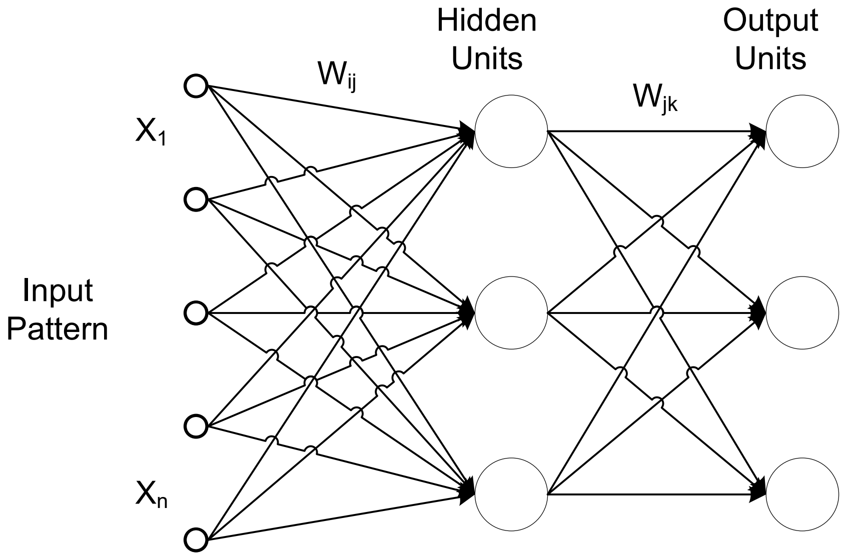


Figure 3.15: Two-Layer Perceptron [88]

The two-layer perceptron indicates that it has two layers of weights. Additionally, I = the number of inputs (five in this figure), J = the number of hidden units (three in this figure) and K = number of outputs (three in this figure). For this K class problem, the error function now becomes:

$$E = \frac{1}{2}\sum_{p=1}^{p}\sum_{k=1}^{k}(y_k^p - t_k^p)^2 = \frac{1}{2}\sum_{p=1}^{p}\sum_{k=1}^{k}(g\sum_{j} w_{jk}y_j^p - t_k^p)^2 \qquad \text{(Equation 3.22)}$$

which expands to:

$$E = \frac{1}{2}\sum_{p}\sum_{k}(g\sum_{j} w_{jk} g(\sum_{i} w_{ij} x_i^p) - t_k^p)^2 \qquad \text{(Equation 3.23)}$$

The subject E can now be differentiated with respect to each weight to calculate the gradient of the error function. The back propagation error algorithm thus arises and is used to update the input weights in the multi-layer network.

The update equations for the input-to-hidden layer is thus derived [89] as:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = -\eta \delta_j y_i \qquad \text{(Equation 3.24)}$$

where:

$$\delta_j = \frac{\partial E}{\partial_{aj}} = \sum_{k} \delta_k w_{jk} y_j (1 - y_j) \qquad \text{(Equation 3.25)}$$

and for the hidden-to-output layer:

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \delta_k y_j \qquad \text{(Equation 3.26)}$$

where:

$$\delta_k = \frac{\partial E}{\delta a_k} = (y_k - t_k) y_k (1 - y_k) \qquad \text{(Equation 3.27)}$$

The back propagation error algorithm derives its name from the fact that the errors δ_k of the output layer need to be calculated first, as they are required in the calculation on the input layer updates.

An ANN is “trained” in three distinct steps, each requiring its own data, as cross contamination of data will affect the results of the network. Step one is where a set of data is presented to the network and the network is trained. Step two is a validation step to determine the error. This is done with new data. The final step, step 3, is the test step where a last set of data is presented to the network.

Artificial Neural Networks do require considerable computational power to train a network, which might mean that the process is time consuming. They require good quality data, although they accept different types of data. There is no requirement for the relationship between the input and output to be known.

3.6 Previous Educational Data Mining Research

Although Educational Data Mining Research in South Africa has been falling behind other regions such as Western Europe, Australia, North America and New Zealand, there has recently been a steady increase in the priority for high quality research into the technologies, methods and models in data analytics in education, teaching and learning in Africa [93]. In a study at a Nigerian University, in 2019, in which learners' fifth-year grade point average was used to determine the impact of engineering learners' performance in the first three years of their graduation result. It was found that a maximum accuracy of 89.15% was achieved using Logistic Regression [76]. In 2020 three learners from the University of the Witwatersrand, used six classifiers, namely Naïve Bayes (NB), Decision Trees (DT), Random Forests (RF), Linear Regression (LR), Support Vector Machines (SVM) and Logistic Model Tress (LMT) to determine learner success at a HEI. They managed to obtain very similar results with 83% accuracy, using all six classifiers. Several similar studies are summarised in Table 3.1.

Table 3.1: Previous Educational Data Mining Research

Researcher	Year	Classifier	Accuracy
Taodzera [73]	2018	DT	66%
Shahari, Husain, Rashid [76]	2019	LR	89%
Ndou, Ajoodha, Jadhav [94]	2020	NB, DT, RF, LR, SVM, LMT	83%
Mashiloane [95]	2015	NB	47%
Fynn, Adamiak [96]	2018	DT	76 – 90%
Masethe, Masethe [97]	2014	NB	82%
Vambe, Sibanda [98]	2017	DT	75%
Masangu, Jadhav, Ajoodha [99]	2021	SVM	71%

Educational Data Mining is an experimental science and continued work will produce further results.

The future of the South African economy is dependent in a large part on the education sector focusing on the quality and relevance of education and skills [100]. Educators need to use technology, where possible, to improve the academic ambit. For example, since poor throughput rates is one of the historical problems plaguing universities [19] and artificial neural networks (ANN) have achieved an accuracy of 90% [101] in studies, when selecting successful learners for university entrance, ANN could be a useful tool to employ, from as early as selection, through to the facilitation of pedagogical and instructional interventions.

As computing storage space technology has increased, the ability to store big data has become a trend. This big data collection is increasing in velocity, variety, and volume. Data mining is the processing of this data to determine hidden insights and improving knowledge. Educational data mining, being a subset of data mining, may use all data from the education sector data bases, namely, MIS, NLRD and HEMIS. At DUT all learners' information is retained on the management information system (MIS). Using artificial neural networks or similar artificial intelligence computing systems can analyse this data or use data mining techniques to gain insight into useful knowledge. These systems, due to the size and complexity of the data, can draw conclusions that may be outside of the normal scope or ability of humans. Neural networks, use the application of artificial intelligence techniques or deep learning, to make decisions or predictions about the future - using information from a similar situation from the past [102].

3.7 Knowledge Discovery Process (KDD)

As mentioned previously the Knowledge Discovery Process (KDD) is one method that can be used in data mining to discover new knowledge. As part of this study, the KDD process is followed where possible, to ensure compliance, transparency and ensure reputable results. The process can be a reiterative one using the steps outlined in the sub-paragraphs below.

3.7.1 Pre-Knowledge

Pre-knowledge is essentially understanding what is required from the input data, understanding the situation, and having the relevant knowledge of the researched topic. This might be understood as having situational awareness.

All learners during the period that spans this investigation would have been preselected and possibly admitted by the writer. Most of the credits/exemptions would have been administered by the writer. The writer was involved in the selection of learners and therefore has a clear understanding of how learners were selected into the Department of Electrical Power Engineering. The years 2013 – 2017, were selected for this research, as they were the most recent before the National Diploma was phased out and therefore the most recent data available.

3.7.2 Selection

This step involves selecting the data set and understanding what discovery is required. No discovery/investigation should take place without having an idea of what is being investigated.

The data analysed in this research is readily available data from DUT's management information system (MIS). This is extracted, analysed, and used to develop predictive models to determine the success/failure of a cohort of learners articulating from TVET colleges into the National Diploma: Electrical Engineering at DUT from January 2013 until July 2017. This model should enable the prediction of future learner success before entering DUT, with possible generalisation to other HEI's.

Data was analyzed from the academic records obtained from the DUT MIS system. In total there are 10 semesters of data, for the years 2013 through to 2017. The learner enrollment data and the results for first level subjects were captured for analysis. All data is extracted manually from a simple text document. Random checks are carried out to check for accuracy. All data is displayed in an MS Office Excel Spreadsheet, as shown in Annexure 1 and Annexure 2. Editing and normalisation is only carried out in the data and processing section. All personal information, except the student number is deleted. The student number is retained in case additional information is required at a later point; however not required for any analytical purposes. All information that can be used for identification purposes is deleted.

3.7.3 Data Cleaning and Processing

This step is required as no data or information will be exactly in the format that is require for every use. Some manipulation will need to take place. This includes handling missing information and noise.

It is assumed that the audited learner data in the DUT MIS is accurate. This assumption is reasonable as the data was utilised frequently in learner data analysis for generating other official results and audited reports at DUT. Any errors and corrections are assumed to be minor and will not have any significant effect on the findings of this research. All attempts were made to locate missing data, by browsing learner information in learner files or by contacting the Central Applications Office (CAO). Any missing data is therefore assumed to be missing completely at random. Any errors are purely unintentional, and the utmost care is taken during capturing, manipulating, and analysing the data.

All data file cleaning is carried out in a systematic method, to clearly identify any changes that are made. The following steps constitute the process:

1. Any exemptions given for a 1st year subject that were awarded, is given a mark of 50%. These exemptions are indicated in green on the spreadsheets Annexure 1 and 2.
2. All National Senior Certificate (NSC) grades or percentages are converted to a value between 8 and 1, according to the Table 3.2 below. These values are indicated in brown on Annexure 2.
3. All TVET grades or percentages are converted to a value between 1 and 8, as indicated in Table 3.2. The Durban University General Handbook [9] , Rule G7(2)(c), specifies that the TVET N3 certificate is equivalent to a standard grade (SG) pass and therefore TVET N4 is equivalent of a higher grade (HG) pass and as such the HG equivalence is used. These values are indicated in brown on Annexure 1.

Table 3.2: Symbol / Numeric Value Conversion [8]

Symbol	A	B	C	D	E	F
HG	8	7	6	5	4	3
SG	6	5	4	3	2	1
%	80 - 100	70 - 79	60 - 69	50 - 59	40 - 49	30 - 39

A total of 154 TVET N4 learners that entered the Department of Electrical Power Engineering were captured. More than 50%, a total of 87 learners' information was excluded, due to missing or incomplete data. These learners had been captured on the MIS system as PASSED and no symbol was attached to the subject. Every attempt was made to obtain these marks from MIS and CAO, unfortunately the CAO records are deleted after one year and therefore unobtainable. This drastically reduced the number of learners in the dataset to 67. A total of 646 SC/NSC learners were captured from the MIS system. Only 100 learners' information was excluded, due to missing information, reducing the dataset to 546,. All learners are shown in Annexure 1 and 2, with the excluded leaners appearing at the bottom of the spreadsheet.

3.7.4 Data Reduction

In this stage data might need to be reduced or some unneeded data deleted. This requires the transforming of the data.

The Data is transformed in the different iterations depending on what analysis is being done. All subjects other than English, Mathematics and Physical Science were removed from the SC/NSC data. This is done because the Department of Electrical Power Engineering focuses on these subjects when selecting learners. They are also the only subjects that are consistently done by all learners entering engineering programmes at DUT. The TVET results, i.e. N4 Engineering Science, N4 Mathematics, N4 Electrotechnics and N4 Industrial Electronics are combined with the SC/NSC English as all TVET leaners are still required to have passed English. This is usually done at the SC/NSC level. One such transformation of the data is the combining of the English A (First Language) and English B (Second Language) to obtain a combined English mark.

3.7.5 Data Mining Type

Waikato University has embraced machine-learning (ML) with the development of Waikato Environment for Knowledge Analysis (WEKA). This software incorporates several ML techniques into a workbench that can derive useful knowledge from databases too large or complex to analyse by hand. It contains tools for data preparation, classification, regression, clustering, association rules, mining, and visualisation [103].

The writer attended a 5-week online training course, entitled “Data Mining with WEKA” offered by Professor Ian Witten from Waikato University, to obtain a better understanding of data mining and the use of WEKA.

A decision on whether the goal of the KDD process is classification, regression or clustering is required. These are three types of data mining, used depending on the outcomes required. The Waikato Environment for Knowledge Analysis (WEKA) classification tool, Version 3.8.5, seen in Annexure 3 [104, 105], is used as a data mining tool. Several classification and filtering tools within WEKA is used for this study. The various classification methods or data mining algorithms have been previously discussed and the results are be shown in Chapter 4.

3.7.6 Data Mining Algorithms

Each data algorithm is best suited for different applications. Examples of these algorithms are Linear Regression (LR); Naïve Bayes (NB); Artificial Neural Networks (ANN); Support Vector Machines (SVM) and Decision Tree (DT). These various data mining techniques have been previously discussed. Data mining determines correlation but does not imply causation. The results of the various data mining results will be presented in Chapter 4.

3.7.7 Actual Data Mining

The actual data mining techniques to find patterns using one or more algorithms are employed at this stage. The data is exported from Excel in a CSV file which WEKA is able to import [104]. Various classification and filtering methods are employed. These methods are tabulated in the results section. The following points are considered when utilising WEKA:

- Be skeptical of results as it is easy to misinterpret the results.
- Be careful of overfitting the algorithm.
- Pitfalls and Pratfall's – minimising mistakes both due to the WEKA / Data and human error.

The WEKA tool is used for two sections of the analysis:

1. Input Attribute Correlation Evaluation – to determine if any single input attribute has more of an effect on the output variable/class.
2. Data Mining/Classification – Using previously listed data mining tools or others to determine if any data mining tool can be used to predict the success of learners entering the Department of Electrical Power Engineering.

3.7.8 Interpretation

Once the data mining algorithms and data mining techniques is utilised, the results are scrutinised. This is presented in Chapter 4 and Chapter 5 of this thesis.

3.7.9 Consolidation

As the new knowledge might not be what one set out to find or even be immediately apparent, consolidating the mined knowledge and using the knowledge would be the final step in the process.

The final step of the KDD process would be the rolling out of the selected data mining process, using new data and the possibility of changing mindsets, for example changing perceptions on articulation and rules relating to the admission from TVET Colleges, to improve the future of the learners.

3.8 Rules for Progression

The success rate refers to the total number of courses passed by learners in a given academic year or semester relative to course enrolments. It is calculated by dividing the total number of FTE degree credits (courses completed) by FTE enrolments. These calculations, for a programme or for an institution, produce weighted average success rates. Apart from being a measure of the overall performance of learners, they are one of the bases upon which HEIs are funded.

One method of determining the success of the learners is by checking their progression into the second semester (S2) of study. In the first semester they are all exposed to the same curriculum and therefore can be compared. Only the first semester of the learners' studies is used as thereafter, too many variables may influence performance. The comparison is done according to the Departmental rules, as written in the Electrical Power Engineering Departmental Rule Book [8], using the HEMIS credit weighting per subject.

Table 3.3: Semester 1 Subjects / Subject Codes and FTE Credit Value

Subject	Projects I	Electrical Engineering I	Electronics I	Mechanics I	Computer Skills I	Communication Skills I	Mathematics I
Subject Code	PRJT101	ELEN101	ETRS101	MECH101	CSKL101	CSCC101	MATH101
Credit Value	0.1	0.1	0.1	0.1	0.05	0.05	0.1

These rules extracted from the Electrical Engineering Departmental Handbook for the different years are listed below. The above Table 3.3 was used, in conjunction with the Departmental Rules, listed below by year, to determine the progression of the learners into the second semester of study.

2013 and 2014

Promotion to a Higher Level

1) Semester 1 (S1) to Semester 2 (S2) - The student must pass at least 0.4 credits at level S1 before registering for subjects in the level S2.

Monitoring of Academic Progress

1) A student who does not complete half of the subjects of semester S1 in the first semester of registration will be given a warning of "UNSATISFACTORY PROGRESS".

2015 to 2018

Promotion to a Higher Study Level

1) Deleted with effect 01 January 2015.

Monitoring of Academic Progress

Students registered for this Diploma are required to meet the following progression rules:

1) The student must pass a minimum number of subjects to accumulate at least 0.25 credits for each semester of registration of the first four semesters of registration. In addition, the student must accumulate at least 1.5 credits by the end of fifth semester of registration. (This does not include Work Experience registrations.)
2) For each of study levels S1 to S3, a student must pass a minimum of 0.25 credits to be promoted to the next level.

In semesters 2013A&B and 2014A&B learners studied a course Mechanics I. This subject had a very high failure rate and was therefore deemed a bottle neck subject. It was moved to S3 on the curriculum and the progression rule was adjusted accordingly from 0.4 to 0.25. A similar situation occurred in The Department of Mechanical Engineering, where Electrotechnology I was moved to the S3 curriculum [106].

3.9 CONCLUSION

This chapter outlined the various statistical and data mining techniques that are used in the study of a cohort of learners entering DEPE using SC/NSC or TVET N4. Statiscal methods are used to complete a longitudinal analysis, while EDM, using WEKA, is used to determine which subjects influence throughput, using the DEPE Rules for progression. WEKA is also used to develop a selection tool. The chapter also outlined a comprehensive process tool for working with all forms of data. This process tool was used successfully with these learners' data. Results are presented in Chapter 4.

Chapter 4 - Results

4.1 Introduction

This chapter presents the results and describes the findings obtained from the study. As it is advisable to dichotomise the analysis [107], the results are divided into two sections, namely statistical analysis and data mining.

4.2 Statistical Analysis

The statical analysis is undertaken using SPSS version 27.0. The descriptive statistics are presented in the form of graphs, cross tabulations and other figures for the quantitative data that was collected. Inferential techniques include the use of regression and the chi-square test values, which are interpreted using the p-values. The traditional approach to reporting a result requires a statement of statistical significance. A p-value is generated from a test statistic. A significant result is indicated with "$p < 0.05$".

4.2.1 The Sample

In total, the results of 546 out of a possible 646 (85%) SC/NSC learners and 67 out of a possible 154 (44%) N4 learners, as discussed previously in Section 3.5.3, are analysed.

4.2.2 Data Analysis

This section uses descriptive statistics to evaluate the patterns in the data. The Kolmogorov tests indicate that the data is not normally distributed. Central measures are described using means and medians, with the spread estimated by the standard deviation and inter-quartile ranges, respectively. (The inter-quartile range is the difference between the 3rd quartile and 1st quartile values.)

Prior to the commencement of the testing procedures, a comparison of the English First (English A) and Second (English B) language results is undertaken within each group. The null hypothesis claims that there is no difference between English A and English B results. The alternate hypothesis claims that there is a difference between the two. The results are shown below in Table 4.1.

Table 4.1: English A and English B Results Comparison

Group	p-value
SC/NSC	0.054
TVET N4	0.424

In both instances, there is no significant differences in the scores for English A and B. Hence, the English score is treated as a single entity (henceforth referred to as Combined English).

4.2.3 Comparisons of Scores between the Groups

The descriptive measures for each of the cognate subjects completed by each of the two groups TVET N4 and SC/NSC are listed in Table 4.2 below.

Table 4.2: Results Table

	Group											
	TVET N4						SC/NSC					
	Count	Mean	Standard Deviation	Median	1st Quartile	3rd Quartile	Count	Mean	Standard Deviation	Median	1st Quartile	3rd Quartile
English (B)	67	4.47	0.65	4.00	4.00	5.00	546	5.16	0.93	5.00	4.00	6.00
English (A)	67	4.33	0.59	4.00	4.00	5.00	546	5.32	0.79	5.00	5.00	6.00
Combined English	67	4.43	0.63	4.00	4.00	5.00	546	5.20	0.89	5.00	5.00	6.00
Physical Science	67	2.95	1.25	3.00	2.00	3.00	546	5.10	0.89	5.00	4.00	6.00
Mathematics	67	3.05	1.37	3.00	2.00	4.00	546	5.38	0.89	5.00	5.00	6.00
TVET N4 Engineering Science	67	5.45	0.74	5.00	5.00	6.00	546					
TVET N4 Industrial Electronics	67	6.00	0.98	6.00	5.00	7.00	546					
TVET N4 Mathematics	67	6.10	0.96	6.00	5.00	7.00	546					
TVET N4 Electrotechnics	67	6.60	1.09	7.00	6.00	8.00	546					
Projects I (0.1) PRJT101	67	61.30	15.38	65.00	58.00	70.00	546	55.92	19.56	60.00	50.00	69.00
Electrical Engineering I (0.1) ELEN101	67	54.39	15.15	57.00	45.00	66.00	546	56.96	15.49	58.00	46.00	68.00
Electronics I (0.1) ETRS101	67	51.57	13.34	52.00	45.00	61.00	546	48.81	14.03	51.00	39.00	58.00
Mechanics I (0.1) MECH101	67	38.28	10.22	38.00	33.00	46.00	546	41.15	11.58	39.00	34.00	47.00
Communication Skills I (0.05) CSKL101	67	60.79	9.96	61.00	58.00	65.00	546	61.54	10.52	62.00	55.00	68.00
Computer Skills I (0.05) CSCC101	67	73.06	15.45	75.00	61.00	87.00	546	72.84	15.06	75.00	64.00	84.00
Mathematics I (0.1) MATH101	67	61.80	16.22	63.50	52.00	73.00	546	60.95	16.42	61.00	50.00	72.00

The SC/NSC and TVET N4 scores are the category ratings as displayed on the SC/NSC or TVET N4 certificates. For example, 80% or more is rated as 7 for the SC/NSC group, see Table 3.2. The university subjects are reflected by the mark obtained as a percentage. It is clear from Table 4.2 that the SC/NSC learners performed better than the TVET learners for Mathematics, 5.38 vs 3.05, Physical Science, 5.10 vs 2.95 and English, 5.2 vs 4.43 in the SC/NSC Examinations. What is important is that although the learners did not have the required competency, they did have the interest and relevant subjects.

4.2.4 School Subject Comparison

Figure 4.1 below, indicates the mean and standard deviation per school subject.

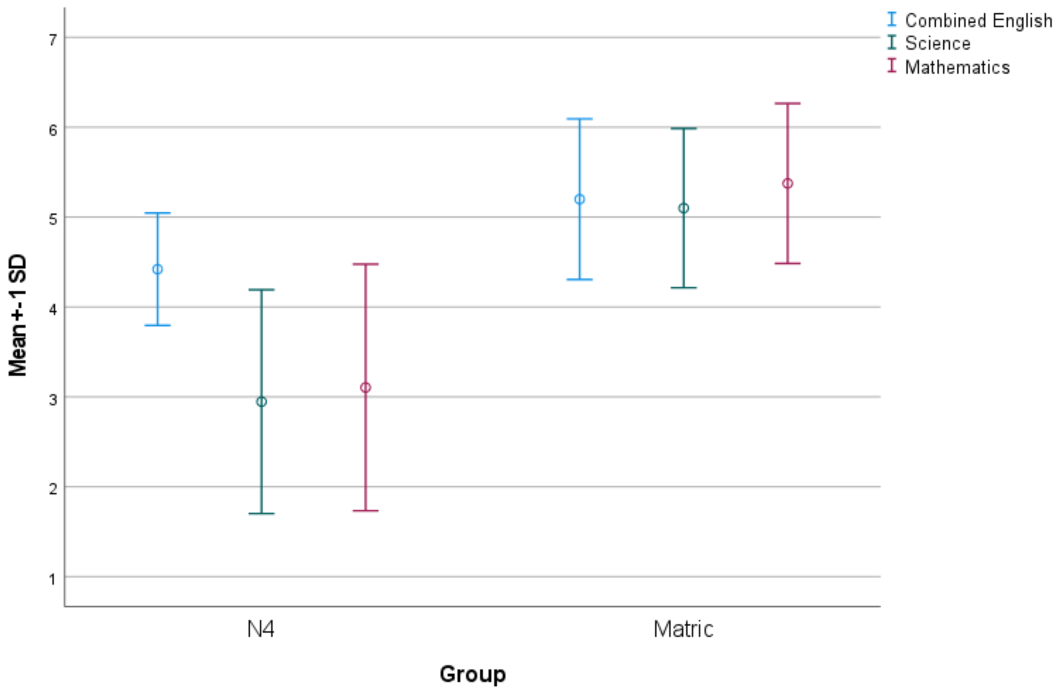


Figure 4.1: Entrance Subjects for TVET N4 & SC/NSC (Matric)

The SC/NSC university learners, those learners entering HE directly, obtained higher marks than those of the TVET N4 learners upon completion of SC/NSC, in all subjects. There is only a marginal overlap in terms of the spread, with the lower end scores of the SC/NSC university learners being in line with the upper end scores of the TVET N4 learners' SC/NSC marks. This is understandable as the entrance criteria to enroll at university is higher than that for a vocational college and the reason the learners did not get accepted directly into the HEI was because of their lower marks.

The test of significance for the differences is shown in Table 4.3 below.

Table 4.3: Significance Test

	English (B)	English (A)	Combined English	Physical Science	Mathematics
Mann-Whitney U	5803.500	467.000	9632.000	2621.000	2798.000
Wilcoxon W	7028.500	638.000	11910.000	4274.000	4751.000
Z	-4.937	-4.691	-6.658	-10.849	-11.278
Asymp. Sig. (2-tailed)	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**

The p-values are all significant ($p < 0.001$). This implies that there is a significant difference in the central values. It is noted in the Table 4.3, as well as Figure 4.1, that the SC/NSC university learners had significantly higher scores than the TVET N4 learner.

4.2.5 University Subject Comparison

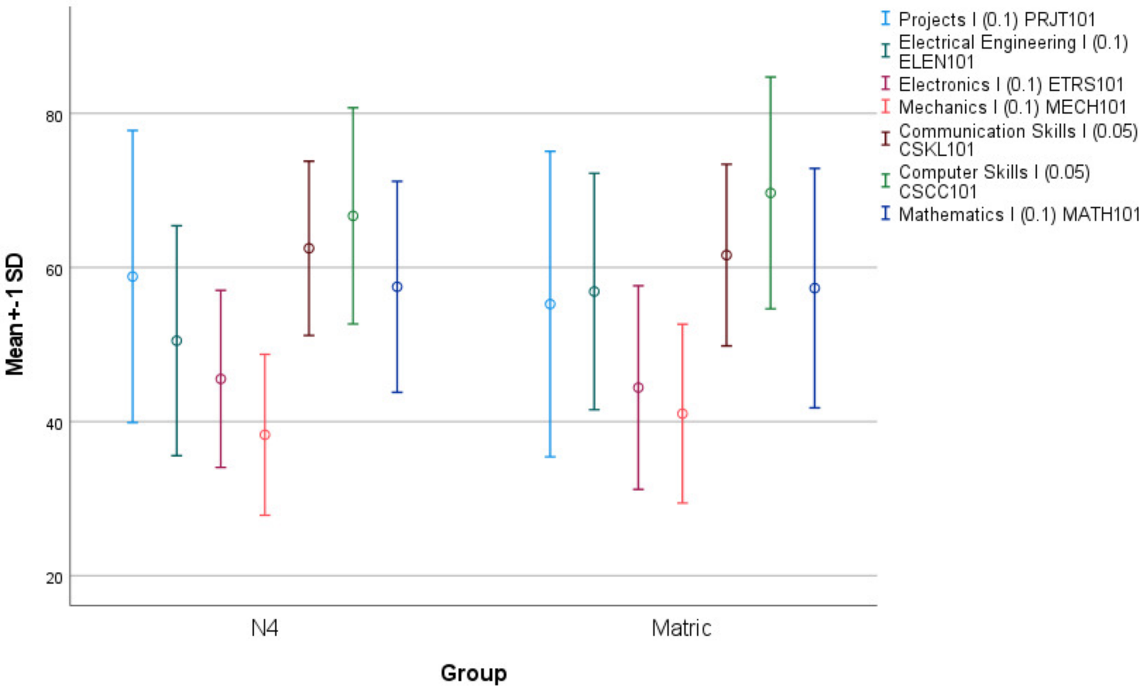


Figure 4.2: University Subject Comparison

Although there is variation in the marks between the subjects, as seen in Figure 4.2, the marks obtained per subject seem similar between the groups. Most of the subject means (medians) are similar, as is the spread. To test the claim that there is no significant difference in the central values by subject between the groups, a Mann-Whitney Test is done. The results are shown in Table 4.4.

Table 4.4: Significance Test University Subjects

	Projects I (0.1) PRJT101	Electrical Engineering I (0.1) ELEN101	Electronics I (0.1) ETRS101	Mechanics I (0.1) MECH101	Communication Skills I (0.05) CSKL101	Computer Skills I (0.05) CSCC101	Mathematics I (0.1) MATH101
Mann-Whitney U	14723.000	16073.500	15954.500	3304.500	17122.500	17326.500	16436.500
Wilcoxon W	160793.000	18284.500	164739.500	3629.500	19400.500	155927.500	156621.500
Z	-2.496	-1.189	-1.687	-0.938	-0.353	-0.223	-0.775
Asymp. Sig. (2-tailed)	**0.013**	0.235	0.092	0.348	0.724	0.824	0.438

The results show that there are no significant differences for the subjects between the groups for all subjects except Projects 1 (as all the p-values are greater than 0.05). This implies that the TVET N4 learners are performing just as well as the SC/NSC university learners in the subjects.

From the Figure 4.2 and Table 4.4, it is noted that the TVET N4 learners had marginally higher scores in 3 of the 7 subjects with the SC/NSC university learners having higher scores in 3 others. The TVET N4 learners have a significantly higher score for Projects 1 ($p = 0.013$), possibly due to these learners having completed practicals and industrial exposure at TVET colleges. This result talks to aims 1 and 2.

4.2.6 Cross Tabulations

A Chi-square test of independence is performed to determine whether there is a statistically significant relationship between the group and passing. The null hypothesis states that there is no association between the two. The alternate hypothesis indicates that there is an association.

Table 4.5 summarises the results of the chi-square tests.

Table 4.5: Chi-Squared Test

			Group		Total
			TVET N4	**SC/NSC**	
Progressed	**No**	**Count**	15	164	179
		% within Progressed	8.4%	91.6%	100.0%
		% within Group	22.4%	30.0%	29.2%
		% of Total	2.4%	26.8%	29.2%
	Yes	**Count**	52	382	434
		% within Progressed	12.0%	88.0%	100.0%
		% within Group	**77.6%**	**70.0%**	70.8%
		% of Total	8.5 %	62.3%	70.8%
Total		**Count**	67	546	613
		% within Progressed	10.9%	89.1%	100.0%
		% within Group	100.0%	100.0%	100.0%
		% of Total	10.9%	89.1%	100.0%

Overall, the ratio of TVET N4 learners to SC/NSC university learners is approximately 1:9 (10.9%: 89.1%) ($p < 0.001$). Within the category of learners who progressed (Yes), 12.0% were TVET N4 learners. Within the category of TVET N4 learners (only), 77.6% passed. This category of TVET N4 learners who passed formed 8.5% of the total sample. (This is an expected smaller overall number in terms of learners group ratio.) It is also noted that a greater proportion of TVET N4 learners progressed than SC/NSC university learners (77.6% compared to 70.0%). This result talks to aims 1 and 2.

4.2.7 Correlations

Bivariate correlation is also performed on the data. The results are shown below in Table 4.6 and Table 4.7. Positive values indicate a directly proportional relationship between the variables and a negative value indicates an inverse relationship. All significant relationships are indicated by a * or **.

Table 4.6: TVET N4 Bivariate Correlation

		Combined English	N4 Engineering Science	N4 Engineering Mathematics	N4 Electrotechnics	N4 Industrial Electronics	Projects I (0.1) PRJT101	Electrical Engineering I (0.1) ELEN101	Electronics I (0.1) ETRS101	Mechanics I (0.1) MECH101	Communication Skills I (0.05) CSKL101	Computer Skills I (0.05) CSCC101	Mathematics I (0.1) MATH101
Combined English	**Pearson Correlation**	1											
	N	67											
N4 Engineering Science	**Pearson Correlation**	0.033	1										
	N	57	57										
N4 Engineering Mathematics	**Pearson Correlation**	0.173	.274**	1									
	N	67	67	67									
N4 Electrotechnics	**Pearson Correlation**	0.005	0.102	-0.048	1								
	N	67	67	67	67								
N4 Industrial Electronics	**Pearson Correlation**	0.001	0.228**	0.208**	0.360*	1							
	N	67	67	67	67	67							
Projects I (0.1) PRJT101	**Pearson Correlation**	-0.052	0.099	-0.039	0.084	-0.103	1						
	N	67	67	67	67	67	67						
Electrical Engineering I (0.1) ELEN101	**Pearson Correlation**	-0.021	0.017	0.048	0.158	-0.276	.563**	1					
	N	67	67	67	67	67	67	67					
Electronics I (0.1) ETRS101	**Pearson Correlation**	-0.018	0.034	-0.009	0.135	-.082	.526**	.770**	1				
	N	67	67	67	67	67	67	67	67				
Mechanics I (0.1) MECH101	**Pearson Correlation**	0.050	0.187	0.006	0.174	0.021	.610**	.783**	.783**	1			
	N	25	25	25	25	25	25	25	25	25			
Communication Skills I (0.05) CSKL101	**Pearson Correlation**	.346**	0.049	0.186	0.11	-0.083	0.189	.402**	.484**	.283**	1		
	N	67	67	67	67	67	67	67	67	25	66		
Computer Skills I (0.05) CSCC101	**Pearson Correlation**	0.036	0.152	-0.006	0.137	-0.18	.387**	.513**	.685**	.282**	.556**	1	
	N	67	67	67	67	67	67	67	67	25	67	67	
Mathematics I (0.1) MATH101	**Pearson Correlation**	0.110	0.195	0.089	0.124	-.054	.312*	.675**	.643**	.281**	.423**	.584**	1
	N	67	67	67	67	67	67	67	67	25	67	67	67

**. Correlation is significant at the 0.01 level (2-tailed).*. Correlation is significant at the 0.05 level (2-tailed).a. Group = TVET N4

The null hypothesis states that there is no correlation between the variables. The alternate states that there is a correlation between the variables. The various correlations are highlighted below.

SC/NSC Subjects vs Semester 1 Subjects

There is a significant correlation between school-level English and Communication ($r = 0.346$, $p = 0.004$).

TVET Subjects

Each N4 subject has a good correlation with the other N4 subjects. This shows that subject content, subject level, and teaching are congruent and that there is synergy.

TVET vs Semester 1 Subjects

The correlations for the TVET subjects to the Semester 1 subjects are not significant. One correlation of interest is that N4 Electrotechnics does have a correlation to Electrical Engineering I. This is due to the contents of the subjects being aligned.

Semester 1 Subjects

Communications Skills I and Computer Skills I show good correlation to all of the Semester 1 subjects. These subjects are used within the “main subjects”. It is essential that learners can communicate, both verbally and written in the technology fields. Positive correlation between the “Main Subjects” indicates subject content, teaching methodology are congruent. This result talks to aim 3.

Table 4.7: SC/NSC Bivariate Correlation

		Combined English	Physical Science	Mathematics	Projects I (0.1) PRJT101	Electrical Engineering I (0.1)	Electronics I (0.1) ETRS101	Mechanics I (0.1) MECH101	Communicati on Skills I (0.05)	Computer Skills I (0.05) CSCC101	Mathematics I (0.1) MATH101
Combined English	**Pearson Correlation**	1									
	N	546									
Physical Science	**Pearson Correlation**	.246**	1								
	N	546	546								
Mathematics	**Pearson Correlation**	0.056	.399**	1							
	N	546	546	546							
Projects I (0.1) PRJT101	**Pearson Correlation**	-0.049	-0.002	0.019	1						
	N	540	540	540	540						
Electrical Engineering I (0.1) ELEN101	**Pearson Correlation**	0.043	.272**	.256**	.407**	1					
	N	535	535	535	530	535					
Electronics I (0.1) ETRS101	**Pearson Correlation**	0.005	.206**	.190**	.399**	.601**	1				
	N	545	545	545	539	535	545				
Mechanics I (0.1) MECH101	**Pearson Correlation**	-0.019	.311**	.380**	.280**	.641**	.537**	1			
	N	298	298	298	294	292	297	298			
Communicati on Skills I (0.05) CSKL101	**Pearson Correlation**	.345**	0.081	0.008	.255**	.394**	.404**	.251**	1		
	N	525	525	525	522	523	525	285	525		
Computer Skills I (0.05) CSCC101	**Pearson Correlation**	.173**	0.057	0.081	.264**	.344**	.396**	.215**	.523**	1	
	N	526	526	526	523	523	526	284	523	526	
Mathematics I (0.1) MATH101	**Pearson Correlation**	-0.057	.273**	.387**	.313**	.636**	.635**	.659**	.307**	.368**	1
	N	529	529	529	524	527	529	289	522	523	529

For this group of learners, there are significant directly proportional correlations between English and Science and, also Science and Mathematics, at matric exit level. In addition to a positive correlation between English and Communication, there are also correlations between Science and Mathematics, and various university subjects. Amongst the university subjects, there are also significant positive correlations.

SC/NSC Subjects

There is a positive correlation, 0.246, between English and Physical Science, presumably since proficiency in English enables better the understanding of the concepts, understanding and being able to communicate the correct terminology of Physical Science. There is also a good correlation, 0.399, between Mathematics and Physically Science presumably due the Mathematic content and analytical concepts in both subjects.

SC/NSC vs Semester 1 Subjects

English has a positive correlation, 0.345, to Communication Skills I and Computer Skills I, 0173. Both these subjects are 'soft skills' and are predominately about communicating in English. Mathematics has a strong correlation to all the 'main" semester 1 subjects, namely. Electrical Engineering I - 0.256, Electronics I - 0.190, Mechanics I - 0.380 and Mathematics I - 0.387. All these subjects have a theory based solidly around analysis and calculations, therefore the stronger a learner is at the basic language of mathematics the more manageable the subject will be. Physical Science also has strong a correlation to these four subjects due to similar reason.

The fact that Mathematics and Physical Science have a strong correlation to the four main subjects of Semester 1 indicates that these subjects are important selection criteria for learners entering the Electrical Engineering Department.

Semester 1 Subjects

Communications Skills I and Computer Skills I show good correlations to all the Semester 1 subjects. These subjects are used within the "main subjects". It is essential that learners can communicate, both verbally and in writing, in the technology fields. Positive correlations between the "Main Subjects" indicate subject content, teaching methodology are congruent. This result talks to aim 3.

4.2.8 Binary Logistic Regression Model

A binomial logistic regression predicts the probability that an observation falls into one of two categories of a dichotomous dependent variable based on one or more independent variables that can be either continuous or categorical. The Progression of the learner (Pass or Fail) is used as the dependent variable.

The outputs are explained in Table 4.8 below.

Table 4.8: Binary Logistic Regression Model Output Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	127.596[a]	0.602	0.815

a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.

Table 4.8 contains the Cox & Snell R Square and Nagelkerke R Square values, which are both methods of calculating the explained variation. These values are sometimes referred to as *pseudo* R^2 values (and will have lower values than in multiple regression). However, they are interpreted in the same manner, but with more caution. Therefore, the explained variation in the dependent variable based on the model ranges from 60.2% to 81.5%, depending on whether the Cox & Snell R^2 or Nagelkerke R^2 methods are referenced, respectively.

Table 4.9: Classification Table

Observed			Predicted		
			Progressed		Percentage Correct
			No	Yes	
Step 1	Progressed	No	104	16	86.7
		Yes	11	173	94.0
	Overall Percentage				91.1

a. The cut value is .500

There is a high percentage of cases that can be correctly classified (91.1%), as seen in Table 4.9.

Table 4.10: Variable in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	**Group (1)**	1.600	0.992	2.600	1	0.107	**4.951**	0.708	34.606
	Projects I (0.1) PRJT101	0.095	0.020	22.259	1	**0.000**	1.100	1.057	1.145
	Electrical Engineering I (0.1) ELEN101	0.133	0.028	22.975	1	**0.000**	1.142	1.082	1.206
	Electronics I (0.1) ETRS101	0.120	0.028	18.095	1	**0.000**	1.128	1.067	1.192
	Mechanics I (0.1) MECH101	0.100	0.035	8.367	1	**0.004**	1.105	1.033	1.183
	Communication Skills I (0.05) CSKL101	-0.053	0.029	3.444	1	0.063	0.948	0.896	1.003
	Computer Skills I (0.05) CSCC101	0.029	0.019	2.236	1	0.135	1.029	0.991	1.068
	Mathematics I (0.1) MATH101	0.136	0.028	24.392	1	**0.000**	1.146	1.085	1.209
	Constant	-27.604	4.166	43.896	1	0.000	0.000		

a. Variable(s) entered on step 1: Projects I (0.1) PRJT101, Electrical Engineering I (0.1) ELEN101, Electronics I (0.1) ETRS101, Mechanics I (0.1) MECH101, Communication Skills I (0.05) CSKL101, Computer Skills I (0.05) CSCC101, Mathematics I (0.1) MATH101.

The Wald test is used to determine statistical significance for each of the independent variables. The statistical significance of the test is found in the "**Sig.**" column, as seen in Table 4.10. From these results it is noted that Projects I (0.1) PRJT101, Electrical Engineering I (0.1) ELEN101, Electronics I (0.1) ETRS101, Mechanics I (0.1) MECH101 and Mathematics I (0.1) MATH101 added significantly to the model/prediction, but the others did not.

The probability of an event occurring based on a one-unit change in an independent variable when all other independent variables are kept constant, can be predicted using the odds ratio (Exp(B)). For example, Table 4.10 shows that the odds of passing semester 1, is 4.951 times greater for TVET N4 learners as opposed to the SC/NSC university group, even though the p-value is not significant.

Most coefficients (B) are positive, implying that the probability of passing is greater with higher marks. This may be obvious, but it indicates the highlighted p-value subjects show greater changes in passing compared to the non-highlighted ones. Even though it is not significant, Communication Skills shows a marginal decrease in the pass rate (-0.053). The odds ratio for all other subjects is approximately 1.

4.3 Educational Data Mining

The Waikato Environment for Knowledge Analysis (WEKA) classification tool, Version 3.8.5 [104, 105], as described in section 3.5.5, is used as a data mining tool. Several classification and filtering tools within WEKA are used for this study. WEKA is an easy-to-use collection of machine learning algorithms for data mining tasks. The algorithms can be applied directly to a dataset. WEKA contains tools for data pre-processing, classification, regression, clustering, association rules, and visualisation.

4.3.1 SC/NSC Subject Input Evaluator

Several input evaluator algorithms, namely Pearsons, Chi-squared, Support Vector Machine, Decision Tree Wrapper and Information Gain are used to evaluate the ranking of the input subjects English, Mathematics and Physical Science against the Department of Electrical Power Engineering progression rule into Semester 2 (S2), using WEKA, shown in Annexure 4.

The input attributes are evaluated using the Correlation Attribute Evaluator within WEKA. This evaluator applies the Pearsons correlation method [108] which ranks all the input attributes in order of their correlation to the class variable or output variable. The output variable is a created variable which denotes if the learners were successful or deemed to have been promoted to the second semester. This is done using The Department of Electrical Power Engineering's Promotion rules as seen in Section 3.6.

A Chi-square method shows the relationship between two categorical variables. This is one number that indicates how much difference exists between the observation and if there was no correlation. A low value for chi-square indicates a high correlation. Although a low value indicates a high correlation, deciding whether the chi-square test is large enough to indicate a statistically significant difference is not easy. It may not be the case that, if the value is greater than 10, there is a big difference.

The SVM evaluator performs a greedy forward or backward search through the space of attribute subsets. It implements a sequential minimal optimization algorithm for the support vector classifier to determine which input variable/attribute best matches the output class.

The DT wrapper subset evaluator uses the J48 decision tree to split the attributes heuristically on a top-down approach evaluating the input attributes against the output class.

This information gain evaluator evaluates the worth of the attribute by measuring the information gain with respect to the output class.

Table 4.11: Summary of SC/NSC Attribute Evaluation

Selected Feature	**SVM**	**DT Wrapper**	**Pearsons Correlation**	**Chi Squared**	**Information Gain**
Mathematics	Selected (2) (0.7)	Selected (3) (0.692)	0.19	19.359	0.0239
Physical Science	Selected (3) (0.7)	Selected (2) (0.7)	0.12	0	0
English	Selected (1) (0.7)	Selected (1) (0.7)	0.07	0	0

The only attribute evaluator that produces a set of results is the Pearsons Correlations Rank Evaluator, seen in Table 4.11 and it ranks Mathematics as having the greatest effect on the progress of the learners into S2. An examination of the content of the subjects in S1, Electrical Engineering I, Mechanics I, Electronics I and Mathematics I, all include significant portions of mathematical content. It is therefore feasible and plausible that learners that perform better in SC/NSC Mathematics will do better in S1. This result talks to aim 3.

4.3.2 TVET Subject Input Evaluator

Several input evaluator algorithms, e.g. Pearsons, Chi-squared, Support Vector Machine, Decision Tree Wrapper and Information Gain, seen in Annexure 5 are used to evaluate the ranking of the input subjects English, N4 Engineering Mathematics, N4 Engineering Science, N4 Electrotechnics and N4 Industrial Electronics against the progression rule into S2.

The same tests are performed for the TVET N4 subjects as SC/NSC tests.

Table 4.12: Summary of TVET N4 Attribute Evaluation

Selected Feature	SVM	DT Wrapper	Pearsons Correlation	Chi Squared	Information Gain
N4 Engineering Science	0.776 (2)	0.776 (2)	0.1801	0(2)	0 (2)
N4 Mathematics	0.776 (4)	0.776 (4)	0.1347	0(4)	0 (4)
N4 Electrotechnics	0.776 (5)	0.776 (5)	0.1677	0(1)	0 (1)
SC/NSC English	0.776 (1)	0.776 (1)	0.0859	0(5)	0 (5)
N4 Industrial Electronics	0.776 (3)	0.776 (3)	0.0366	0(3)	0 (3)

The Pearson Correlation method is selected, as it was the only evaluator that produced results that are able to be used. Although the Pearsons Correlations Evaluator, shown in Table 4.12, ranked N4 Engineering Science, followed by N4 Electrotechnics and then N4 Mathematics, the results are very close together. This might be because the content of N4 Engineering Science overlaps or is in alignment with the other engineering subjects; and the N4 Electrotechnics assisted the learners with Electrical Engineering I and all these learners were building on previous knowledge from SC/NSC. This result talks to aim 3.

4.3.3 Evaluator Ranking Results Comparison

The summary of both entrance subject evaluator results, as seen in Table 4.13, for the SC/NSC evaluation ranks Mathematics, then Science and finally SC/NSC English as indicators. The evaluator ranks N4 Engineering Science, N4 Electrotechnics, N4 Mathematics, SC/NSC English and finally N4 Industrial Electronics. This result talks to aim 3.

Table 4.13: TVET N4 and SC/NSC Input Results Comparison

TVET N4 Pearsons Correlation Evaluation Ranking Results	SC/NSC Pearson Correlation Evaluation Ranking Results
N4 Engineering Science	Mathematics
N4 Electrotechnics	Physical Science
N4 Mathematics	English
SC/NSC English	
N4 Industrial Electronics	

4.4 TVET N4 Educational Data Mining

The WEKA Data mining tool [104], seen in Annexure 6,is used for the Educational Data Mining using the following classification models Logistic Regression, Naïve Bayes, Artificial Neural Networks, Support Vector Machine and Decision Trees. These classification methods have been discussed in detail in Chapter 3. An extra classification model, that of ZeroR is also used. ZeroR is a base classification model that uses the “most popular’ class within the data to determine the output variable. All classification models should be viewed against ZeroR, as this is the minimum accuracy expected. The input variables used are the four TVET N4 subjects namely N4 Engineering Mathematics, N4 Engineering Science, N4 Industrial Electronics and N4 Electrotechnics, as well as SC/NSC English and the output variable is the DEPE progression Rule. Each classifier is run to determine which could use the input variables to predict the output variable which the highest accuracy.

4.4.1 Summary TVET N4 Success Predictor

The results for the data mining are tabulated in Table 4.14, Table 4.15 and Table 4.16.

Table 4.14: TVET N4 Data Mining Accuracy Table

	Algorithm					
	ZeroR	**LR**	**NB**	**ANN**	**SVM**	**DT (J48)**
Correctly Classified Instances	52	51	53	52	52	51
Incorrectly Classified Instances	15	16	14	15	15	16
Prediction Accuracy	77.61	76.12	79.10	77.61	77.6	76.12

Table 4.15: TVET N4 Data Mining Error Table

	Algorithm					
	ZeroR	**LR**	**NB**	**ANN**	**SVM**	**DT (J48)**
Mean Absolute Error	0.35	.35	0.34	0.37	0.22	0.36
Kappa Statistic	0	-.0.3	0.10	0	0	-0.0288
Root Mean Squared Error	0.42	0.43	0.42	0.44	0.47	0.44
Relative Absolute Error	100%	99.96%	96.15%	104.07%	63.39%	101.58%
Relative Squared Error	100%	103.59%	100.06%	104.79%	113.17%	104.26%

Table 4.16: TVET N4 Data Mining Statistics Table

Algorithm	**TP Rate**	**FP Rate**	**Precision**	**Recall**	**F-Measure**	**MCC**	**ROC Area**	**PRC Area**
ZeroR	**0.766**	**0.766**		**0.776**			**0.403**	**0.619**
LR	**0.761**	**0.780**	**0.600**	**0.761**	**0.671**	**-0.066**	**0.513**	**0.642**
NB	**0.791**	**0.724**	**0.835**	**0.791**	**0.712**	**0.229**	**0.575**	**0.697**
ANN	**0.776**	**0.776**		**0.776**			**0.378**	**0.606**
SVM	**0.776**	**0.776**		**0.776**			**0.500**	**0.652**
DT (J48)	**0.761**	**0.780**	**0.600**	**0.761**	**0.671**	**-0.066**	**0.384**	**0.611**

Although NB gave the best prediction accuracy of 79.10%, better than the ZeroR classifier's base line accuracy of 77.61%, the Kappa statistic was only 0.1. The 0.1 Kappa statistic is better than the 0 of the ZeroR classifier. A Kappa of between 0.2 – 0.40 is considered fair. Thus, a better classifier was sought.

The reason for the low Kappa values and prediction accuracy is the small data set. Unfortunately, almost 50% of the learners' results were eliminated from the test set due to missing information. These results will improve on larger test sets or more accurate information being captured. This result talks to aims 4 and 5.

4.4.2 TVET N4 IBK Alternate Classifier

The Instance Based K-Nearest Neighbour (IBK) classifier within WEKA, shown in Annexure 7, is used as an alternate to those already discussed. IBK is a lazy based algorithm which delays all predictions until called upon to do so, therefore doing no actual learning. The algorithm then uses a distance-based measure to locate k closest instances in the training data to make a prediction. In basic terms, it seeks a match to the given input sample and then provides an output. This result talks to aims 4 and 5.

4.4.3 TVET N4 IBK Alternate Classifier Summary

Table 4.17: TVET N4 Data Mining Accuracy Table (Alternate)

	Algorithm
	IBK
Correctly Classified Instances	52
Incorrectly Classified Instances	15
Prediction Accuracy	77.61%

Table 4.18: TVET N4 Data Mining Error Table (Alternate)

	Algorithm
	IBK
Kappa Statistic	0.23
Mean Absolute Error	0.33
Root Mean Squared Error	0.45
Relative Absolute Error	94.77%
Relative Squared Error	107 89%

Table 4.19: TVET N4 Data Mining Statistics Table (Alternate)

Algorithm	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
IBK	0.776	0.586	0.743	0.776	0.749	0.244	0.509	0.671

The IBK classifier in this case has given a prediction accuracy of 77,61%, see Table 4.17, Table 4.18 and Table 4.19, but the Kappa statistic is 0.23, indicating that it has performed fairly and better than all the other classifiers. As mentioned previously the TVET N4 dataset was very small, only 67 learners and therefore other classifiers did not have enough training data.

4.5 SC/NSC Educational Data Mining

As mentioned in the previous section, the WEKA Data mining tool [104], shown in Annexure 8, is used for the following classification models Logistic Regression, Naïve Bayes, Artificial Neural Networks, Support Vector Machine and Decision Trees . ZeroR, the base classification model that uses the "most popular' class within the data to determine the output variable, is used again. All classification models should be viewed against ZeroR, as this is the minimum accuracy expected. The input variables used are the three subjects from SC/NSC used to select learners namely Mathematics, Physical Science and English. The output variable is the DEPE progression Rule. Each classifier is run to determine which could use the input variables to predict the output variable with the highest accuracy. This result talks to aims 4 and 5.

4.5.1 Summary SC/NSC Success Predictor

Table 4.20: SC/NSC Data Mining Accuracy Table

	Algorithm					
	ZeroR	**LR**	**NB**	**ANN**	**SVM**	**DT (J48)**
Training Instances (80%)	546	546	546	546	546	546
Test Instances (20%)	109	109	109	109	109	109
Correctly Classified Instances	75	75	75	78	75	75
Incorrectly Classified Instances	34	34	34	31	34	34
Prediction Accuracy	68.81%	68.81%	68.71%	71.56%	68.81%	68.81%

Table 4.21: SC/NSC Data Mining Error Table

	Algorithm					
	ZeroR	**LR**	**NB**	**ANN**	**SVM**	**DT (J48)**
Kappa Statistic	0	0.02	0.02	0.20	0	0
Mean Absolute Error	0.42	0.40	0.40	0.39	0.31	0.42
Root Mean Squared Error	0.46	0.44	0.445	0.44	0.56	0.46
Relative Absolute Error	100%	94.83%	94.57%	92.46%	73.54%	99.92%
Relative Squared Error	100%	95.98%	96.00%	95.87%	120.50%	100.00%

Table 4.22: SC/NSC Data Mining Statistics Table

Algorithm	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
ZeroR	0.688	0.688	?	0.688	?	?	0.500	0.571
LR	0.688	0.672	0.632	0.688	0.577	0.056	0.695	0.716
NB	0.688	0.672	0.632	0.688	0.577	0.056	0.694	0.714
ANN	0.716	0.547	0.694	0.716	0.670	0.241	0.665	0.690
SVM	0.688	0.688	?	0.688	?	?	0.500	0.571
DT (J48)	0.688	0.688	?	0.688	?	?	0.500	0.571

The EDM results are tabulated above in Table 4.20, Table 4.21 and Table 4.22. The Artificial Neural Network Classifier (ANN) performed the best out of all the models tested with a prediction accuracy of 71.58%, well above the baseline ZeroR accuracy of 68.81%. The Kappa statistic for this model is 0.2, which is fair.

4.6 SC/NSC Mathematics vs TVET N4 Mathematics Syllabi Review

All learners articulating into qualifications offered by the Department of Electrical Power Engineering at DUT from any TVET college have either failed SC/NSC Mathematics or Physical Science or both, but they have all completed the academic year and obtained the required institutional entrance requirements. More specifically, 66% of learners had failed Mathematics at SC/NSC, see Table 4.24. These learners have then attended a TVET college and through mutual agreement, not only completed N4 Mathematics and N4 Engineering Science but two other electrical related subjects; N4 Electrotechnics and N4 Industrial Electronics. This result talks to aim 6.

Table 4.23: TVET N4 vs SC/NSC Mathematics Syllabi

TVET N4 Mathematics Syllabi	SC/NSC Mathematics Syllabi
Equations, Manipulation, and word problems	Algebra
Determinants	
Complex Numbers	
Trigonometry	Trigonometry
Sketch Graphs	Functions
Limits and Differentiation	Differential Calculus
Integration	
	Numbers Patterns, Sequences, Series
	Finance, growth, and decay
	Probability
	Euclidean Geometry and Measurement
	Analytical Geometry
	Statistics

The SC/NSC Mathematics syllabi [109] shown in the right-hand side column of Table 4.23. All learners intending to enter engineering at a HE institution would have completed the above or similar syllabus. The TVET N4 Engineering Mathematics Syllabi [110], shown on the LHS is what would have been covered during the trimester at the TVET College. Since all the learners have already completed four of the seven sections, see Table 4.23, from the Mathematics syllabus at school, the syllabus presented at the TVET college might be seen as revision, ensuring that the learner is able to obtain the required level of competency previously missed. The N4 Mathematics syllabus [110] is clearly aligned to engineering, as listed in the general subject aims for the subject, where engineering is specified. This syllabus omits sections such as finance and statistics.

Many reasons have been proffered for the poor performance of learners in the SC/NSC Mathematics subject. Reddy has indicated that South African Mathematics and Science educators' standards, benchmarked internationally as amongst the lowest [111].

Table 4.24: TVET N4 Mathematics vs Semester 1 Mathematics Success Rate

	Failed SC/NSC Mathematics	**TVET N4 Mathematics**	**Mathematics I (MATH101)**
Number of Learners	89/135	135	135
Average %			59.85%
Number Passed			112
% Passed			83%
% Standard Deviation			14.85%

The writer is of the view that having already completed most sections of the N4 syllabus for the SC/NSC, the articulating learners have a reduced workload, increased maturity and because of the focus on engineering, they have a far better understanding of their chosen career path. This enables the learners to achieve the required competency level as borne out by the comparative results for the mathematics S1 results, shown in Table 4.24.

4.7 SC/NSC Physical Science vs TVET N4 Engineering Science Syllabi Review

Like Mathematics, all learners articulating into the Department of Electrical Power Engineering at DUT from any TVET college have either failed SC/NSC Mathematics or Physical Science or both, but they have all completed the academic year and obtained a Diploma pass which allowed them to enter HE. These learners have then attended a TVET college and through mutual agreement, not only completed N4 Mathematics and N4 Engineering Science, but two other electrical related subjects, namely, N4 Electrotechnics and N4 Industrial Electronics. Both subjects have portions of science embedded in their content.

Since all the learners have already completed four of the seven sections from the Physical Science syllabus at school, the syllabus presented at the TVET college may be construed as revision, ensuring that the learner is able to obtain the required level of competency previously not achieved. The N4 Engineering Science syllabus [112], seen in Table 4.25, is clearly aligned to engineering, as listed in the general subject aims for the subject, where engineering is specified, but certain sections such as electricity and magnetism are omitted. These sections are covered in more detail in Electrotechnics N4.

For reasons similar to those listed for Mathematics the writer espouses the view that the learners have been able to achieve the required competency level and borne out by the comparative results for S1. It is noteworthy that N4 Engineering Science ranks first, as a subject impacting the learners Semester 1 progression, as seen in the next section.

Table 4.25: TVET N4 Engineering Science vs SC/NSC Physical Science Syllabi

TVET N4 Engineering Science Syllabi	SC/NSC Physical Science Syllabi
Kinetics	Mechanics
Dynamics	
Angular Motion	
Statics	
Hydraulics	
Stress, Strain and Young's Modulus	
Heat	Change (Chemistry)
	Systems (Chemistry)
	Waves, Sound and Light
	Electricity and Magnetism

The subject N4 Engineering Science does not have a direct subject comparison in Semester 1. It has some overlap with Mechanics 1. Although the subject Mechanics 1 was moved to Semester 3 in 2015 (due to the high failure rate), the topics covered in Engineering Science are general in nature and form the basics every engineer is required to know. The SC/NSC results for these TVET learners show that 107 learners failed Physical Science (nearly 80%); more than the number who failed Mathematics. As mentioned previously, Reddy has indicated that South African Mathematics and Science educators' standards, benchmarked internationally as amongst the lowest [88]. This result talks to aim 6.

4.8 Tracked Cohort of First-time registering learners

All the previous analyses, both statistical and EDM, have only concentrated on the data of the learners in semester 1 and the progression of the learners from semester 1 to semester 2. No analysis has been done between the cohorts of learners further than that, namely semesters 2 to 4, due to the complexity and diverse nature of the learners' paths from semester 2 to 4. It would be extremely difficult to track and compare the diverse nature of each learner's record accurately and draw any accurate conclusions. There are multiple variables that affect the data such as dropout rates, the number of courses chosen from one semester to the next, and the various Working Integrated Learning (WIL) options afforded students across the subsequent semesters. The most reliable analysis of performance beyond semester 2 is the overall tracking of student performance for reporting purposes to DHET. These statistics are available from the Department of Management Information Systems (MIS), which tracks a cohort of students registering for the first time at a tertiary institution and completing in minimum time (3 years), minimum time + 1 (4 years) and minimum time + 2 (5 Years). These results may then be compared across the two cohorts of learners, to obtain a final overall impression of the learners' ability and comparability in the National Diploma: Electrical Engineering as seen in Table 4.26, Table 4.27, Table 4.28, Table 4.29 and Table 4.30. This analysis includes all learners, compared to the S1 analysis were certain learners were excluded due to missing information.

Table 4.26: Cohort Study of 3 Year Programmes First-time Entering Students for 2013

SC/NSC / TVET	No. First-time Entering 2013	No. Graduated in			No. Returning in 2018	Throughput Rate			No. Dropped out	Dropout Rate	Still in Progress
		2015	2016	2017		Min Time	Min Time + 1	Min Time + 2			
SC/NSC	171	8	44	22	10	5%	26%	13%	87	51%	6%
TVET	37	1	9	8	7	3%	24%	22%	12	32%	19%
Total	**208**	**9**	**53**	**30**	**17**	**4%**	**25%**	**14%**	**99**	**48%**	**8%**

For the first cohort of learners, who registered for the first time in 2013, the throughput rates are comparable. These results show that the TVET learners graduate in relatively comparable numbers as the SC/NSC learners.

Table 4.27: Cohort Study of 3 Year Programmes First-time Entering Students for 2014

SC/NSC / TVET	No. First-time Entering 2014	No. Graduated in			No. Returning in 2019	Throughput Rate			No. Dropped out	Dropout Rate	Still in Progress
		2016	2017	2018		Min Time	Min Time + 1	Min Time + 2			
SC/NSC	188	8	26	19	11	4%	14%	10%	124	66%	6%
TVET	21	0	1	5	1	0%	5%	24%	14	67%	5%
Total	**209**	**8**	**27**	**24**	**12**	**4%**	**13%**	**11%**	**138**	**66%**	**6%**

For this cohort there are appreciable differences in the throughput rates for minimum time and minimum time + 1, but a huge improvement in the throughput rate for minimum time + 2.

Table 4.28: Cohort Study of 3 Year Programmes First-time Entering Students for 2015

SC/NSC / TVET	No. First-time Entering 2015	No. Graduated in			No. Returning in 2020	Throughput Rate			No. Dropped out	Dropout Rate	Still in Progress
		2017	2018	2019		Min Time	Min Time + 1	Min Time + 2			
SC/NSC	123	11	23	18	9	9%	19%	15%	62	50%	7%
TVET	34	1	1	10	4	3%	3%	29%	18	53%	12%
Total	**157**	**12**	**24**	**28**	**13**	**8%**	**15%**	**18%**	**80**	**51%**	**8%**

The TVET learners in this cohort fell behind the SC/NSC learners in minimum time and minimum time + 1 but catch up tremendously in minimum time + 2.

Table 4.29: Cohort Study of 3 Year Programmes First-time Entering Students for 2016

SC/NSC / TVET	No. First-time Entering 2016	No. Graduated in			No. Returning in 2021	Throughput Rate			No. Dropped out	Dropout Rate	Still in Progress
		2018	2019	2020		Min Time	Min Time + 1	Min Time + 2			
SC/NSC	89	4	14	12	11	4%	16%	13%	48	54%	12%
TVET	21	2	8	2	0	10%	38%	10%	9	43%	0%
Total	**110**	**6**	**22**	**14**	**11**	**5%**	**20%**	**13%**	**57**	**52%**	**10%**

In this cohort, the TVET learners out-perform the SC/NSC learners in both minimum time and minimum time + 1 and having a similar result for minimum time + 2.

Table 4.30: Cohort Study of 3 Year Programmes First-time Entering Students for 2017

SC/NSC / TVET	No. First-time Entering 2017	No. Graduated in			No. Returning in 2021	Throughput Rate			No. Dropped out	Dropout Rate	Still in Progress
		2019	2020			Min Time	Min Time + 1				
SC/NSC	76	2	10		20	3%	13%		44	58%	26%
TVET	25	0	4		6	0%	16%		15	60%	24%
Total	**101**	**2**	**14**		**26**	2%	14%		**59**	**58%**	**26%**

For the 2017 cohort the throughput rates for minimum time and minimum time + 1 are comparable. These results talk to aims 7 and 8.

Chapter 5 - Conclusion

5.1 Introduction

The three most common pathways for learners to be admitted to DUT are via: (i) SC/NSC or the 'traditional' route on exiting school; (ii) transfer from a traditional university – usually to avoid academic exclusion from that institution; and (iii) articulation from a TVET College. The second route did not form part of this investigation as learners transferring from degrees at a traditional university, usually enter from the second or third semester onwards depending on credits obtained. The first group is the usual route and was used as a control group. The focus of this study is the cohort of learners articulating from the TVET/FET College into the DEPE at DUT. The findings are summarised in the subsequent sections.

5.2 Summary

1 *Conduct a longitudinal study on the performance of articulated TVET college learners, from semester 1 into semester 2 of the National Diploma: Electrical Engineering programme at DUT.*

The statistical analysis has shown that learners entering DEPE from a TVET College on completion of the TVET N4 Electrical Certificate with N4 Engineering Mathematics, N4 Engineering Science, N4 Electrotechnics and N4 Industrial Electronics performed as well as, if not better in all the S1 subjects for the National Diploma: Electrical Engineering. This indicates that this method of articulation has demonstrated success and may be considered a viable option for both learners and HEIs to promote as an alternate access instrument.

2 *Conduct a comparative study on the performance, from semester 1 into semester 2, between the aforementioned learners and those learners who were admitted into the National Diploma: Electrical Engineering programme at DUT directly having achieved a SC or NSC.*

From the comparative analysis, both cohorts of learners are able to complete the DEPE curriculum to the required competency. It would be of interest to investigate their further development in industry. There is a supply shortage of learners with the required STEM subjects, at the required competences, for all academic institutions. Learners need to be assisted to obtain those competences if they don't obtain them when exiting the schooling system. This study has indicated that articulation is a viable alternative admission to the National Diploma: Electrical Engineering.

3 *Review and rank, using Educational Data Mining (EDM), for both TVET and SC/NSC, the subjects affecting progression from semester 1 into semester 2.*

The subject ranking varies for the TVET and SC/NSC, but curriculum relevance in the fields of STEM is a significant factor. It is the writer's view that vocational guidance and psychometric testing is necessary to enable correct career choice.

4 *Review and analyse methods of Educational Data Mining (EDM) that can be used to select learners from TVET Colleges, based on progression from semester 1 into semester 2 using the results from the TVET subjects as input variables.*

5 *Review and analyse methods of Educational Data Mining (EDM), that can be used to select learners from SC/NSC, based on progression from semester 1 into semester 2 using the results from the SC/NSC subjects as input variables.*

Although a different EDM method was obtained for each cohort, they were tested successfully. Educational Data Mining is a relatively new and exciting method of data analysis but is an experimental science and an art. The art is related to the experience of the user working with classifiers and the experimental part is using the classifier to obtain the best results from the data. The tools available are extremely useful, powerful and with the right information, vast amounts of data may be changed into new knowledge.

Educational Data Mining was used in this research on two separate cohorts of learners with differing results. An Artificial Neural Network was able to predict a 71.56% success, as shown in Table 4.21, for the SC/NSC learners, while not being able to do a similar prediction for the TVET learners. An IBK classification returned an accuracy of 77.61%, as shown in Table 4.17 for the TVET learners. These tools may in the future be able to select learners for engineering programmes.

6 *A basic qualitative comparative analysis of the Science and Mathematics (STEM) subjects from FET and Umalusi curricular will be undertaken.*

The National Curriculum Statement Grades R-12 (NCS) stipulates policy on curriculum and assessment in the schooling sector. To improve implementation, the National Curriculum Statement was amended, with effect from January 2012. A single comprehensive Curriculum and Assessment Policy document was developed for each subject to replace Subject Statements, Learning Programme Guidelines and Subject Assessment Guidelines in Grades R-12. These documents for the subject Mathematics and Physical Science are extremely comprehensive and if delivered correctly should equip learners for HE. Unfortunately, this is often not the case, and learners enter HE ill-equipped for the rigours of undergraduate programmes or often under-achieving on the NCS outcomes. The curriculum statements for the TVET subjects have not been updated since 1996.

7 *Conduct a longitudinal analysis of throughput of articulating TVET learners and SC/NSC learners, in National Diploma: Electrical Engineering over separate 5 cohorts as per the DHET reporting standard. (Tracked cohort of students registering for the*

first time at a tertiary institution and completing in minimum time (3 years), minimum time + 1 (4years), and minimum time + 2 (5 years).

8 It is suggested that the results for minimum time (3 years), minimum time + 1 (4 years), and minimum time + 2 (5 years) be viewed collectively due to the difficulty learners experienced with WIL. The WIL component of course seriously affected throughput. The overarching question would be, “How do these learners do compared to others at the end of the day?” The DHET cohort tracking report demonstrates that there is no definitive difference between the throughput rates of an articulating TVET learner and a SC/NSC learner.

5.3 Further Research

To further enhance and remove any doubt regarding the findings on articulating learners, similar research will need to be applied to the new degree programme - the BEng Tech Power Engineering. The Department of Electrical Power Engineering had its first intake of learners complete this qualification at the end of the 2020 academic year. An obvious limitation that will have to be factored in is the affect that Covid-19 may have had on learners’ results. Similar studies could also be undertaken in other Departments to determine if similar trends exist. Additional research with EDM also needs to be done involving some of the following factors: socioeconomic factors – family income, parents education, head of household occupation, dwelling value, dwelling location, financial support; psycho-social factors – academic self-efficacy, stress and time pressure, class communication, HE activity participation, organization and attention to study, sense of loneliness; individual attributes – age, work status, home language, home province, home country, interest in sports, academic results [94], as these factors have been shown to affect learner performance at HEI in other EDM studies.

5.4 Concluding Statement

Many learners leave the school system without the necessary entrance requirements to be admitted directly into a Higher Educational Institution (HEI). Through articulation, they have a second chance to attend an HEI. Articulation, as defined in the Articulation Policy for Post-School Education and Training System of South African stipulates, amongst others, that if a learner completes a course at one institution and has gained the relevant knowledge and skills at the necessary levels, this must be recognised by other institutions if the knowledge gained is sufficient to allow epistemological access to programme(s) that the learner wants to enter [47]. Learners must be supported in their individual learning and work pathways. Through articulation and being directed to TVET colleges to acquire relevant subjects, learners can gain access to their chosen programme.

References

[1] Department of Statistics (DOS), "Quarterly Labour Force Survey (QLFS) - Quarter 1: 2021," ed. Pretoria: Republic of South Africa (RSA): Department of Statistics (DOS), 2021.

[2] S. von Solms and H. Nel, "STEM project based learning: Towards improving secondary school performance in mathematics and science," *IEEE AFRICON 2017,* pp. 770 - 775, 30th July 2021 2017, doi: 10.1109/AFRCON.2017.8095580.

[3] Department of Labour (DOL), "Skills Shortages in South Africa: Case Studies of Key Professions," 2010.

[4] Sasol, "Reaching new frontiers in STEM education," 30th July 2021 2012. [Online]. Available: http://www.sasolinzalofoundation.org.za/sasol_inzalo/downloads/Inzalo_brochure_1332954308300.pdf.

[5] S. Moodley, "Poor maths science education at heart of SA's skills problem," in *Engineering News*, ed. Johannesburg: Creamer Media, 2014.

[6] The Presidency: Republic of South Africa, "ASGISA Annual Report 2007," Government Printers, Pretoria: Republic of South Africa (RSA), 2007.

[7] Department of Basic Education (DBE), "National Senior Certificate Examination Technical Report," Department of Basic Education, Pretoria: Republic of South Africa (RSA), 2012.

[8] *Department of Electrical Power Engineering Handbook,* D. U. o. T. (DUT), 2013 - 2017.

[9] *Durban University of Technology General Rule Book,* D. U. o. T. (DUT), 2013 - 2017.

[10] M. Cosser, "The Skills cline: Higher Education and the Supply-demand Complex in South Africa," *Higher Education Journal,* vol. 59, pp. 43 - 53, 2010.

[11] *National Plan for Higher Education*, March 2001.

[12] L. Bunting, "Measuring Institutional change: The Application of Two Theoretical Models to Two South African Higher Education Institutions," Masters, Department of Education, University of Cape Town, Cape Town, 2002.

[13] Organisation for Economic Co-operation and Development (OECD), "Education at a Glance 2008," *OECD 2008,* 2008. [Online]. Available: https://www.oecd.org/education/skills-beyond-school/41284038.pdf.

[14] R. du Pre, J. Reddy, and G. Scott, "The philosophy of a University of Technology in South Africa: An Introduction," *Sediba sa Thuto,* vol. 1, pp. 19-37, 2004.

[15] South African Qualifications Authority (SAQA), "Articulation Between Technical and Vocational Education and Training (TVET) Colleges and Higher Education Institutions (HEIs): National Articulation Baseline Study Report," October 2017.

[16] *Management Plan for The Administration and Management of The Certificate of Competency (Mines and Factories),* D. o. H. E. a. T. (DHET), 2020.

[17] *Occupational Health and Safety Act (Act 85 of 1993),* D. o. H. (DOH), 1993.

[18] *Policy for Credit Accumulation and Transfer within the National Qualifications Framework,* S. A. Q. A. (SAQA), 2014.

[19] J. Case, "Issues Facing Engineering Education in South Africa," presented at the 3rd African Regional Conference on Engineering Education, Pretoria, 2006.

[20] South African Qualifications Authority (SAQA), "Report on the 3rd NQF Conference: Strengthening Learning-and-Work Pathways in Community Development, Early Childhood Development and Engineering," presented at the 3rd NQF Conference, Johannesburg, 2019.

[21] IEEE, "Spectrum Forecasters STEM Survey Report," IEEE Spectrum, New York, Third Quarter 2013.

[22] *Funding of Public Higher Education, Higher Education Act of 1997 (Act 101 of 1997),* D. o. E. (DOE), 2003.

[23] C. Bundy, "Global patterns, local options? Changes in higher education internationally and some implications for South Africa," *Ten years of Higher Education under Democracy,* vol. Winter 2001, no. 4, pp. 1 - 20, 2006. CHE Kagisano.

[24] M. Hart, *Gold. The Race for the World's most Seductive Metal*. London: Simon & Schuster, 2013.

[25] K. Hart and V. Padayachee, "A history of South African capitalism in national and global perspective," *Transformation 81/82,* pp. 55 - 85, 2013.

[26] South African History Online. "South African History Online (SAHO), South Africa: 1900 – 1976. From the Union to the Republic, 2016." http://www.sahistory.org.za/topic/union-south-africa-movement-towards-republic (accessed 25th November 2020).

[27] *Criteria for Programme Accreditation,* C. o. H. E. (CHE), 2004.

[28] *Criteria for Programme Accreditation (Revised),* C. o. H. E. (CHE), 2012.

[29] I. Bunting, "The Higher Education Landscape Under Apartheid," in *Transformation in Higher Education. Higher Education Dynamics*, vol. 10, N. Cloete, P. Maassen, R. Fehnel, T. Moja, T. Gibbon, and H. Perold Eds. Dordrecht,: Springer, 2006.

[30] *Criteria for Programme Accreditation,* C. o. H. E. (CHE), 1998.

[31] N. Shepstone, "Placing the Curriculum in a Feedback Loop," presented at the SE4 Biennial Conference, 1991.

[32] *White Paper for Post-School Education and Training,* D. o. H. E. a. T. (DHET), November 2013.

[33] *The South African Schools Act (Act 84 of 1996),* D. o. Education, 1996.

[34] Republic of South Africa (RSA). "Constitution of the Republic of South Africa, Chapter 2: Bill of Rights (Rights)." https://www.gov.za/documents/constitution/chapter-2-bill-rights#7 (accessed 25th November 2020).

[35] Republic of South Africa (RSA). "Constitution of the Republic of South Africa, Chapter 2: Bill of Rights (Education)." Republic of South Africa (RSA). https://www.gov.za/documents/constitution/chapter-2-bill-rights#29 (accessed 25th November 2020).

[36] *National Qualifications Framework (Act 67 of 2008),* D. o. H. E. a. Training, June 2009.

[37] D. Cooper and G. Subotzky, "The skewed revolution: Trends in South African Higher Education, 1988-1998.," 2001.

[38] A. Fataar, "Higher Education Policy Discourse in South Africa: A Struggle for Alignment with Macro Development Policy," *South African Journal of Higher Education.,* vol. 17, no. 2, pp. 31 - 39, 2003.

[39] R. du Toit and J. Roodt, *Engineers in a developing country: the profession and education of engineering professionals in South Africa*. Cape Town: HSRC Press, 2009.

[40] A. Lawless, "Numbers & Needs : addressing imbalances in the civil engineering profession," 2005.

[41] Boston College - Lynch School of Education. "The Trends in International Maths and Science Study (TIMSS)." Boston College. http://timssandpirls.bc.edu/timss2015/international-results/timss2015/mathematics/student-achievement/distribution-of-mathematics-achievement (accessed 30th November 2020).

[42] National Commission on Higher Education (South Africa), *NCHE Discussion Document: a Framework for Transformation*. Pretoria: Republic of South Africa (RSA): The Commission (1996), 1996.

[43] *Green Paper on Higher Education Transformation,* D. o. E. (DOE), 1996.

[44] *White Paper 3 - A Programme for Higher Education Transformation,* D. o. E. (DOE), 1997.

[45] *National Education Policy (Act 27 of 1996),* D. o. E. (DOE), 1996.

[46] A. Kraak, "The South African Context: Planning and Legislation in Higher Education," HSRC, Johannesburg, 2001.

[47] *The Articulation Policy for the Post-School Education and Training System of South Africa (PSET)*, 13th January 2017.

[48] *General and Further Education and Training Quality Assurance Act (Act 58 of 2001),* D. o. B. Education, 2001.

[49] *Higher Education Act (Act 101 of 1997),* D. o. H. E. a. T. (DHET), 1997.

[50] *The Skills Development Act (Act 97 of 1998),* D. o. L. (DOL), 1998.

[51] H. Lotz-Sisitka, "Close-out Report for SAQA-Rhodes Universdity Reseach Partnership for Learning Pathways Research," 2015.

[52] R. Blom, "Articulation in the South African Education and Training System," July 2013, doi: DOI: 10.13140/RG.2.2.21748.14720.

[53] Austrlia. (2013). *Australian Qualifications Framework*.

[54] J. Brennan, "Flexible Learning Pathways in British Higher Education: A Decentalized and Market-Based System," Quality Assurance Agency for Higher Education, Gloucester, United Kingdom, 2021.

[55] *E-23-P Enginneering Qualifications in the Higher Education Qualifications,* ECSA, 2019.

[56] *Engineering Professions Act (Act 46 of 2000),* D. o. P. W. a. Infrastructure, 2000.

[57] R. B. Burns, *Introduction to research methods.*, 4 ed. London: Routledge, 1997.

[58] Z. O'Leary, Sage, Ed. *The essential guide to doing research.* London: Sage, 2004.

[59] R. C. Bogdan and S. K. Biklin, *Qualitative research for education: An introduction to theory and methods.*, 3 ed. Boston: Allyn and Bacon, 1998.

[60] D. Mertens, *Research methods in education and psychology: Integrating divesity with quantitative and qualitative approaches.*, 2 ed. Thousand Oaks: Sage, 2005.

[61] N. Mackenzie and S. Knipe, "Research dilemmas: Paradigms, methods and methodology. Issues in Education..." *Acadenia: Issues in Educational Research,* vol. 16, no. 2, 2006.

[62] J. W. Creswell, *Research design: Qualitative, quantitaive, and mixed methods approaches.*, 2 ed. Thousand Oaks: Sage, 2003.

[63] T. Cook and D. Campbell, *Quasi-experimentation: design and analysis issues for field settings.* Boston: Houghton Mifflin, 1979.

[64] J. McMillan and S. Schumacher, *Research in Education: Evidence-based Inquirey*, 6 ed. Cape Town: Pearson, 2006.

[65] R. N. Prince, V. Frith, S. Steyn, and A. F. Cliff, "Academic and Quantitative Literacy in Higher Education: Relationship with Cognate School-Leaving Subjects.," *South African Journal of Higher Education.,* vol. 35, no. 3, pp. 163 - 181, 2021.

[66] M. Rijmenam. "A Short History of Big Data." https://datafloq.com/read/big-data-history/239#:~:text=90%25%20of%20the%20available%20data,O'Reilly%20Media%20in%202005 (accessed 30th October 2020).

[67] Cloudmoyo. "Where does it come from?" https://www.cloudmoyo.com/blog/data-architecture/what-is-big-data-and-where-it-comes-from/ (accessed 25th March 2021).

[68] S. Kaufhold and J. Kaufhold, *Basic Statistics for Educational Research*, 2 ed. Bloomington: iUniverse LLC, 2013.

[69] D. J. Hand and N. M. Adams, "Data Mining," in *Wiley Stats Ref: Statistics Reference Online*, 2015, pp. 1-7.

[70] I. Witten, E. Frank, and M. Hall, *Data Mining - Practical Machine Learning Tools and Techniques.* Amsterdam: Morgan Kaufmann, 2016.

[71] N. Altinok, N. Angrist, and H. Patrinos, "Global Data Set on Education Quality (1965 - 2015)," in "World Bank Policy Research Working Paper 8314," World Bank, Washington, USA, 2018.

[72] The Tec Series, *Digital Teaching Platforms: Customizing Classroom Learning for Each Student.* New York & London: Teachers College Press, 2012.

[73] T. Taodzera, B. Twala, and J. Carroll, "Predicting Engineering Student Success Using Machine Learning," 2018.

[74] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "The KDD process for extracting useful knowledge from volumes of data," *Communications of the ACM,* vol. 39, no. 11, pp. 27 - 34, 1996. [Online]. Available: https://dl.acm.org/doi/10.1145/240455.240464.

[75] T. Taodzera, "Predicting Student Performance Using Machine Learning Analytics," Masters in Engineering, Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg, 2018.

[76] A. M. Shahari, W. Husain, and N. A. Rashid, "A review on Predicting Students performance using Data Mining Techniques," presented at the The third Information Systems International Conference, Malaysia, 2015.

[77] D. C. Montgomery, *Design and Analysis of Experiments.* John Wiley & Sons, 2019.

[78] P. Ayush. "Introduction to Logistic Regression." Towards Data Science. https://towardsdatascience.com/introduction-to-logistic-regression-66248243c148 (accessed 1st February 2021).

[79] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. Waltham, USA: Morgan Kaufmann Publishers, 2011.

[80] N. Avinash. "Decision Tree Classification in Python." https://www.datacamp.com/community/tutorials/decision-tree-classification-python (accessed 5th February 2021).

[81] F. Pedregosa *et al.* "Decision Trees." https://scikit-learn.org/stable/modules/tree.html (accessed 30th October 2020).

[82] S. Sayad, *Real Time Data Mining*. Canada: Self-Help Publishers, 2017.

[83] M. Campbell-Kelly and W. Aspray, *Computer: A History of the Information Machine*. New York: Basic Books, 1996.

[84] L. Gerritsen, "Predicting student performance with Neural Networks " Master of Science in Data Science: Business and Governance, School of Humanities, Tilburg University, May 2017.

[85] J. E. Dayhoff, *Neural Network Architechtures – An Introduction*. New York: Van Nostrarand Reinhold, 1990.

[86] D. Rumelhart, B. Widrow, and M. Lehr. (1994, March 1994) The Basic Ideas in Neural Networks. *Communications of the ACM*. 87 - 92.

[87] N. King. "Introducing Deep Learning and Neural Networks - Deep Learning for Rookies (1)." https://towardsdatascience.com/introducing-deep-learning-and-neural-networks-deep-learning-for-rookies-1-bd68f9cf5883 (accessed 1st June 2021).

[88] N. King. "Multi-Layer Neural Networks with Sigmoid Function - Deep Learning for Rookies (2)." https://towardsdatascience.com/multi-layer-neural-networks-with-sigmoid-function-deep-learning-for-rookies-2-bf464f09eb7f (accessed 1st June 2021).

[89] L. Tarassenko, *A Guide to Neural Computing Applications. HHodding Headline Group.* London: Butterworth-Heinemann, 1998.

[90] J. Pauck, "Neural Network Modelling and Prediction of the Flotation Deinking Behaviour of Complex Recycled Paper Mixes.," UKZN, Durban, 2011.

[91] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1999.

[92] M. Hajek, "Neural Networks Lecture Notes Department of Computer Science," UKZN, 2008.

[93] G. Siemens and R. Baker, "Learning analytics and educational data mining: Towards communication and collaboration.," *ACM International Conference Proceeding Series,* 2012, doi: 10.1145/2330601.2330661.

[94] N. Ndou, R. Ajjodha, and A. Jadhav, "Educational Data-mining to Determine Student Success at Higher Education Instutions.," *IMITEC Conference 2020,* 2020.

[95] L. Mashiloane, "Educational Data Mining (EDM) in a South African University: A Longitudinal Study of Factors that Affect the Academic Performance of Computer Science I students.," MSC, Faculty of Science, University of Witwatersrand, Johannesburg, 2015.

[96] A. J. Fynn and A. Adamiak, "A Comparison of the Utility of Data Mining Algorithms in An Open Distance Learning Context.," *South African Journal of Higher Education.,* vol. 32, no. 4, 2018.

[97] M. Masethe and H. Masethe, "Prediction of Work Integrated Learning Placement Using Data Learning Algorithms.," presented at the World Congress on Engineering and Computer Science, San Francisco, 2014.

[98] W. Vambe, "Using Data Mining Techniques for the Prediction of Student Dropouts from University Science Programmes," in *IIMC International Information Management Corporatiopn - IST-Africa 2017*, Africa, 2017.

[99] L. Masangu, A. Jadhav, and R. Ajoodha, "Predicting Student Academic Performance Using Data Mining Techniques.," *ASTES Journal,* vol. 6, no. 1, pp. 153 - 163, 2021.

[100] E. McGivney and R. Winthrop, "Education's Impact on Economic Growth and Productivity.," July 2016.

[101] A. Adewale, A. Bamidele, and U. Lateef, "Predictive Modelling and Analysis of Academic Performance of Secondary School Students: Artificial Neural Network Approach," *International Journal of Computer Science and Information Security (IJCSIS),,* vol. 9, no. 1, pp. 1 - 8, 2018.

[102] B. A. Kalejaye, O. Folorunso, and O. L. Usman, "Predicting Students' Grade Scores Using Training Functions of Artificial Neural Network," *Journal of Natural Sciences, Engineering and Technology,* vol. 14, 2015.

[103] G. Holmes, A. Donkin, and I. Witten, "WEKA: A Machine Learning Workbench," in *Second Australia and New Zealand Conference on Intelligent Information Systems*, Brisbane, Australia, 1994.

[104] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten, "The WEKA Data Mining Software: An Update.," *SIGKDD Explorations,* vol. 11, no. 1, 2009.

[105] E. Frank, M. Hall, and I. Witten, *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques*. New Zealand: Morgan Kaufmann, 2016.

[106] B. R. Graham, "Analysis of Student Succcess in Mechanical Engineering at The Durban University of Technology," Masters of Engineering Degree, Department of Mechanical Engineering, The Durban University of Technology, Durban, South Africa, August 2015.

[107] R. Nisbet, J. Elsder, and G. Miner, *Handbook of Statistical Analysis & Data Mining Applications*. Burlington (USA), 2009.

[108] J. Rodgers and A. Nicewander, "Thirteen Ways to Look at the Correlation Coefficient.," *The American Statistician,* vol. 42, no. 1, pp. 59-66, 1988. [Online]. Available: https://www.stat.berkeley.edu/~rabbee/correlation.pdf.

[109] Department of Basic Education (DBE), "Curiculum and Assessement Policy Statement Grades 10 – 12 Mathematics," Government Printers, Pretoria Republic of South Africa (RSA), 2011.

[110] *Syllabus for Mathematics N4 National Certificate (TVET),* D. o. E. (DOE), 1996.

[111] V. Reddy, "Mathematics and Science achievement at South African Schools in TIMSS 2003 " HSRC Press, Cape Town, 2006. [Online]. Available: https://www.hsrcpress.ac.za/books/mathematics-and-science-achievement-at-south-african-schools-in-timss-2003

[112] *Syllabus for N4 Engineering Science National Certicate (TVET),* D. o. E. (DOE), 1996.

Annexure 1 – TVET N4 Results 2013 –2017

TVET N4 Results 2013 - 2017

	English (B)	English (A)	English Combined	Science	Mathematics		N4 Communication	N4 Science	N4 IE	N4 Mathematics	N4 ET	Projects I (0.1) PRJT101	Electrical Engineering I (0.1) ELEN101	Electronics I (0.1) ETRS101	Mechanics I (0.1) MECH101	Communication Skills I (0.05) CSKL101	Computer Skills I (0.05) CSCC101	Mathematics I (0.1) MATH101	Progression > 0.4 Credits		Progressed
1	4		4	3	2			6	6	5	8	68	29	30	27	38	50	22	0.15	0	No
2	5		5	3	2			5	7	6	5	40	31	40	31	61	59	46	0.10	0	No
3	5		5	2	3			4	6	5	5	60	41	42	41	61	56	45	0.20	0	No
4		5	5	3	3			5	6	6	8	0	36	46	29	75	89	63	0.20	0	No
5		4	4	3	4			5	5	7	5	20	30	27	12	62	54	52	0.20	0	No
6	5		5	4	4			5	7	6	6	42	34	27	33	65	50	54	0.20	0	No
7	3		3	1	2			5	8	6	8	68	39	30	23	38	63	52	0.25	0	No
8	6		6	3	3			5	5	7	5	68	40	42	35	69	77	50	0.30	0	No
9	4		4	4	5			5	4	5	7	75		32	38	41	71		0.15	0	No
10	4		4	3	3			5	6	5	7	42	33	38	30	45	61	57	0.15	0	No
11		6	6					5	5	6	6	55	42	37	34	71	56	39	0.20	0	No
12	5		5	2	3			7	7	6	6	65	41	38		56	58	45	0.20	0	No
13	4		4	3	4			6	6	7	6	68	39	45		61	72	17	0.20	0	No
14	4		4	3	3			5	6	5	6	42	45	41		59	52	50	0.20	0	No
15		4	4	2	3			5	7	6	6	0	2	0		16	8	4	0.00	0	No
16	4		4	7	7			4	6	8	5	70	68	55	45	47	46	63	0.40	1	Yes
17		4	4		3			5	6	6	8	72	50	58	47	62	75	65	0.50	1	Yes
18		4	4	3	1			6	5	5	6	58	65	57	51	61	64	73	0.60	1	Yes
19		4	4	2	2			5	6	5	8	68	78	63	56	83	79	61	0.60	1	Yes
20		4	4	1	2			5	7	6	7	70	50	46	48	70	60	69	0.40	1	Yes
21	4		4	3	3			6	5	8	6	68	64	26	36	60	77	50	0.40	1	Yes
22	6		6	5	1			6	7	7	9	72	69	50	46	78	88	75	0.50	1	Yes
23	5		5	3	4			7	7	7	8	68	52	50	39	61	77	67	0.50	1	Yes
24		4	4	1	2			6	6	7	7	75	61	55	45	61	70	61	0.50	1	Yes
25	4		4	3	4			6	7	8	8	78	73	58	51	66	52	71	0.60	1	Yes
26		5	5	6	5			5	5	5	5	72	56	59	52	65	88	70	0.60	1	Yes
27	4		4	2	2			6	5	5	6	42	58	52	33	61	59	55	0.40	1	Yes
28	5		5	2	2			5	5	7	5	68	50	50	36	65	58	38	0.40	1	Yes
29	6		6	3	3			5	5	8	7	68	63	55	39	75	93	82	0.50	1	Yes
30	4		4	2	1			6	6	6	8	60	45	52		62	82	50	0.40	1	Yes
31	5		5		1			5	6	5	8	60	71	65		59	71	78	0.50	1	Yes
32	4		4					5	7	6	7	75	55	69		64	89	50	0.50	1	Yes
33	4		4	4	4			5	5	5	7	65	68	51		62	73	68	0.50	1	Yes
34	4		4	2	1			6	6	6	8	60	45	52		62	82	50	0.40	1	Yes
35	5		5		1			5	6	5	8	60	71	65		59	71	78	0.50	1	Yes
36	4		4					5	7	6	7	75	55	69		64	89	50	0.50	1	Yes
37	4		4	4	4			5	5	5	7	65	68	51		62	73	68	0.50	1	Yes
38	5		5		3			5	5	5	6	70	58	58		61	78	55	0.50	1	Yes
39	5		5	5	3			7	6	7	5	70	59	47		59	58	70	0.40	1	Yes
40	5		5	3	3			5	5	7	6	70	50	58		52	63	58	0.50	1	Yes
41	5		5	1	2			5	7	6	8	65	58	52		61	66	58	0.50	1	Yes
42	4		4	4	3			5	7	7	7	65	50	52		58	64	64	0.50	1	Yes
43	5		5	3	5			5	5	5	6	70	62	52		58	84	71	0.50	1	Yes
44	4		4	5	4			7	4	6	5	65	77	65		63	89	84	0.50	1	Yes
45		5	5	5	4			7	6	7	8	70	72	73		68	92	84	0.50	1	Yes
46		5	5	4	7			5	8	7	8	42	58	59		75	84	58	0.40	1	Yes
47		4	4		4			6	6	7	7	60	77	68		65	81	75	0.50	1	Yes
48		4	4					4	4	5	6	75	68	70		75	87	57	0.50	1	Yes
49	4		4	3	5			5	5	5	8	72	78	62		62	84	59	0.50	1	Yes
50		4	4	2	4			5	5	6	7	55	62	51		54	88	60	0.50	1	Yes
51		4	4	2	5			5	5	7	4	72	80	65		61	89	84	0.50	1	Yes
52	4		4	3	2			6	6	6	5	70	68	69		55	90	68	0.50	1	Yes
53	4		4	2	2			5	6	6	7	72	66	63		65	75	72	0.50	1	Yes
54	4		4	3	4			6	7	7	7	64	41	45		61	59	52	0.30	1	Yes
55	4		4	3	5			6	6	5	6	56	29	31		68	84	66	0.30	1	Yes
56	4		4					6	6	6	6	62	46	57		60	89	72	0.40	1	Yes
57	4		4	3	3			6	6	7	5	52	35	51		63	87	52	0.40	1	Yes

58	5		5	3	2			7	8	5	7	68	48	52		56	70	76	0.40	1	Yes
59		4	4	1	3			6	7	6	7	64	58	62		57	88	75	0.50	1	Yes
60	5		5	2	4			6	8	8	7	50	58	51		66	67	82	0.50	1	Yes
61	4		4	3	1			5	6	5	6	72	53	56		56	82	64	0.50	1	Yes
62	5		5	1	3			7	6	8	7	56	50	54		70	89	90	0.50	1	Yes
63	4		4	3	3			5	5	6	6	50	71	61		63	85	84	0.50	1	Yes
64	4		4	4	2			6	8	6	6	68	64	68		60	91	75	0.50	1	Yes
65		4	4	2	2			5	6	6	7	80	58	70		68	91	86	0.50	1	Yes
66	5		5	4	3			5	6	5	6	58	56	55		53	78	75	0.50	1	Yes
67	5		5	2	1			5	6	6	8	62	63	55		53	71	65	0.50	1	Yes
68	5			2	1		4	6	7	7	8										
69	5			2	1		4	6	7	7	8										
70							6	7	8	8	6										
71	3			1	2			5	5	5	5	60	40	42	23	53	63	32	0.20		
72	4							5	5	5	5	75	55		37	50	77		0.30		
73	5			2	3			5	5	5	5	70	50	50	32	55	73		0.40		
74	4			2	2			5	5	5	5	72	70	65	41	51	54	60	0.50		
75		4		3	3			5	5	5	5	58	50	50	38	63	81	57	0.50		
76	5			4	3		5	5	7	8	8	70	84	82	75	58	64	50	0.60		
77	4			3	4			5	5	5	5	50	25	4	24	60	68	50	0.30		
78		4		3	3			5	5	5	5	55	45	46	38	68	76	50	0.30		
79	5			2	3			5	5	5	5	70	63	56	43	65	70	47	0.40		
80	4			5	4			5	5	7	8	72	75	66	45	69	71	80	0.50		
81		5		3	4			5	5	5	5	72	89	75	80	73	93	92	0.60		
82	4			1	2			5	5	5	5	75	53	37	36	50	81	58	0.40		
83		5		2	3			5	5	5	5	60	50	43	30	60	81	65	0.40		
84	4			3	2			5	5	5	5	68	55	37	33	59	53	41	0.30		
85	4			4	3			5	5	5	5	0	43	22	25	46	35	32	0.00		
86		4			2		6	6	6	6	5	50	35	26	21	50	60	36	0.20		
87	6			5	3			5	5	5	5	42	43	45	34	63	88	68	0.20		
88		4		6	5			4	6	5	5	55	50	37	38	50	50	50	0.40		
89		5		2	2			5	5	5	5	75	53	56	38	57	65	62	0.50		
90	5			3	4			5	5	5	5	42	38	42	29	62	85	45	0.10		
91	5			4	4			5	5	5	5	42	40	27	31	28	61	50	0.15		
92	5			3	3			5	5	5	5	55	46	47	37	47	61	47	0.15		
93	4			3	3			5	5	5	5	70	43	35		51	73	73	0.30		
94	5			2	2			5	5	5	5	70	71	46		45	72	51	0.35		
95	5			2	4			5	5	5	5	65	53	40		51	80	72	0.40		
96		5		3	3			5	5	5	5	42	70	50		61	75	59	0.40		
97	5			2	3			5	5	5	5	75	70	58		67	86	63	0.50		
98	5			2	3			5	5	5	5	72	52	58		57	82	67	0.50		
99	4			3	4			5	5	5	5	55	68	51		54	77	50	0.50		
100	5			2	2			5	5	5	5	52	54	65		59	70	64	0.50		
101	5			5	4			6	6	6	7	70	62	67		56	81	72	0.50		
102	4			3	3			5	5	5	5	70	43	35		51	73	73	0.30		
103	5			2	2			5	5	5	5	70	71	46		45	72	51	0.35		
104	5			2	4			5	5	5	5	65	53	40		51	80	72	0.40		
105		5		3	3			5	5	5	5	42	70	50		61	75	59	0.40		
106	5			2	3			5	5	5	5	75	70	58		67	86	63	0.50		
107	5			2	3			5	5	5	5	72	52	58		57	82	67	0.50		
108	4			3	4			5	5	5	5	55	68	51		54	77	50	0.50		
109	5			2	2			5	5	5	5	52	54	65		59	70	64	0.50		
110	4			3	1			5	5	5	5	50	45	28		57	56	41	0.20		
111		5		2	2			5	5	5	5	50	43	25		59	63	42	0.20		
112		6		4	3			5	5	5	5	60	46	30		81	86	41	0.20		
113	4			1	2			5	5	5	5	60	45	31		47	47	50	0.20		
114	5			3	2			5	5	5	5	70	47	54		65	72	63	0.40		
115		5		2	4			5	5	5	5	65	50	57		61	67	56	0.50		
116		4		3	2			5	5	5	5	72	59	52		66	87	61	0.50		
117	4			3	2			5	5	5	5	68	58	56		67	60	64	0.50		
118	6			2	1			5	6	6	5	42	48	21		60	68	36	0.10		
119	5			2	2			5	5	5	5	55	48	54		63	85	46	0.30		
120		6		2	4			5	6	5	5	75	70	50		75	64	50	0.50		
121		6		3	4			5	5	5	5	75	50	50		68	69	50	0.50		
122		5		3	4			5	5	5	5	60	71	69		73	73	62	0.50		
123		4		4	3			5	5	5	5	72	73	70		76	87	65	0.50		
124		4			1			5	5	5	5	75	75	67		70	95	61	0.50		
125		4		3	3			5	5	5	5	70	52	59		59	83	52	0.50		
126	5			2	2			5	5	5	5	60	61	58		69	63	54	0.50		
127	3			2	1							0	8	20				18	0.00		
128		6		6	6							0	13	0					0.00		
129		4		4	6							88	66	66					0.30		
130	4			3	6			6	7	6	7	60	47	45		61	83	82	0.30		
131	4			2	2			5	5	5	5	52	37	51		54	82	67	0.40		
132	4			5	4			5	5	5	5	54	62	63		58	84	63	0.50		
133	5			3	3			5	5	5	5	68	55	60		70	86	65	0.50		
134	4			3	2			6	6	6	6	54	50	60		52	80	61	0.50		

135		4		3	4			5	6	6	6	68	64	68		79	94	51	0.50		
136	4			3	4			5	5	5	5	52	55	58		53	84	63	0.50		
137	4				2			7	7	7	8	64	54	64		58	90	70	0.50		
138	6			4	6								54	46				46	0.10		
139	4			4	4							0	54	31				42	0.10		
140	6			2	2		5	5	5	7	6	55	40	46	31	52	91	60	0.30		
141	5			2	5			5	5	5	6	72	67	51	39	60	71	68	0.50		
142	4			4	4		7	7	6	7	8	68	64	52	30	65	83	63	0.50		
143	5			5	4			6	6	6	7	70	62	67		56	81	72	0.50		
144	6			6	7			7	7	8	7	60	72	66		77	91	87	0.50		
145	4			2	3			5	6	5	6	50	64	61		57	72	68	0.50		
146		4		2	4		4	5	5	7	6	75	71	59	29	59	66	50	0.50	1	
147		5		2	1			7	7	8	8	68	76	70	61	71	76	50	0.60	1	
148	3			2	3		5	5	5	6	5	42	36	26	25	50	77	57	0.20	1	
149	5			4	3		5	5	6	7	6	72	50	43	37	50	84	55	0.40	1	
150	4			3	4			5	5	4	9	72	52	50	32	47	81	53	0.45	1	
151	4			2	4			6	7	5	5	68	72	45	50	53	78	50	0.50	1	
152		5		3	3			5	5	7	5	75	50	66	50	72	50	54	0.60	1	
153	4			2	3			6	6	5	6	75	27	36		54	60	50	0.30	1	
154		4		3	4			7	6	6	8	50	33	48		8	45	50	0.20	1	

Annexure 2 – SC/ NSC Results 2013 - 2017

SC/NSC Results 2013 - 2107

	English (B)	English (A)		Combined English	Mathematics	Science			Matric	N4	Projects I (0.1) PRJT101	Electrical Engineering I (0.1) ELEN101	Electronics I (0.1) ETRS101	Mechanics I (0.1) MECH101	Communication Skills I (0.05) CSKL101	Computer Skills I (0.05) CSCC101	Mathematics I (0.1) MATH101	Progression > 0.4 Credits		Progressed
1	4		4	4	5	5			1		40	18	24	44	42	33	34	0.00	0	No
2	5		5	5	6	5			1		65	40	46	28	43	26	34	0.10	0	No
3		6	6	6	4	5			1		0	29	38	37	55	50	36	0.10	0	No
4	7		7	7	7	7			1		0	36	43	25	75	79	37	0.10	0	No
5		6	6	6	5	5			1		40	30	36	27	73	69	48	0.10	0	No
6	4		4	4	4	5			1		68	46	28	39	39	27	46	0.10	0	No
7	5		5	5	5	5			1		70	29	32	23	52	43	33	0.15	0	No
8	4		4	4	4	6			1		78	45	33	38	48	58	40	0.15	0	No
9	5		5	5	5	5			1		72	43	32	30	51	66	46	0.20	0	No
10	7		7	7	4	6			1		0	53	36	39	70	52	47	0.20	0	No
11	5		5	5	7	5			1		52	40	29	38	43	21	52	0.20	0	No
12		4	4	4	4	4			1		55	35	45	31	67	88	48	0.20	0	No
13		6	6	6	5	5			1		65	45	46	40	73	88	33	0.20	0	No
14		5	5	5	5	6			1		40	33	45	24	52	78	55	0.20	0	No
15	5		5	5	6	6			1		50		58	46	51	51		0.30	0	No
16		5	5	5	5	5			1		70	50	47	29	61	65	43	0.30	0	No
17		5	5	5	5	4			1		68	50	38	33	61	63	37	0.30	0	No
18	6		6	6	5	5			1		50	41	53	38	55	57	42	0.30	0	No
19	5		5	5	6	5			1		52	53	53	31	45	40	48	0.30	0	No
20		7	7	7	4	6			1		68	45	51	45	86	79	42	0.30	0	No
21	5		5	5	5	6			1		68	45	39	47	59	58	65	0.30	0	No
22		6	6	6	5	5			1		50	37	37	33	68	63	50	0.30	0	No
23		5	5	5	4	4			1		68	60	41	36	73	80	45	0.30	0	No
24		5	5	5	5	6			1		75	75	36	35	71	80	48	0.30	0	No
25	5		5	5	6	6			1		55	50	40	40	53	60	45	0.30	0	No
26		5	5	5	5	6			1		70	66	47	16	63	80	43	0.30	0	No
27		6	6	6	4	5			1		52	46	48	34	69	87	52	0.30	0	No
28	4		4	4	5	6			1		62	51	38	34	46	53	51	0.35	0	No
29	4		4	4	4	6			1		55	45	52	39	58	39	52	0.35	0	No
30	6		6	6	5	5			1		40	65	52	46	61	30	66	0.35	0	No
31	6		6	6	5	5			1		40	47	37	45	85	83	42	0.10	0	No
32		6	6	6	4	5			1		70	45	13	26			31	0.10	0	No
33	7		7	7	8	8			1		0		66					0.10	0	No
34	5		5	5	5	4			1		0	36	19	15	68	63	28	0.10	0	No
35	5		5	5	6	4			1		72	34	48	40				0.10	0	No
36	5		5	5	5	4			1		60	30	23	21	42	55	17	0.15	0	No
37	5		5	5	4	4			1		60	37	39	35	58	64	37	0.20	0	No
38	5		5	5	6	5			1		0	40	49	28	68	66	56	0.20	0	No
39	5		5	5	6	6			1		55	45	17	43	64	69	47	0.20	0	No
40		6	6	6	5	5			1		50	54	23	37			42	0.20	0	No
41	4		4	4	4	4			1		50	33	31	28	51	60	43	0.20	0	No
42	5		5	5	4	4			1		60	25	7	26	57	42	62	0.25	0	No

43	6		6	6	5	5			1		60	51	43	39	77	89	43	0.30	0	No
44		5	5	5	5	4			1		42	65	45	37	79	63	67	0.30	0	No
45	5		5	5	6	5			1		72	38	38	45	52	83	56	0.30	0	No
46	7		7	7	7	7			1		0	58	51	37			56	0.30	0	No
47		6	6	6	5	4			1		60	35	19	31	72	50	50	0.30	0	No
48		5	5	5	5	4			1		0	48	50	48	69	56	69	0.30	0	No
49	4		4	4	6	5			1		52	45	34	35	46	81	70	0.25	0	No
50	6		6	6	5	5			1		68	38	47	36	73	73	63	0.30	0	No
51	5		5	5	5	5			1		60	50	32	32	61	65	34	0.30	0	No
52	6		6	6	6	6			1		40	25	8	32	45	91	27	0.05	0	No
53	4		4	4	7	5			1		60	45	45	42	68	54	50	0.30	0	No
54	7		7	7	4	5			1		72	42	39	34	71	85	72	0.30	0	No
55	7		7	7	5	5			1		0	25	0	9	31	38	27	0.00	0	No
56	5		5	5	4	5			1		72	32	36	34	51	78	42	0.20	0	No
57	4		4	4	5	4			1		65	51	34	34	57	77	40	0.30	0	No
58		7	7	7	8	7			1		65	37	21	24	63	84	19	0.20	0	No
59		6	6	6	4	5			1		60	45	37	29	71	78	38	0.20	0	No
60	6		6	6	4	5			1		65	46	29	31		55	46	0.15	0	No
61	4		4	4	5	5			1		0	70	28	30	51	69	48	0.20	0	No
62	7		7	7	5	5			1		50		32	32	58	65	19	0.20	0	No
63		7	7	7	5	4			1		40	45	33	29	78	75	45	0.10	0	No
64	6		6	6	4	4			1		40	50	26	29	58	52	36	0.20	0	No
65	6		6	6	6	6			1		0	83	47	45	61	75	69	0.30	0	No
66	6		6	6	5	6			1		65	45	37	28	55	65	54	0.30	0	No
67		5	5	5	5	4			1			62	35	36	56	61	47	0.20	0	No
68		4	4	4	4	4			1		60	43	26	39	65	71	45	0.20	0	No
69		6	6	6	5	4			1		42	60	45	28	57	75	68	0.30	0	No
70		5	5	5	5	5			1		72	53	43	42	55	76	42	0.30	0	No
71	5		5	5	5	5			1		60	62	38	33	52	56	47	0.30	0	No
72	4		4	4	5	5			1		65	38	19	34	47	61	46	0.15	0	No
73		5	5	5	5	5			1		50	68	36	33	62	75	38	0.30	0	No
74	6		6	6	6	5			1		0	57	24	35	54	88	50	0.30	0	No
75	6		6	6	5	6			1		68	54	39	37	59	39	51	0.35	0	No
76	6		6	6	4	5			1		70	68	46	36	62	83	45	0.30	0	No
77	5		5	5	4	5			1		0	38	30	32	70	52	36	0.10	0	No
78	5		5	5	4	4			1		68	43	24	38	53	70	33	0.20	0	No
79		5	5	5	5	5			1		0	18	11	22	0	28	30	0.00	0	No
80		6	6	6	6	5			1		42	60	36	36	53	84	47	0.20	0	No
81		6	6	6	4	5			1		70	52	29	29	78	73	45	0.30	0	No
82		6	6	6	4	4			1		55	62	36	26	56	53	35	0.30	0	No
83		6	6	6	5	5			1		55	35	21	27	61	60	40	0.20	0	No
84	4		4	4	5	4			1		0	9	14	21	11	12	25	0.00	0	No
85	5		5	5	4	4			1		70	44	34	32	42	51	39	0.15	0	No
86	5		5	5	5	6			1		0	6	19	7	16	28	4	0.00	0	No
87		6	6	6	6	4			1		70	66	33	41	73	70	47	0.30	0	No
88	6		6	6	5	4			1		78		55	61				0.30	0	No
89		5	5	5	6	5			1		40	75	41	46	72	55	65	0.30	0	No
90	6		6	6	5	6			1		60	55	25	37	58	66	45	0.30	0	No
91	7		7	7	5	5			1		0	41	22	35	65	60	45	0.10	0	No
92	4		4	4	6	5			1		60		57	42				0.20	0	No
93		5	5	5	6	5			1		70		62					0.20	0	No
94	5		5	5	5	4			1		0	37	20	17	45	78	34	0.05	0	No
95	7		7	7	4	5			1		42	36	27	27	50	68	25	0.10	0	No
96		6	6	6	5	5			1			45	39	28			59	0.10	0	No
97	6		6	6	4	4			1		42	26	39	38	56	65	36	0.10	0	No
98	5		5	5	4	4			1		0	49	35	26	60	73	30	0.10	0	No
99	5		5	5	4	5			1		0	38	42	28	54	71	35	0.10	0	No
100		5	5	5	5	4			1		0	37	45	42	54	82	47	0.10	0	No
101	5		5	5	6	6			1		42	45	25	42	45	50	73	0.15	0	No
102	5		5	5	5	4			1		42	50	46	46	58			0.15	0	No
103	4		4	4	4	4			1		0	43	29	32	45	68	52	0.15	0	No
104	6		6	6	6	6			1		73	43	48	56				0.20	0	No
105	5		5	5	6	4			1		42	37	40	27	55	54	51	0.20	0	No
106	7		7	7	4	5			1		50	45	37	35	54	50	38	0.20	0	No
107	6		6	6	6	5			1		59			56				0.20	0	No
108	5		5	5	5	6			1		42	54	40	32	75	75	46	0.20	0	No
109		6	6	6	4	4			1				53	33			50	0.20	0	No
110	5		5	5	5	4			1		42	24	38	34	55	50	57	0.20	0	No
111	4		4	4	6	4			1		68	36	34	31	61	71	40	0.20	0	No
112	7		7	7	5	6			1		0	48	46	34	66	83	60	0.20	0	No
113	5		5	5	5	4			1		42	45	48	32	62	94	56	0.20	0	No
114	4		4	4	7	5			1			61	61	51				0.30	0	No
115		6	6	6	5	4			1		60	31	31	30	57	73	51	0.30	0	No
116		4	4	4	6	6			1		0	54	45	35	54	65	57	0.30	0	No
117	6		6	6	5	4			1		0	50	47	43	70	56	50	0.30	0	No
118		4	4	4	4	5			1		65	45	45	38	69	88	56	0.30	0	No
119	4		4	4	5	4			1		68	46	35	38	61	76	54	0.30	0	No

120	4		4	4	5	4			1		65	42	42	36	58	73	54	0.30	0	No
121	5		5	5	6	5			1		42	45	55	42	51	69	54	0.30	0	No
122		4	4	4	4	5			1		42	64	42	43	75	85	57	0.30	0	No
123		6	6	6	4	5			1		60	48	45	19	67	82	50	0.30	0	No
124	6		6	6	7	6			1		70	47	55	32	67	78	48	0.30	0	No
125	4		4	4	5	4			1		42	54	35	34	71	81	54	0.30	0	No
126	4		4	4	6	6			1		60	34	32	28	64	75	51	0.30	0	No
127	4		4	4	5	4			1		55	53	13	29	60	86	43	0.30	0	No
128		4	4	4	6	5			1		0	57	45	47	67	84	77	0.30	0	No
129	6		6	6	5	6			1		0	23	17		47	47	45	0.00	0	No
130	5		5	5	8	4			1		0	20	17		37	10	9	0.00	0	No
131	4		4	4	6	5			1		0	12	0		24	28	26	0.00	0	No
132	4		4	4	5	6			1		0	0	0		0	0	0	0.00	0	No
133	7		7	7	5	5			1		42	31	32		63	92	47	0.10	0	No
134	5		5	5	4	5			1		42	38	29		55	85	45	0.10	0	No
135	6		6	6	4	5			1		42	31	42		63	55	31	0.10	0	No
136		5	5	5	4	4			1		42	34	31		53	78	27	0.10	0	No
137		6	6	6	5	6			1		65		45			75		0.15	0	No
138	5		5	5	7	5			1		60	36	28		54	75	36	0.20	0	No
139	5		5	5	6	6			1		75		68					0.20	0	No
140		4	4	4	4	4			1		55	40	31		70	91	29	0.20	0	No
141	5		5	5	6	4			1			36	40		54	51	53	0.20	0	No
142	5		5	5	5	4			1		72	29	45		50	56	45	0.20	0	No
143	6		6	6	5	4			1		52	28	41		55	75	42	0.20	0	No
144	6		6	6	6	6			1		42	39	38		56	64	64	0.20	0	No
145	5		5	5	4	5			1		50	30	44		54	61	43	0.20	0	No
146	5		5	5	5	4			1		72	26	39		61	84	35	0.20	0	No
147	6		6	6	5	5			1		0	30	32		56	77	65	0.20	0	No
148		5	5	5	5	4			1		42	40	34		63	88	53	0.20	0	No
149	4		4	4	5	4			1		0	38	35		53	78	52	0.20	0	No
150	7		7	7	6	7			1		55		58					0.20	0	No
151	5		5	5	5	6			1		0	40	14		34	46	48	0.00	0	No
152	6		6	6	5	5			1		0	7	17		42	66	28	0.05	0	No
153	5		5	5	5	5			1		0	37	28		61	58	24	0.10	0	No
154		5	5	5	5	5			1		52	34	45		59	82	17	0.20	0	No
155		5	5	5	5	5			1		5	33	27		62	71	52	0.20	0	No
156	6		6	6	7	5			1		0	45	38		36	91	45	0.05	0	No
157		6	6	6	6	6			1		0	39	39		72	81	50	0.20	0	No
158	5		5	5	4	5			1		50	46	39		57	50	43	0.20	0	No
159	6		6	6	7	6			1		0	33	45		64	70	68	0.20	0	No
160	4		4	4	4	4			1		0	48	45		54	52	62	0.20	0	No
161		4	4	4	6	4			1		42	45	40		59	69	50	0.20	0	No
162	6		6	6	6	4			1		54	37	45		70	92	39	0.20	0	No
163	4		4	4	4	4			1		52	45	36		51	66	38	0.20	0	No
164		6	6	6	6	5			1		32	27	37		64	96	50	0.20	0	No
165	5		5	5	6	6			1		72	53	50	37	45	37	50	0.40	1	Yes
166	5		5	5	5	5			1		50	52	42	37	55	75	65	0.40	1	Yes
167		5	5	5	4	4			1			52	57		75	64	59	0.40	1	Yes
168	6		6	6	6	5			1		60	46	45	50	66	66	60	0.40	1	Yes
169	5		5	5	5	6			1		60	63	30	45	69	87	82	0.40	1	Yes
170	4		4	4	5	5			1		75	38	55	30	53	63	79	0.40	1	Yes
171		6	6	6	6	6			1		68	65	57	39	81	81	46	0.40	1	Yes
172		6	6	6	7	6			1		55	51	42	46	63	80	70	0.40	1	Yes
173	5		5	5	5	5			1		50	50	34	41	63	59	75	0.40	1	Yes
174		4	4	4	6	4			1		60	57	45	33	59	80	62	0.40	1	Yes
175		5	5	5	5	6			1		60	70	50	37	69	88	40	0.40	1	Yes
176	5		5	5	6	5			1		52	51	38	28	55	56	51	0.40	1	Yes
177	4		4	4	6	4			1		55	59	50	38	47	36	64	0.40	1	Yes
178	6		6	6	5	4			1		70	68	54	45	57	54	39	0.40	1	Yes
179	6		6	6	6	7			1		55	58	45	32	78	76	71	0.40	1	Yes
180	5		5	5	4	6			1		68	58	58	45	52	45	60	0.45	1	Yes
181	4		4	4	5	4			1		68	59	56	47	43	70	52	0.45	1	Yes
182	4		4	4	5	5			1		65	59	53	41	47	55	61	0.45	1	Yes
183	7		7	7	5	5			1		65	58	62	46	47	62	56	0.45	1	Yes
184		6	6	6	6	5			1		40	68	50	65	70	84	69	0.50	1	Yes
185	5		5	5	5	5			1		70	54	60	35	51	59	54	0.50	1	Yes
186		6	6	6	6	6			1		65	56	63	43	75	56	59	0.50	1	Yes
187	6		6	6	6	6			1		40	75	55	58	65	73	65	0.50	1	Yes
188		6	6	6	4	6			1		75	67	67	45	79	70	54	0.50	1	Yes
189	5		5	5	6	5			1		55	53	50	40	51	53	61	0.50	1	Yes
190	4		4	4	4	5			1		50	59	52	38	53	58	54	0.50	1	Yes
191	6		6	6	5	6			1		60	51	51	37	70	51	54	0.50	1	Yes
192	5		5	5	5	4			1		50	56	52	31	61	64	52	0.50	1	Yes
193		5	5	5	4	6			1		58	66	65	46	65	75	65	0.50	1	Yes
194	7		7	7	6	6			1		68	62	52	43	66	56	57	0.50	1	Yes
195	4		4	4	6	5			1		65	75	73	46	61	64	92	0.50	1	Yes
196	4		4	4	6	5			1		65	53	50	34	58	56	57	0.50	1	Yes

197	6		6	6	5	4			1		50	51	59	35	51	52	60	0.50	1	Yes
198	6		6	6	5	5			1		55	64	50	38	61	71	50	0.50	1	Yes
199	6		6	6	5	5			1		65	52	55	35	62	73	50	0.50	1	Yes
200		5	5	5	5	5			1		65	70	58	48	56	61	64	0.50	1	Yes
201	6		6	6	6	4			1		55	75	58	46	53	66	58	0.50	1	Yes
202	4		4	4	7	6			1		60	54	56	43	51	51	76	0.50	1	Yes
203	6		6	6	6	5			1		68	71	53	47	70	80	68	0.50	1	Yes
204	5		5	5	7	7			1		52	77	59	46	58	58	89	0.50	1	Yes
205	6		6	6	6	4			1		70	65	63	43	72	65	50	0.50	1	Yes
206	6		6	6	5	6			1		60	60	51	51	69	43	64	0.55	1	Yes
207	5		5	5	7	7			1		72	72	70	78	50	63	87	0.60	1	Yes
208	4		4	4	5	6			1		70	80	78	68	71	75	69	0.60	1	Yes
209	5		5	5	5	5			1		70	73	70	65	73	81	91	0.60	1	Yes
210	4		4	4	7	4			1		65	66	57	56	71	80	63	0.60	1	Yes
211	5		5	5	7	6			1		60	60	56	52	59	54	68	0.60	1	Yes
212	5		5	5	6	5			1		68	73	56	55	54	76	66	0.60	1	Yes
213	7		7	7	7	6			1		70	71	67	63	78	84	95	0.60	1	Yes
214	7		7	7	5	4			1		70	50	53	51	67	56	62	0.60	1	Yes
215	5		5	5	5	6			1		60	68	61	57	66	70	87	0.60	1	Yes
216	6		6	6	5	5			1		55	75	66	61	75	82	58	0.60	1	Yes
217	6		6	6	6	6			1		60	65	52	58	69	63	73	0.60	1	Yes
218	7		7	7	7	7			1		70	91	67	64	71	85	95	0.60	1	Yes
219		5	5	5	5	5			1		55	69	61	51	57	67	63	0.60	1	Yes
220	4		4	4	5	6			1		65	75	56	56	50	80	51	0.60	1	Yes
221		5	5	5	6	5			1		55	75	54	54	70	71	63	0.60	1	Yes
222		5	5	5	6	6			1		55	71	55	54	72	77	72	0.60	1	Yes
223	5		5	5	5	5			1		55	67	60	51	70	69	81	0.60	1	Yes
224	6		6	6	6	7			1		60	66	61	81	62	57	73	0.60	1	Yes
225	5		5	5	5	4			1		50	64	36	38	62	81	77	0.40	1	Yes
226	5		5	5	5	4			1		70	56	45	38	57	78	54	0.40	1	Yes
227		5	5	5	5	6			1		60	59	62	54				0.40	1	Yes
228	4		4	4	5	5			1		72	58	37	45	55	59	62	0.40	1	Yes
229	4		4	4	4	4			1		78	43	50	37	72	87	64	0.40	1	Yes
230	6		6	6	7	4			1		50	58	53	39	83	86	59	0.50	1	Yes
231		5	5	5	4	5			1		50	59	50	51	73	65	47	0.50	1	Yes
232	7		7	7	4	4			1		62	64	58	28	71	76	63	0.50	1	Yes
233	4		4	4	5	5			1		75	69	49	53	57	66	57	0.50	1	Yes
234	4		4	4	6	5			1		75	61	57	43	56	77	68	0.50	1	Yes
235		5	5	5	5	4			1		70	56	45	53	60	57	58	0.50	1	Yes
236		4	4	4	5	4			1		50	51	55	45	77	89	61	0.50	1	Yes
237	5		5	5	5	4			1		70	51	52	14	64	55	64	0.50	1	Yes
238	6		6	6	5	5			1		68	68	51	38	81	93	68	0.50	1	Yes
239		6	6	6	6	5			1		58	70	45	52	90	83	73	0.50	1	Yes
240	5		5	5	6	4			1		50	53	55	52	75	63	62	0.60	1	Yes
241	4		4	4	5	5			1		70	63	57	57	61	75	70	0.60	1	Yes
242	6		6	6	7	7			1		70	78	58	69	77	86	92	0.60	1	Yes
243	6		6	6	6	6			1		70	72	69	61	72	81	77	0.60	1	Yes
244	4		4	4	7	7			1		68	73	50	67	57	64	90	0.60	1	Yes
245	4		4	4	5	5			1		68	65	53	47	39	65	64	0.45	1	Yes
246	6		6	6	5	5			1		75	75	58	59	65	85	87	0.60	1	Yes
247	4		4	4	6	4			1		60	52	35	45	60	85	70	0.40	1	Yes
248	6		6	6	6	7			1		78	23	61	62	68		69	0.45	1	Yes
249	4		4	4	6	6			1		75	64	51	38	56	45	67	0.45	1	Yes
250	6		6	6	5	5			1		75	65	50	40	63	61	38	0.40	1	Yes
251	4		4	4	6	5			1		60	70	34	50	42	72	75	0.45	1	Yes
252	6		6	6	7	5			1		65	81	40	57	51	85	80	0.50	1	Yes
253		6	6	6	7	6			1		60	79	39	54	70	91	71	0.50	1	Yes
254	7		7	7	5	5			1		60	84	35	41	80	87	68	0.40	1	Yes
255	5		5	5	6	6			1		65	87	62	70	71	80	88	0.60	1	Yes
256		6	6	6	6	5			1		60	75	39	55	59	87	56	0.50	1	Yes
257	6		6	6	6	6			1		75	51	56	39	61	68	53	0.50	1	Yes
258		5	5	5	6	6			1		75	73	42	41	72	89	79	0.40	1	Yes
259	4		4	4	5	5			1		0	58	54	55	59	65	56	0.50	1	Yes
260		5	5	5	6	5			1		60	84	64	76	64	86	92	0.60	1	Yes
261	6		6	6	4	4			1		65	57	24	28	62	77	50	0.40	1	Yes
262		6	6	6	5	5			1		50	64	39	37	80	83	51	0.40	1	Yes
263	6		6	6	6	6			1		65	78	58	59	66	88	70	0.60	1	Yes
264	6		6	6	6	7			1		68	79	56	51	61	79	77	0.60	1	Yes
265	4		4	4	5	5			1		60	61	38	57	51	57	51	0.50	1	Yes
266	7		7	7	7	7			1		70	96	69	54	72	79	100	0.60	1	Yes
267		5	5	5	5	4			1		68	50	31	34	72	71	53	0.40	1	Yes
268		4	4	4	5	5			1		42	66	37	51	67	79	62	0.40	1	Yes
269	5		5	5	6	6			1		70	72	40	52	66	54	50	0.50	1	Yes
270		7	7	7	7	7			1		62	73	29	53	73	63	51	0.50	1	Yes
271	6		6	6	6	6			1		65	66	41	38	51	62	56	0.40	1	Yes
272		7	7	7	5	6			1		60	73	36	38	70	75	50	0.40	1	Yes
273		7	7	7	6	7			1		55	67	31	38	70	81	62	0.40	1	Yes

274	5		5	5	7	7			1		55	83	38	61	67	83	84	0.50	1	Yes
275	5		5	5	6	4			1		70	77	55	61	50	78	78	0.60	1	Yes
276	6		6	6	6	5			1		70	78	53	45	62	86	68	0.50	1	Yes
277	4		4	4	6	7			1		70	75	59	45	50	79	60	0.50	1	Yes
278		5	5	5	6	5			1		50	59	52	31	50	59	59	0.50	1	Yes
279	5		5	5	6	6			1		65	56	51	50	59	55	67	0.60	1	Yes
280		6	6	6	6	5			1		65	78	47	39	75	77	51	0.40	1	Yes
281	6		6	6	6	5			1		80	83	54	57	75	72	82	0.60	1	Yes
282	7		7	7	5	4			1		70	71	36	45	73	78	67	0.40	1	Yes
283		6	6	6	7	7			1		42	86	46	62	71	91	81	0.40	1	Yes
284		6	6	6	6	5			1		65	61	53	26	63	72	58	0.50	1	Yes
285	5		5	5	6	5			1		70	56	48	35	54	63	52	0.40	1	Yes
286		4	4	4	5	5			1		60	75	42	47	77	63	60	0.40	1	Yes
287	6		6	6	4	6			1		50	67	45	43	66	72	64	0.40	1	Yes
288	5		5	5	6	5			1		70	80	50	48	76	90	62	0.50	1	Yes
289	4		4	4	7	5			1		42	94	70	59	61	75	89	0.50	1	Yes
290		6	6	6	5	5			1		68	72	51	36	67	68	53	0.50	1	Yes
291		5	5	5	6	5			1		55	71	50	50	63	61	58	0.60	1	Yes
292	6		6	6	5	5			1		42	62	35	52	62	81	50	0.40	1	Yes
293		6	6	6	6	6			1		50	79	56	46	87	83	69	0.50	1	Yes
294	5		5	5	6	5			1		55	63	36	46	59	52	55	0.40	1	Yes
295		6	6	6	7	5			1		68	72	57	51	73	81	72	0.60	1	Yes
296	6		6	6	5	5			1		55	85	41	43	55	79	70	0.40	1	Yes
297	5		5	5	5	5			1		55	78	50	58	61	73	81	0.60	1	Yes
298	5		5	5	7	4			1		65	67	16	48	50	76	67	0.40	1	Yes
299	4		4	4	5	5			1		55	69	44	62	47	62	64	0.45	1	Yes
300	6		6	6	6	6			1		65	79	32	58	63	71	59	0.50	1	Yes
301	4		4	4	6	5			1		60	67	39	39	60	60	65	0.40	1	Yes
302	4		4	4	7	4			1		65	83	50	64	51	69	62	0.60	1	Yes
303	5		5	5	6	7			1		68	55	38	50	59	64	84	0.50	1	Yes
304	5		5	5	5	4			1		68	64	38	41	50	66	61	0.40	1	Yes
305	5		5	5	6	4			1		70	53	38	30	53	64	50	0.40	1	Yes
306		4	4	4	6	5			1		70	54	34	36	66	77	57	0.40	1	Yes
307		6	6	6	6	6			1		78	66	69	45	85	91	42	0.40	1	Yes
308	7		7	7	4	4			1		70	61	40	39	68	90	60	0.40	1	Yes
309	6		6	6	7	7			1		50	45	10	50	67	84	84	0.40	1	Yes
310		4	4	4	6	4			1		50	58	34	40	55	50	73	0.40	1	Yes
311	4		4	4	6	5			1		50	52	41	30	53	77	64	0.40	1	Yes
312	6		6	6	5	4			1		68	50	36	37	60	84	63	0.40	1	Yes
313	4		4	4	4	4			1		65	36	51	38	52	81	50	0.40	1	Yes
314	4		4	4	4	4			1		50	60	46	37	50	57	52	0.40	1	Yes
315		6	6	6	4	4			1		70	66	62	43	62	88	37	0.40	1	Yes
316	4		4	4	6	4			1		72	54	48	37	63	71	68	0.40	1	Yes
317	6		6	6	5	5			1		60	57	53	39	59	85	39	0.40	1	Yes
318	4		4	4	5	4			1		72	58	45	39	64	71	51	0.40	1	Yes
319	5		5	5	5	5			1		42	58	51	45	66	84	57	0.40	1	Yes
320	4		4	4	5	4			1		58	58	47	31	64	77	61	0.40	1	Yes
321	4		4	4	5	4			1		60	55	30	40	54	76	62	0.40	1	Yes
322	5		5	5	5	4			1		42	56	59	37	61	72	64	0.40	1	Yes
323		4	4	4	6	4			1		60	56	46	38	85	87	79	0.40	1	Yes
324		6	6	6	6	6			1		70	58	42	33	75	98	58	0.40	1	Yes
325	6		6	6	6	5			1		55	50	43	37	66	80	70	0.40	1	Yes
326		5	5	5	6	5			1		75	86	67	46	73	84	93	0.50	1	Yes
327	6		6	6	4	4			1		70	53	51	32	75	87	52	0.50	1	Yes
328		6	6	6	4	6			1		68	64	58	40	69	98	60	0.50	1	Yes
329	5		5	5	5	5			1		60	50	58	39	77	81	50	0.50	1	Yes
330		5	5	5	5	5			1		68	50	50	43	58	89	70	0.50	1	Yes
331	5		5	5	6	4			1		55	61	59	39	57	88	57	0.50	1	Yes
332	7		7	7	4	5			1		68	59	52	32	83	89	58	0.50	1	Yes
333	6		6	6	4	4			1		70	60	61	38	68	81	55	0.50	1	Yes
334	4		4	4	4	4			1		68	62	43	59	50	68	56	0.50	1	Yes
335	5		5	5	5	5			1		70	55	55	39	63	83	72	0.50	1	Yes
336	5		5	5	6	5			1		68	69	59	46	50	89	85	0.50	1	Yes
337	7		7	7	6	5			1		75	70	77	56	86	88	62	0.60	1	Yes
338	4		4	4	6	5			1		50	31	27		52	56	50	0.30	1	Yes
339	5		5	5	5	4			1		0	47	55		66	88	68	0.30	1	Yes
340	6		6	6	6	5			1		70	54	45		55	62	47	0.30	1	Yes
341	5		5	5	5	6			1		55	33	43		50	61	59	0.30	1	Yes
342		5	5	5	5	6			1		55	43	39		58	72	56	0.30	1	Yes
343	7		7	7	7	5			1		42	43	58		71	92	84	0.30	1	Yes
344		5	5	5	6	6			1		42	50	40		52	80	58	0.30	1	Yes
345	4		4	4	7	7			1		50	63	56					0.30	1	Yes
346		6	6	6	4	5			1		60	50	45		72	79	61	0.40	1	Yes
347	5		5	5	6	6			1		42	76	58		61	64	92	0.40	1	Yes
348		5	5	5	5	4			1		50	42	50		64	90	76	0.40	1	Yes
349		5	5	5	5	4			1		60	50	45		63	89	56	0.40	1	Yes
350	4		4	4	5	4			1		52	61	45		51	70	54	0.40	1	Yes

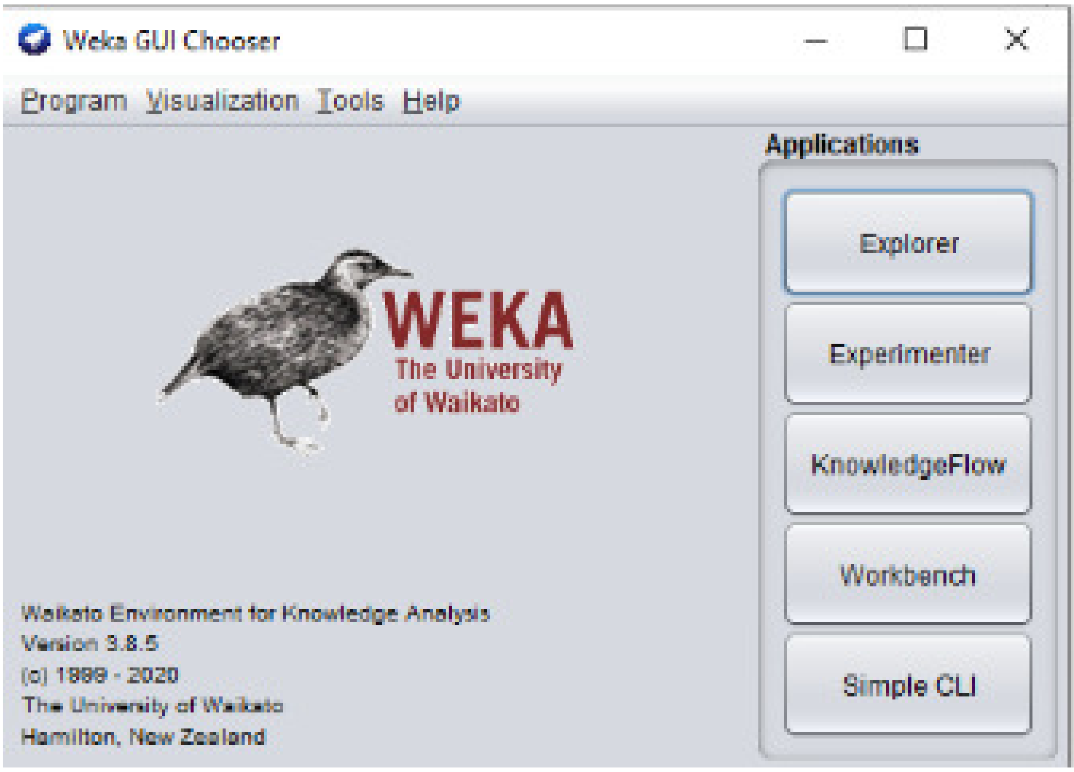
351	5		5	5	5	5			1		70	59	42		54	75	65	0.40	1	Yes
352	6		6	6	5	4			1		68	41	52		66	70	57	0.40	1	Yes
353		5	5	5	6	5			1		60	71	41		51	85	57	0.40	1	Yes
354	5		5	5	5	5			1		50	40	50		60	53	59	0.40	1	Yes
355		5	5	5	5	4			1		75	46	51		58	91	62	0.40	1	Yes
356	6		6	6	5	4			1		72	45	50		65	84	59	0.40	1	Yes
357	6		6	6	4	5			1		72	55	30		62	67	60	0.40	1	Yes
358	6		6	6	6	4			1		65	43	55		65	83	63	0.40	1	Yes
359		5	5	5	4	5			1		70	71	59		68	88	45	0.40	1	Yes
360	6		6	6	7	5			1		72	57	55		50	86		0.40	1	Yes
361	4		4	4	6	5			1		68	71	50		57	91	84	0.50	1	Yes
362	5		5	5	5	4			1		72	71	61		58	61	92	0.50	1	Yes
363		6	6	6	6	4			1		65	57	52		68	55	53	0.50	1	Yes
364	5		5	5	5	4			1		75	64	57		54	84	66	0.50	1	Yes
365	7		7	7	5	6			1		50	67	60		68	77	87	0.50	1	Yes
366	5		5	5	7	7			1		68	64	68		53	84	79	0.50	1	Yes
367	4		4	4	6	6			1		68	75	55		66	62	75	0.50	1	Yes
368	6		6	6	5	6			1		60	63	60		63	76	64	0.50	1	Yes
369	5		5	5	6	5			1		68	51	58		62	80	76	0.50	1	Yes
370	7		7	7	5	5			1		70	89	73		70	93	78	0.50	1	Yes
371	6		6	6	5	5			1		75	79	57		59	95	80	0.50	1	Yes
372	4		4	4	4	5			1		68	83	50		56	75	58	0.50	1	Yes
373	7		7	7	6	7			1		68	80	58		71	93	75	0.50	1	Yes
374		4	4	4	7	5			1		60	67	52		63	81	68	0.50	1	Yes
375	5		5	5	7	7			1		68	67	58		61	80	80	0.50	1	Yes
376		4	4	4	5	4			1		68	52	57		60	92	71	0.50	1	Yes
377	4		4	4	5	5			1		72	56	63		63	54	56	0.50	1	Yes
378	5		5	5	4	4			1		50	60	54		62	85	62	0.50	1	Yes
379	5		5	5	6	6			1		65	68	65		68	83	73	0.50	1	Yes
380	6		6	6	6	5			1		55	60	57		60	61	69	0.50	1	Yes
381	5		5	5	6	5			1		75	75	63		59	94	87	0.50	1	Yes
382	7		7	7	4	5			1		70	54	63		65	89	64	0.50	1	Yes
383	5		5	5	7	7			1		70	84	78		57	90	92	0.50	1	Yes
384	6		6	6	6	5			1		70	58	52		62	86	79	0.50	1	Yes
385	5		5	5	5	6			1		70	59	56		62	82	75	0.50	1	Yes
386		5	5	5	6	5			1		60	66	57		56	89	76	0.50	1	Yes
387	5		5	5	6	7			1		75	64	64		60	91	75	0.50	1	Yes
388	5		5	5	5	5			1		75	68	55		54	70	82	0.50	1	Yes
389		4	4	4	5	5			1		70	64	60		57	79	87	0.50	1	Yes
390	6		6	6	4	6			1		75	63	59		75	93	71	0.50	1	Yes
391		6	6	6	7	6			1		68	82	76		61	95	93	0.50	1	Yes
392	4		4	4	6	5			1		68	67	58		63	71	87	0.50	1	Yes
393	5		5	5	5	7			1		72	66	66		58	85	81	0.50	1	Yes
394		4	4	4	6	5			1		50	50	51		50	86	56	0.50	1	Yes
395	4		4	4	5	5			1		68	81	63		61	83	80	0.50	1	Yes
396	5		5	5	5	6			1		70	57	52		51	85	59	0.50	1	Yes
397	5		5	5	6	5			1		65	67	54		59	78	83	0.50	1	Yes
398	6		6	6	6	7			1		60	54	70		59	87	84	0.50	1	Yes
399		6	6	6	7	7			1		42	59	54		47	79	76	0.35	1	Yes
400		6	6	6	6	5			1		68	61	52		65	89	46	0.40	1	Yes
401	6		6	6	6	6			1		60	56	42		67	79	54	0.40	1	Yes
402	5		5	5	7	5			1		50	52	33		59	65	58	0.40	1	Yes
403		6	6	6	5	5			1		60	50	50		58	82	52	0.50	1	Yes
404	5		5	5	5	5			1		72	51	50		60	69	64	0.50	1	Yes
405	5		5	5	5	5			1		65	64	50		61	62	65	0.50	1	Yes
406	4		4	4	5	5			1		70	63	55		50	68	65	0.50	1	Yes
407	6		6	6	4	5			1		50	72	56		76	78	70	0.50	1	Yes
408	5		5	5	5	5			1		50	69	62		68	82	70	0.50	1	Yes
409		5	5	5	5	5			1		50	75	69		72	78	72	0.50	1	Yes
410		6	6	6	4	4			1		70	89	76		75	94	77	0.50	1	Yes
411	6		6	6	7	7			1		72	75	79		64	89	80	0.50	1	Yes
412	6		6	6	6	7			1		50	75	60		73	78	81	0.50	1	Yes
413	5		5	5	7	6			1		68	75	65		53	85	81	0.50	1	Yes
414	6		6	6	6	5			1		70	67	50		66	81	92	0.50	1	Yes
415		5	5	5	5	5			1		55	43	47		68	86	62	0.30	1	Yes
416	4		4	4	6	5			1		70	34	45		50	59	58	0.30	1	Yes
417	5		5	5	5	5			1		50	48	31		51	67	61	0.30	1	Yes
418	5		5	5	6	6			1		0	43	51		60	56	54	0.30	1	Yes
419	6		6	6	5	5			1		42	56	40		70	66	59	0.30	1	Yes
420	4		4	4	6	5			1		60	45	31		50	58	75	0.30	1	Yes
421	5		5	5	4	5			1		58	48	45		56	67	50	0.30	1	Yes
422	4		4	4	4	5			1		75	55	45		57	83	43	0.30	1	Yes
423	6		6	6	5	4			1		70	68	38		64	72	64	0.40	1	Yes
424	4		4	4	5	5			1		50	64	36		54	64	77	0.40	1	Yes
425	4		4	4	4	5			1		65	60	45		55	53	50	0.40	1	Yes
426	5		5	5	5	5			1		42	58	51		67	90	63	0.40	1	Yes
427	6		6	6	7	6			1		42	86	67		65	56	87	0.40	1	Yes

428		6	6	6	5	6			1		0	68	69		79	90	64	0.40	1	Yes
429	5		5	5	6	5			1		68	70	40		63	84	68	0.40	1	Yes
430		6	6	6	5	5			1		75	38	55		62	94	50	0.40	1	Yes
431	4		4	4	4	4			1		55	40	53		70	57	56	0.40	1	Yes
432	5		5	5	6	4			1		72	45	60		62	76	71	0.40	1	Yes
433	4		4	4	5	4			1		55	58	46		53	66	59	0.40	1	Yes
434	5		5	5	4	4			1		65	54	48		56	64	55	0.40	1	Yes
435	5		5	5	6	6			1		42	54	50		58	68	65	0.40	1	Yes
436	4		4	4	6	5			1		42	55	50		53	59	75	0.40	1	Yes
437	5		5	5	5	6			1		55	47	65		59	71	56	0.40	1	Yes
438	6		6	6	5	4			1		55	71	51		66	52	50	0.50	1	Yes
439		5	5	5	5	4			1		55	63	54		62	80	50	0.50	1	Yes
440	5		5	5	6	6			1		78	75	63		66	70	62	0.50	1	Yes
441	6		6	6	5	5			1		55	64	59		66	93	77	0.50	1	Yes
442	5		5	5	5	5			1		75	68	50		56	75	59	0.50	1	Yes
443		5	5	5	6	6			1		75	75	65		60	89	67	0.50	1	Yes
444	6		6	6	6	6			1		60	75	54		72	87	69	0.50	1	Yes
445		5	5	5	5	6			1		72	80	68		75	85	72	0.50	1	Yes
446	5		5	5	6	7			1		72	76	66		65	89	67	0.50	1	Yes
447	6		6	6	7	4			1		68	64	62		68	61	68	0.50	1	Yes
448	4		4	4	5	4			1		65	63	57		61	78	58	0.50	1	Yes
449		5	5	5	5	4			1		65	51	53		75	82	51	0.50	1	Yes
450	6		6	6	5	6			1		55	72	64		68	93	67	0.50	1	Yes
451		6	6	6	6	6			1		70	68	59		65	87	73	0.50	1	Yes
452	5		5	5	6	6			1		60	69	63		55	85	75	0.50	1	Yes
453		6	6	6	5	5			1		60	63	56		75	75	54	0.50	1	Yes
454	5		5	5	5	6			1		72	58	60		71	72	62	0.50	1	Yes
455	5		5	5	5	4			1		70	68	55		65	94	72	0.50	1	Yes
456		5	5	5	4	4			1		60	71	62		63	67	59	0.50	1	Yes
457	5		5	5	6	5			1		68	83	67		70	79	72	0.50	1	Yes
458		4	4	4	5	6			1		72	81	69		61	80	72	0.50	1	Yes
459		4	4	4	5	5			1		50	76	70		75	79	78	0.50	1	Yes
460	6		6	6	5	5			1		65	62	64		63	92	67	0.50	1	Yes
461	5		5	5	5	6			1		60	73	53		67	78	65	0.50	1	Yes
462	6		6	6	7	6			1		65	63	56		67	72	75	0.50	1	Yes
463	4		4	4	5	5			1		70	59	66		73	72	64	0.50	1	Yes
464	4		4	4	4	5			1		70	70	64		67	61	58	0.50	1	Yes
465	4		4	4	6	5			1		68	75	67		65	83	75	0.50	1	Yes
466	4		4	4	5	4			1		68	56	56		57	62	50	0.50	1	Yes
467	5		5	5	6	6			1		68	72	59		60	72	78	0.50	1	Yes
468		5	5	5	5	5			1		75	85	76		71	96	85	0.50	1	Yes
469	4		4	4	5	5			1		50	62	58		67	75	64	0.50	1	Yes
470	6		6	6	7	7			1		70	86	71		63	88	79	0.50	1	Yes
471	4		4	4	6	7			1		55	70	62		58	70	77	0.50	1	Yes
472	6		6	6	6	5			1		70	68	63		66	87	62	0.50	1	Yes
473	6		6	6	5	5			1		55	72	71		75	89	70	0.50	1	Yes
474	4		4	4	5	5			1		65	60	58		65	90	52	0.50	1	Yes
475	4		4	4	7	5			1		50	83	62		58	69	87	0.50	1	Yes
476	4		4	4	4	4			1		55	70	54		62	73	58	0.50	1	Yes
477	6		6	6	5	5			1		60	65	55		69	63	67	0.50	1	Yes
478	5		5	5	7	7			1		68	76	73		73	79	94	0.50	1	Yes
479	5		5	5	5	5			1		62	75	58		55	69	63	0.50	1	Yes
480		4	4	4	6	4			1		88	66	66					0.30	1	Yes
481	4		4	4	4	5			1		32	33	50		56	76	63	0.30	1	Yes
482		6	6	6	6	6			1		42	37	54		63	96	54	0.30	1	Yes
483		5	5	5	5	5			1		72	43	60		61	99	45	0.30	1	Yes
484		5	5	5	4	5			1		36	47	61		58	61	60	0.30	1	Yes
485		6	6	6	5	5			1		60	46	50		65	73	40	0.30	1	Yes
486	4		4	4	5	6			1		28	62	27		61	53	81	0.30	1	Yes
487	5		5	5	6	6			1		68	45	52		57	73	72	0.40	1	Yes
488	6		6	6	4	4			1		64	42	50		62	78	62	0.40	1	Yes
489	4		4	4	5	4			1		50	30	52		64	90	53	0.40	1	Yes
490	6		6	6	6	6			1		32	52	69		66	93	92	0.40	1	Yes
491	6		6	6	7	7			1		30	64	69		69	84	72	0.40	1	Yes
492	6		6	6	5	5			1		58	48	59		63	67	69	0.40	1	Yes
493	5		5	5	6	5			1		76	45	54		62	76	82	0.40	1	Yes
494	4		4	4	7	5			1		38	57	62		50	71	76	0.40	1	Yes
495	4		4	4	6	5			1		50	47	56		61	95	68	0.40	1	Yes
496		6	6	6	5	5			1		54	45	64		63	94	58	0.40	1	Yes
497	4		4	4	5	5			1		50	47	50		52	82	68	0.40	1	Yes
498	5		5	5	5	5			1		50	33	50		67	93	63	0.40	1	Yes
499		4	4	4	5	4			1		58	43	55		65	92	56	0.40	1	Yes
500	5		5	5	5	5			1		38	52	54		62	77	75	0.40	1	Yes
501	4		4	4	6	4			1		52	53	42		57	88	82	0.40	1	Yes
502	5		5	5	6	6			1		72	77	69			75	73	0.45	1	Yes
503	5		5	5	6	5			1		54	53	59		62	94	80	0.50	1	Yes
504	5		5	5	5	6			1		62	54	53		63	71	50	0.50	1	Yes

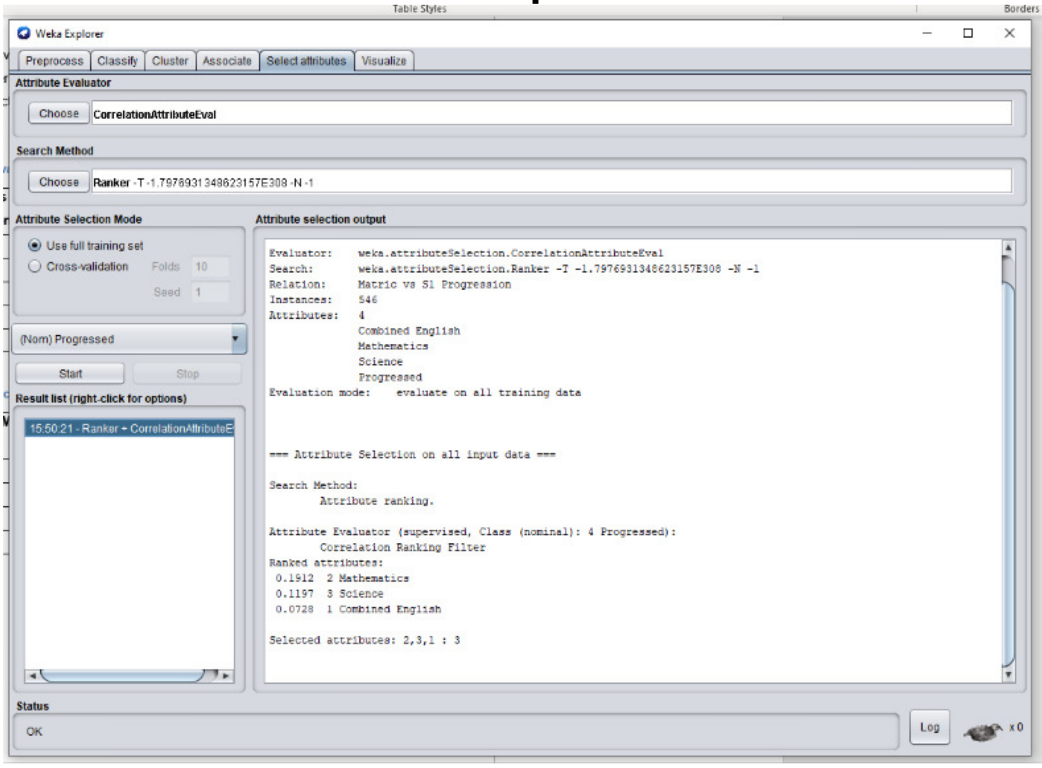
505	5		5	5	7	6			1		68	67	70		60	71	83	0.50	1	Yes
506	4		4	4	5	5			1		62	65	59		60	71	62	0.50	1	Yes
507	5		5	5	8	4			1		60	51	57		60	91	90	0.50	1	Yes
508	4		4	4	7	5			1		68	66	65		62	68	85	0.50	1	Yes
509	5		5	5	6	4			1		72	67	71		80	97	91	0.50	1	Yes
510	6		6	6	6	5			1		50	58	55		72	83	94	0.50	1	Yes
511	5		5	5	6	4			1		64	64	55		69	72	75	0.50	1	Yes
512	6		6	6	6	6			1		68	77	70		60	84	90	0.50	1	Yes
513	5		5	5	5	4			1		68	55	53		71	82	66	0.50	1	Yes
514	5		5	5	5	4			1		52	68	72		65	58	79	0.50	1	Yes
515	5		5	5	5	4			1		82	66	76		62	86	89	0.50	1	Yes
516		6	6	6	5	4			1		86	53	61		80	96	53	0.50	1	Yes
517	6		6	6	5	4			1		56	60	57		63	86	75	0.50	1	Yes
518		6	6	6	5	5			1		50	50	58		76	88	67	0.50	1	Yes
519	4		4	4	5	5			1		66	55	66		65	76	75	0.50	1	Yes
520		5	5	5	5	5			1		56	62	63		76	92	69	0.50	1	Yes
521	4		4	4	5	4			1		56	55	53		58	68	83	0.50	1	Yes
522	4		4	4	6	5			1		80	75	63		66	92	89	0.50	1	Yes
523	4		4	4	5	4			1		72	50	62		60	75	69	0.50	1	Yes
524	7		7	7	6	7			1		50	75	71		71	72	98	0.50	1	Yes
525	7		7	7	4	6			1		56	55	53		65	86	66	0.50	1	Yes
526	4		4	4	4	4			1		82	63	56		57	88	72	0.50	1	Yes
527	6		6	6	5	6			1		68	61	61		59	66	73	0.50	1	Yes
528	6		6	6	6	5			1		75	75	77		67	93	88	0.50	1	Yes
529		5	5	5	5	5			1		75	55	66		72	96	75	0.50	1	Yes
530	5		5	5	6	7			1		58	75	64		65	84	78	0.50	1	Yes
531	5		5	5	6	4			1		52	61	62		63	90	89	0.50	1	Yes
532	5		5	5	6	4			1		62	53	50		62	71	76	0.50	1	Yes
533	4		4	4	6	6			1		64	65	61		68	86	90	0.50	1	Yes
534	5		5	5	4	5			1		72	59	56		64	65	63	0.50	1	Yes
535	7		7	7	6	7			1		68	75	69		77	77	87	0.50	1	Yes
536	5		5	5	5	6			1		68	72	57		61	83	69	0.50	1	Yes
537	4		4	4	5	4			1		72	70	75		62	71	83	0.50	1	Yes
538	4		4	4	6	5			1		72	72	69		60	71	84	0.50	1	Yes
539	5		5	5	6	5			1		56	52	62		70	88	63	0.50	1	Yes
540	4		4	4	5	6			1		50	72	66		53	78	77	0.50	1	Yes
541		5	5	5	6	5			1		68	61	60		70	95	66	0.50	1	Yes
542	4		4	4	7	6			1		62	71	61		67	65	73	0.50	1	Yes
543	4		4	4	5	6			1		68	79	71		73	92	82	0.50	1	Yes
544		5	5	5	4	5			1		56	66	64		75	75	67	0.50	1	Yes
545	4		4	4	5	5			1		68	71	60		51	54	78	0.50	1	Yes
546	4		4	4	5	5			1		58	75	69		53	77	86	0.50	1	Yes
547		6			6	6			1		0	13	0					0.00	0	
548	6				6	4			1			54	46				46	0.10	0	
549	4				4	4			1		0	54	31				42	0.10	0	
550	5				7	6			1		40							0.00	0	
551	4				2	4			1		50	50	50		50	50	50	0.50		
552		4			3	5			1		60	57	55		68	84	57	0.50		
553								1	1		75	62	34	58	21	17	61	0.40		
554								1	1		55	52	43	33	51	63	56	0.40		
555								1	1		68	80	84		68	76	94	0.50		
556								1	1		50	57	64		61	86	50	0.50		
557								1	1		80	63	52		63	91	72	0.50		
558								1	1		52	45	52		52	77	76	0.40		
559		4			1			1	1		72	76	59		86	86	75	0.50		
560	3				1	1		1	1		58	54	53	50	47	63	67	0.55		
561								1	1		68	71	53	58	45	79	83	0.55		
562								1	1		75		38	50	50	95	50	0.40		
563								1	1		78	79	68	64	57	50	86	0.60		
564	6				3	4		1	1		75			30	52			0.15		
565								1	1		70	34	22	25	53	58	14	0.20		
566								1	1			50					50	0.20		
567								1	1						50	50	50	0.20		
568								1	1			97	90		72		92	0.35		
569								1	1		55	33	37		57	53	36	0.20		
570	3				1	2		1	1		0	17	30		51	54	23	0.10		
571								1	1			50			50	50	50	0.30		
572								1	1		55	21	52		51	59	50	0.40		
573		5			4			1	1		42	58	69		66	82	70	0.40		
574	4				4	4		1	1		72	77	67		71	88	58	0.50		
575		4			3	2			1		68	47	45		57	60	41	0.20		
576		7			7	7			1		0	50	39	53			50	0.30	1	
577	5				8	4			1		65	47	52	32	70	58	50	0.40	1	
578	5				4	4			1		55	50	50	47	57	50	50	0.50	1	
579	5				7	6			1		60	50	50	50	50	50	50	0.60	1	
580		7			6	6			1		58	51	50	50	50	50	50	0.60	1	
581		5			5	5			1		75	68	76	73	50	87	50	0.60	1	

582	6				6	5			1		70	57	54	50	50	50	50	0.60	1	
583	7				8	8			1		55	50	50	50	50	50	50	0.60	1	
584	4				6	4			1		55	37	36	36	50	50	50	0.30	1	
585	6				6	5			1		75	50	73	50	50	50	50	0.60	1	
586	6				6	5			1		50	50	50	50	50	50	50	0.60	1	
587	6				5	4			1			50			50	50	50	0.30	1	
588	6				6	7			1			50		50		50	50	0.35	1	
589		6			7	7			1			50		50	50		50	0.35	1	
590	6				1	2			1			50	50			50	50	0.35	1	
591	6				5	5			1			50	50		50	50	50	0.40	1	
592	6				4	6			1			50		50	50	50	50	0.40	1	
593		7			6	5			1		50	50	50	50	50	50	50	0.60	1	
594		6			7	5			1		50	50	50	50	50	50	50	0.60	1	
595	5				4	4			1		65	40	36	28	50	50	50	0.30	1	
596		5			6	5			1		60	67	49	42	50	50	50	0.40	1	
597	4				6	4			1		68	55	53	32	60	78	50	0.50	1	
598		5			5	4			1		55	54	50	36	50	50	54	0.50	1	
599	7				7	7			1		50	79	67	50	50	50	50	0.60	1	
600		6			6	6			1		68	75	62	50	50	50	77	0.60	1	
601		6			7	7			1		72					50		0.15	1	
602	5				6	6			1		72		63			50		0.25	1	
603	6				6	6			1		72	67	40		50	50	50	0.40	1	
604		7			6	7			1		58	50	53		50	50	50	0.50	1	
605	6				6	5			1			50			50			0.15	1	
606	5				7	6			1			50			50		50	0.25	1	
607	4				7	7			1					50	50	50	50	0.30	1	
608		6			5	6			1			50		50	50		50	0.35	1	
609	5				6	6			1			50		50	50		50	0.35	1	
610		6			7	7			1			50	50		50		50	0.35	1	
611		6			7	7			1			50		50	50		50	0.35	1	
612	5				6	6			1			50		50	50	50	50	0.40	1	
613		6			5	5			1			50	50		50	50	50	0.40	1	
614		5			5	4			1			50	50		50	50	50	0.40	1	
615		7			6	7			1			50		50	50	50	50	0.40	1	
616		5			5	6			1			50		50	50	50	50	0.40	1	
617		6			6	5			1			50	50		50	50	50	0.40	1	
618		6			6	6			1		50	50	22		50	85	50	0.40	1	
619		6			7	7			1			50	50	50	50		50	0.45	1	
620		7			7	6			1			50	50	50	50	50	50	0.50	1	
621		5			6	6			1		72	65	64		50	89	50	0.50	1	
622		7			6	6			1			50			50	50	50	0.30	1	
623		6			7	7			1		70	67	56		50	78	71	0.50	1	
624		7			6	6			1		75	83	72		50	97	83	0.50	1	
625		7			6	6			1		80	50	86	50	50	50	50	0.60	1	
626	4				7	7			1						50	50	50	0.20	1	
627	6				5	5			1			50	50		50	50	50	0.40	1	
628	5				6	4			1		58	61	70		59	91	50	0.50	1	
629		7			6	5			1		78	50	75		50	50	50	0.50	1	
630		7			7	7			1		56	64	80		86	98	50	0.50	1	
631	5				5	6			1		0	0	0	0	0	0	0	0.00	0	
632		6			4	5			1		0	0	0	0	0	0	0	0.00	0	
633	6				4	5			1		0	0	0	0	0	0	0	0.00	0	
634		5			5	4			1		0	9	5	4	22	7	3	0.00	0	
635		6			4	5			1		0	12	0	4	27	0	7	0.00	0	
636		7			6	4			1		0	0	0		0	0	0	0.00	0	
637	5				7	7			1		0	0	0		0	0	0	0.00	0	
638	5				5	5			1		0	0	0		0	0	0	0.00	0	
639	6				4	4			1		0	11	18		19	10	20	0.00	0	
640	4				4	5			1		0	34	29		40	27	41	0.00	0	
641	4				8	6			1		0	0	0		0	0	0	0.00	0	
642	5				5	6			1		0	25	20		18	0	20	0.00	0	
643	4				7	6			1		0	29	30		20	26	47	0.00	0	
644		7			7	6			1				45					0.00	0	

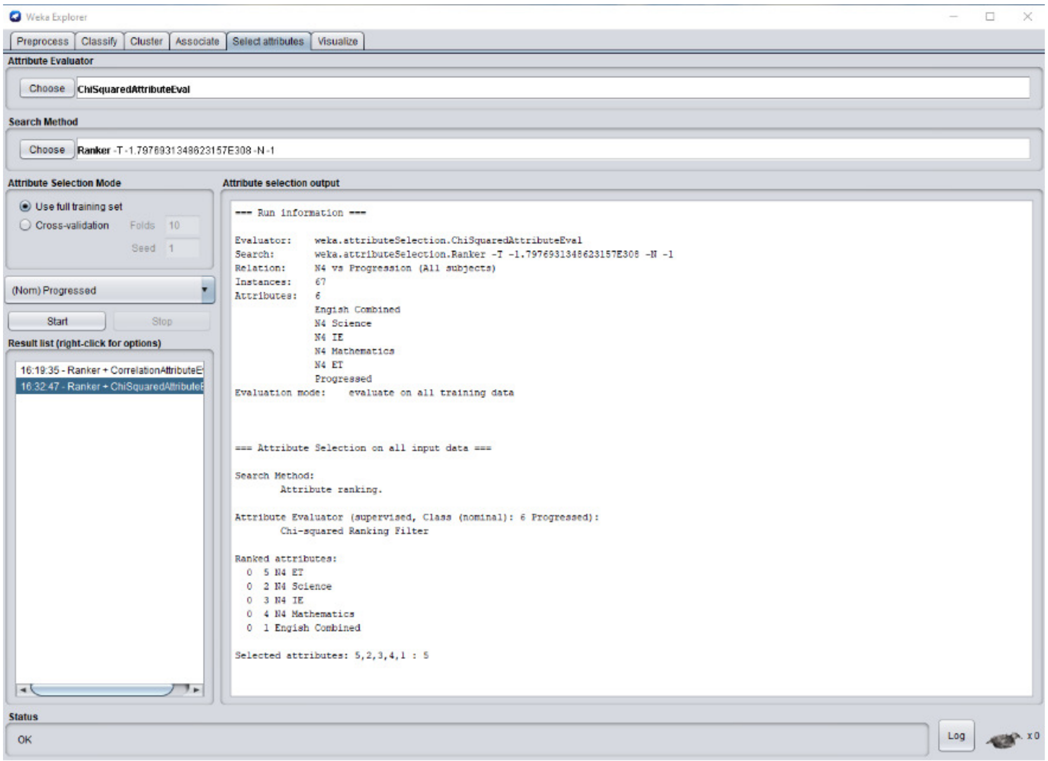
Annexure 3 – WEKA Tool [105]



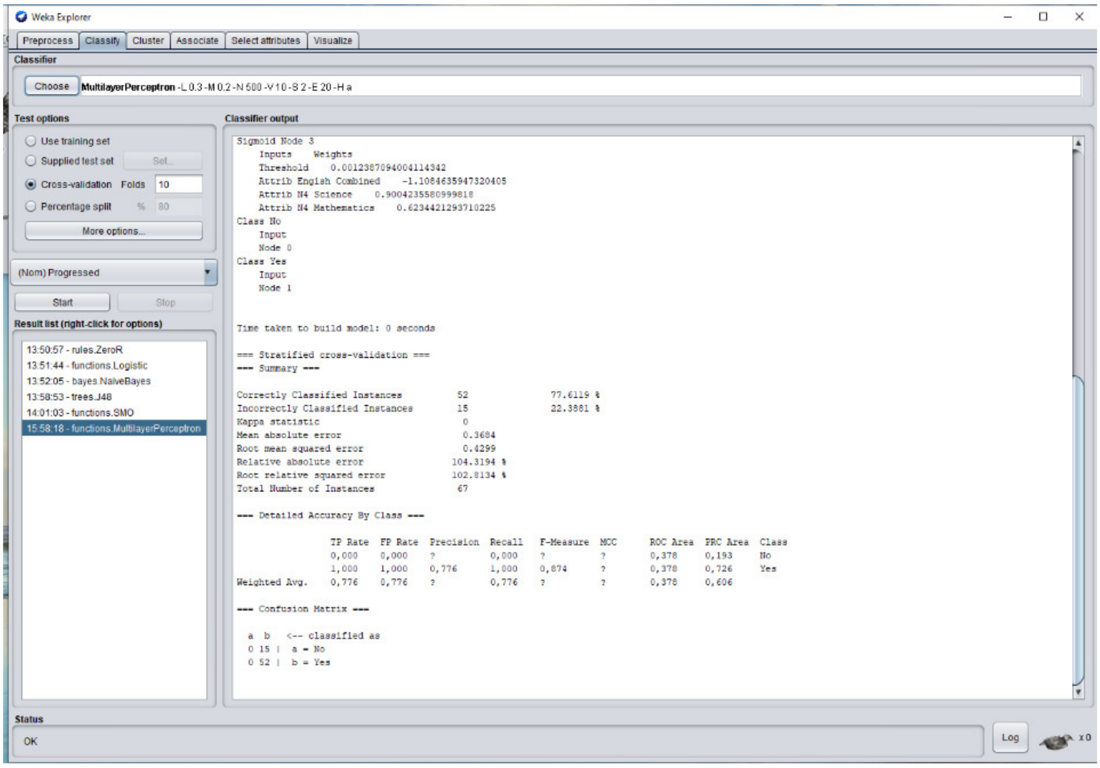
Annexure 4 – WEKA Input Evaluator SC/NSC



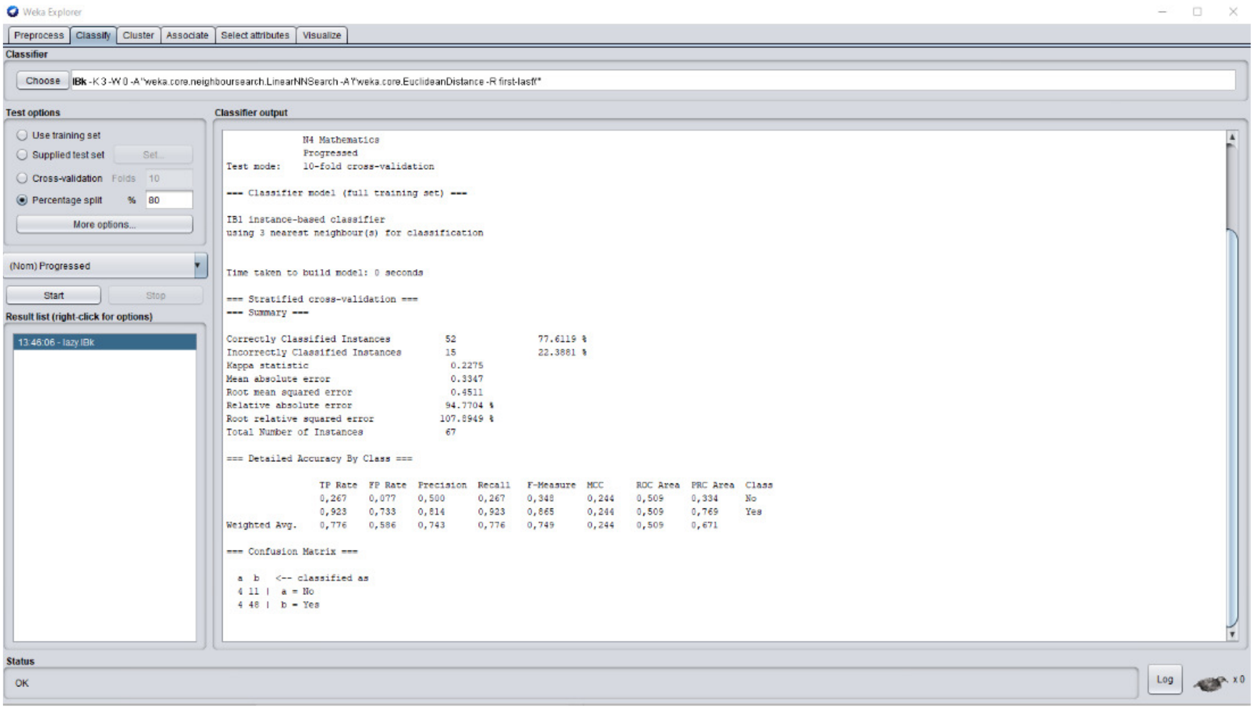
Annexure 5 – WEKA Input Evaluator TVET N4



Annexure 6 – WEKA TVET N4 Classification Models



Annexure 7 – WEKA IBK Classifier



Annexure 8 – WEKA SC/NSC Classifier [104]