

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Figure 1.1: Education System in India	5
Figure 1.2: Research Design	10

CHAPTER 2

INTRODUCTION TO MACHINE LEARNING TECHNIQUES

Figure 2.1: Logistics Regression	17
Figure 2.2: Support Vector Machine	18
Figure 2.3: Decision Tree	20
Figure 2.4: K-Nearest Neighbor	21
Figure 2.5: Random Forest	22
Figure 2.6: Gradient Boosting	23
Figure 2.7: AdaBoost	24
Figure 2.8: Bagged Decision Tree	24

CHAPTER 3

LITERATURE REVIEW

Figure 3.1: Common Factors used for student performance prediction	30
--	----

CHAPTER 4

RESEARCH METHODOLOGY

Figure 4.1: CRISP-DM Process	45
Figure 4.2: Methodology of Research	47
Figure 4.3: The Data Collection Process	50
Figure 4.4: Dependency Relation with Demographic, Socio-economic and Academic Variable	52
Figure 4.5: Data Collected for student analysis based on attributes	53
Figure 4.6: Dataset collected for analysis based on boards	54
Figure 4.7: Dataset collected for analysis of student performance	54
Figure 4.8: Exploratory Data Analysis	57
Figure 4.9: Feature Selection Process	58
Figure 4.10: Filter Method	60

CHAPTER 5

PREDICTION OF STUDENT POTENTIAL IN TECHNICAL PROGRAMS

Figure 5.1:	KNN ROC Curve	74
Figure 5.2:	SVM ROC Curve	75
Figure 5.3:	Naïve Bayes ROC Curve	76
Figure 5.4:	Decision Tree ROC Curve	77
Figure 5.5:	Random Forest ROC Curve	78
Figure 5.6:	Gradient Boosting ROC Curve	79
Figure 5.7:	Bar Graph for Comparison Analysis of Classifier Methods	82
Figure 5.8:	Comparison Analysis of Classification Methods with Filter Methods	90
Figure 5.9:	Gender Ratio	91
Figure 5.10:	Medium of Education	91
Figure 5.11:	Family Income	92
Figure 5.12:	Students in technical vs non-technical programs	92
Figure 5.13:	Comparison for RBSE Dataset	93
Figure 5.14:	Comparison for CBSE Dataset	94
Figure 5.15:	Graphical Comparison for RBSE dataset based on Ensemble Methods	95
Figure 5.16:	Graphical Comparison for CBSE dataset based on Ensemble Methods	96

CHAPTER 6

PREDICTION OF STUDENT PERFORMANCE

Figure 6.1:	Representation of Result Categories	98
Figure 6.2:	Bar chart of Performance for Students dataset with two class labels	103
Figure 6.3:	Bar chart of Performance for three class labels	104
Figure 6.4:	Performance Analysis of Students with five class labels	105

LIST OF TABLES

CHAPTER 4

FRAMEWORK FOR RESEARCH

Table 4.1:	Original attributes in Student Dataset	49
Table 4.2:	Confusion Matrix	69

CHAPTER 5

PREDICTION OF STUDENT POTENTIAL IN TECHNICAL PROGRAMS

Table 5.1:	KNN Confusion Matrix	73
Table 5.2:	KNN result evaluation for analysis of student potential	73
Table 5.3:	SVM Confusion Matrix	74
Table 5.4:	SVM result evaluation for analysis of student potential	75
Table 5.5:	NB Confusion Matrix	76
Table 5.6:	NB result evaluation for analysis of student potential	76
Table 5.7:	Decision Tree Confusion Matrix	77
Table 5.8:	DT result evaluation for analysis of student	77
Table 5.9:	RF Confusion Matrix	78
Table 5.10:	RF result evaluation for analysis of student potential	78
Table 5.11:	GB Confusion Matrix	79
Table 5.12:	GB result evaluation for analysis of student potential	79
Table 5.13:	Correctly and Incorrectly Classified Records	80
Table 5.14:	Comparison based on training dataset for analysis of student potential	80
Table 5.15:	Comparison based on training dataset for analysis of student potential	81
Table 5.16:	Feature ranked by MIG	84
Table 5.17:	Comparing Predictive Accuracy of Six classifiers with MIG filter	85
Table 5.18:	Feature ranked by ANOVA	86
Table 5.19:	Comparing Predictive Accuracy of Six classifiers with ANOVA filter	86
Table 5.20:	Feature ranked by ROC	87

Table 5.21:	Comparing Predictive Accuracy of Six classifiers with ROC filter	87
Table 5.22:	Feature ranked by CHI	88
Table 5.23:	Comparing Predictive Accuracy of Six classifiers with CHI filter	88
Table 5.24:	Classification methods results with Filter Methods	89
Table 5.25:	Accuracy comparison for RBSE Dataset	93
Table 5.26:	Accuracy comparison for CBSE Dataset	94
Table 5.27:	Accuracy comparison for RBSE Dataset based on Ensemble Methods	95
Table 5.28:	Accuracy comparison for CBSE Dataset based on Ensemble Methods	95

CHAPTER 6

PREDICTION OF STUDENT PERFORMANCE

Table 6.1:	Result Categories	99
Table 6.2:	LR result evaluation for analysis of student performance	100
Table 6.3:	SVM result evaluation for analysis of student performance	100
Table 6.4:	AdaBoost result evaluation for analysis of student performance	101
Table 6.5:	DT result evaluation for analysis of student performance	101
Table 6.6:	RF result evaluation for analysis of student performance	102
Table 6.7:	GB result evaluation for analysis of student performance	102
Table 6.8:	Performance Analysis of Students dataset with two class labels	103
Table 6.9:	Performance Analysis of Students dataset with three class labels	104
Table 6.10:	Performance Analysis of Students dataset with five class labels	104