PROACTIVE STUDENT SUPPORT AND NEXT BEST ACTION

Group 25

CLIENT

The University of Melbourne

Prof. Kate Smith Miles
&
Mr. Milad Chenaghlou

TEAM

Milad Chenaghlou
Supervisor

Sudheer Kumar
1051717

Haopeng Yan
962332

Xiaowen Jin
1023499

Yiming Xu
863672

Harshal Shah
1020849

GROUP 25

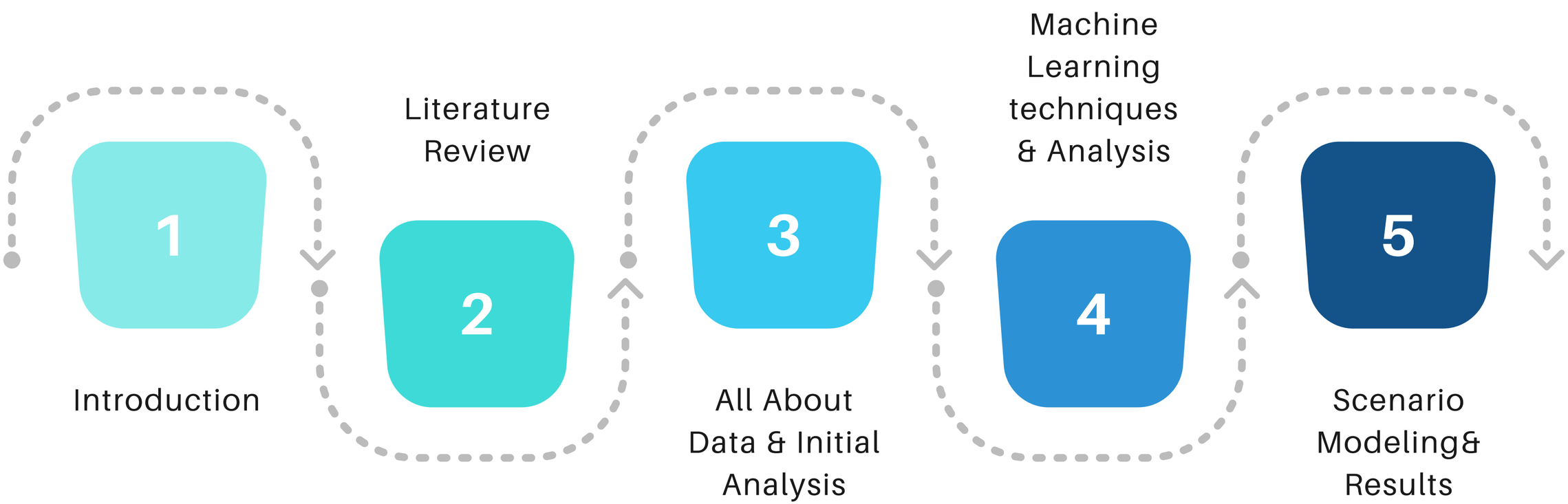
Overview of Roles

Sudheer Kumar Kolla	Data Extrapolation	Regression Models	Scenario Modeling
Haopeng Yan	Initial Data Analysis	Classification For Demographic Data	Scenario Modeling Story Line
Xiaowen Jin	Classification For Demographic Data	Hyperparamater Tuning & Results For Demographic Data	Linear Discriminant Analysis On Demographic Data
Yiming Xu	Initial Data Analysis	Clustering Research	Scenario Modeling Story Line
Harshal Shah	Classification & Regression Models (For Entire Dataset)	Hyperparameter Tuning & Results For Entire Dataset	Linear Discriminant Analysis On Entire Dataset

All the team members contributed equally towards report and presentation

PROACTIVE STUDENT SUPPORT AND NEXT BEST ACTION

Project Flow



INTRODUCTION

CLIENT

The University of Melbourne

DOMAIN

Student Success Prediction & Recommendation

GOAL

Empowering the educational institutions to provide a better support to their students

SCOPE

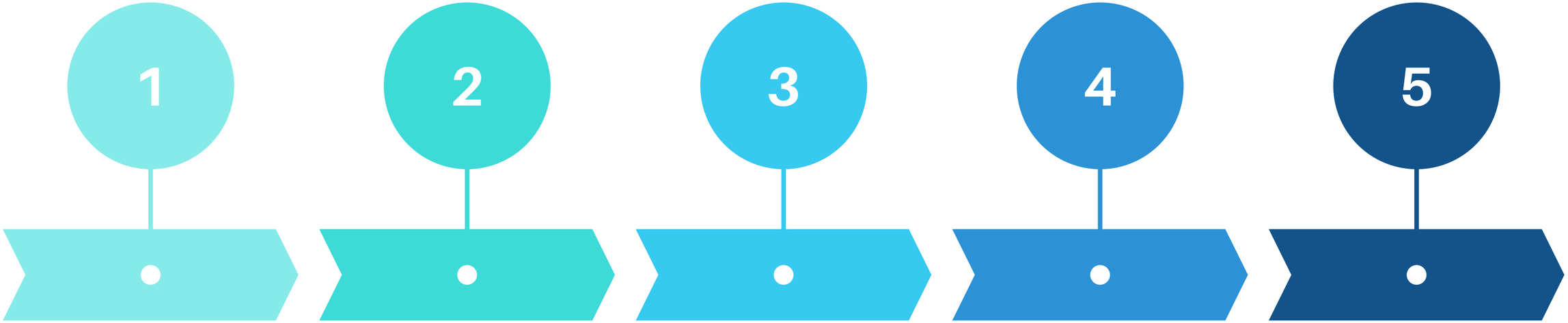
Identify the students at risk and accordingly, reach out to them and provide relevant suggestions

LITERATURE SURVEY

1. Identifying At-Risk Students: Using Machine Learning Techniques: A Case Study with IS 100
2. Dimensions of student success: a framework for defining and evaluating support for learning in higher education
3. Connect for Success: A proactive student identification and support program
4. Analyzing and Predicting Student's Performance by Means of Machine Learning: A review
5. Identifying students at risk of academic failure within the educational data mining framework

DATASET INFORMATION

UCI Machine Learning Repository



Two Schools
(Gabriel Pereira (GP)& Mousinho da Silveira (MS))

Demographics of Students
(Age, Gender, Address, etc.)

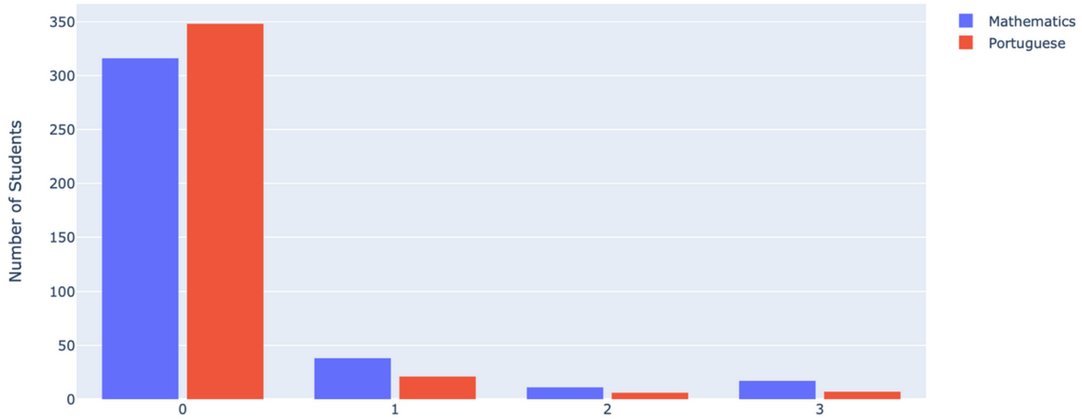
School Related Features
(School Support, Number of failures, absences, etc.)

Social Features
(Free time, go out, Romantic relationship, etc.)

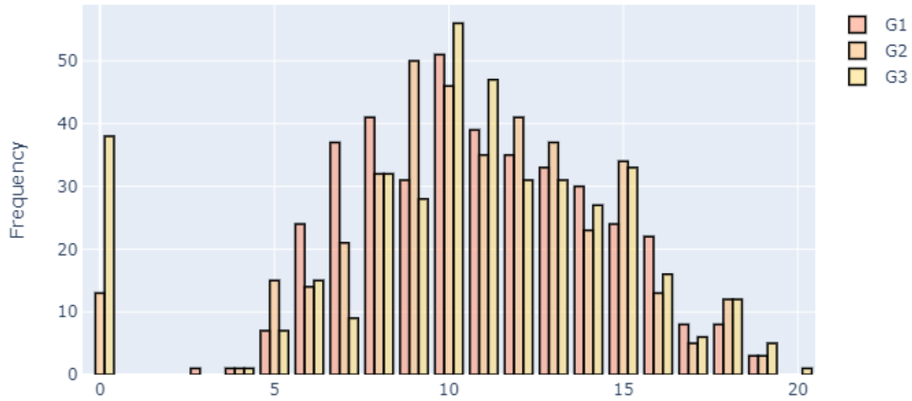
Student Grades
(G1, G2 & G3)

In Total, there are 33 features varying from school to student grades and for two subjects, Mathematics and Portuguese

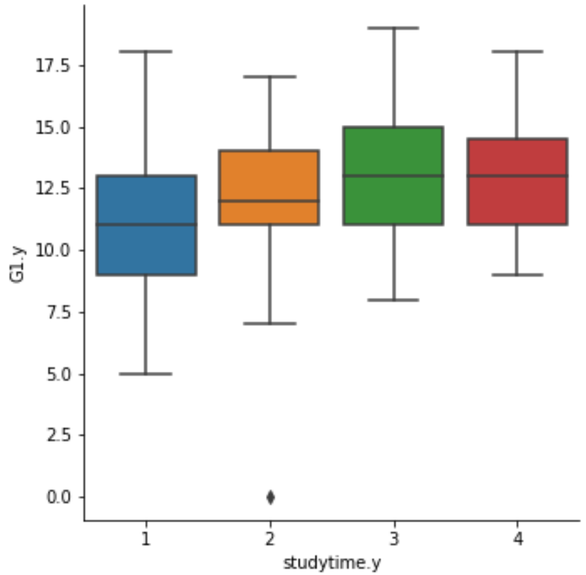
EXPLORATORY DATA ANALYSIS



Number of Failures for Maths and Portuguese



Portuguese Grades Distribution



Study Time vs G1 Grade for Portuguese

Failures vs G3 Grade for Portuguese

DIVISION OF FEATURES & GRADES INTO CLASSES

Grades G1 & G3

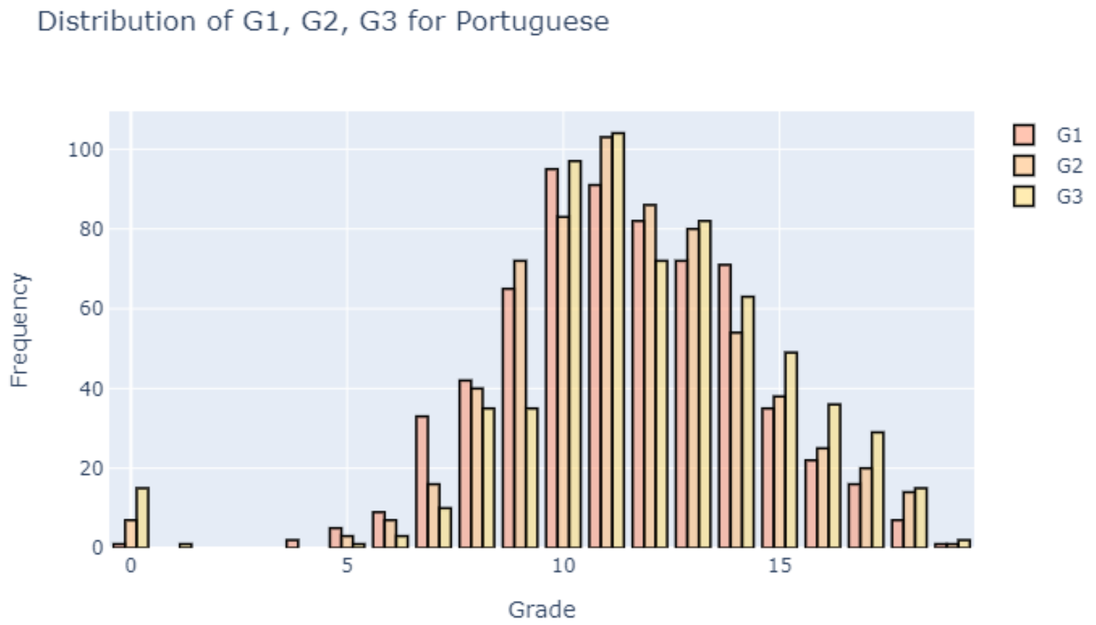
2 classes	5 classes
0 - 9 : B (Fail) 10 - 20 : A	0 - 9 : F (Fail) 10 - 11 : D 12 - 13 : C 14 - 15 : B 16 - 20 : A

Features

Demographic	Demographic + Activity (All Features)
Enrolment Options	Throughout the Semester

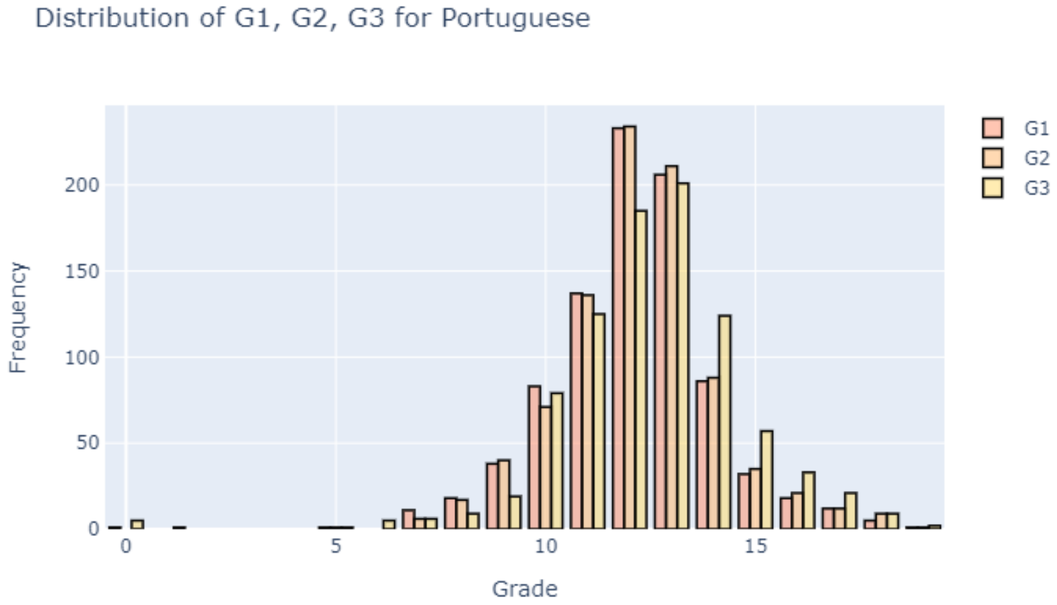
* Applied on all the datasets which are used

DATA EXTRAPOLATION



Original Data

(382 rows)



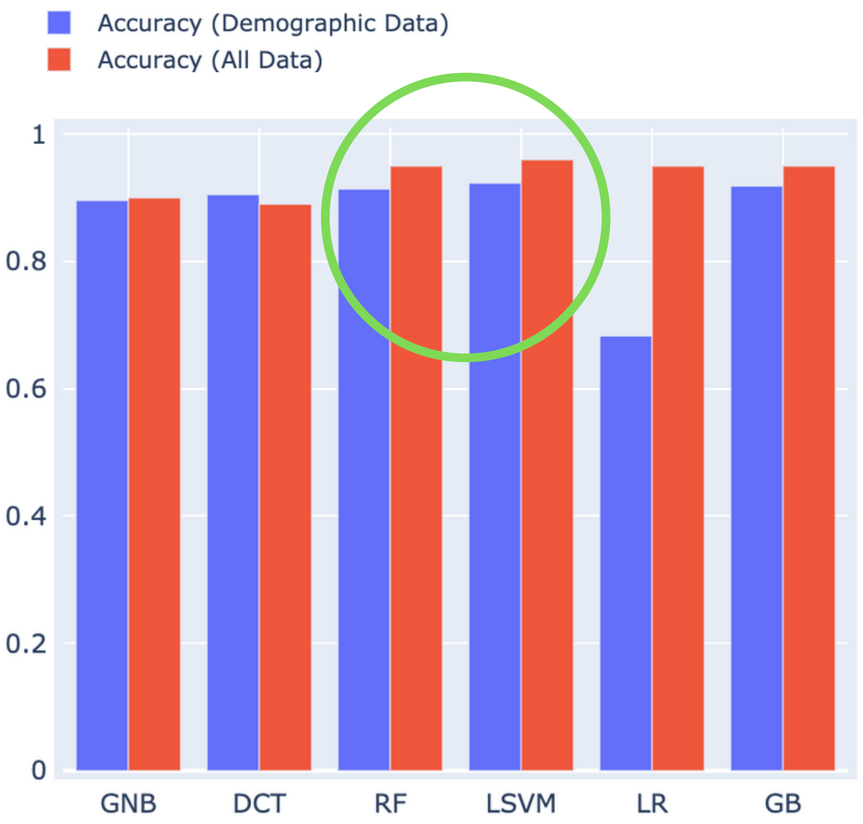
Extrapolated Data

(882 rows)

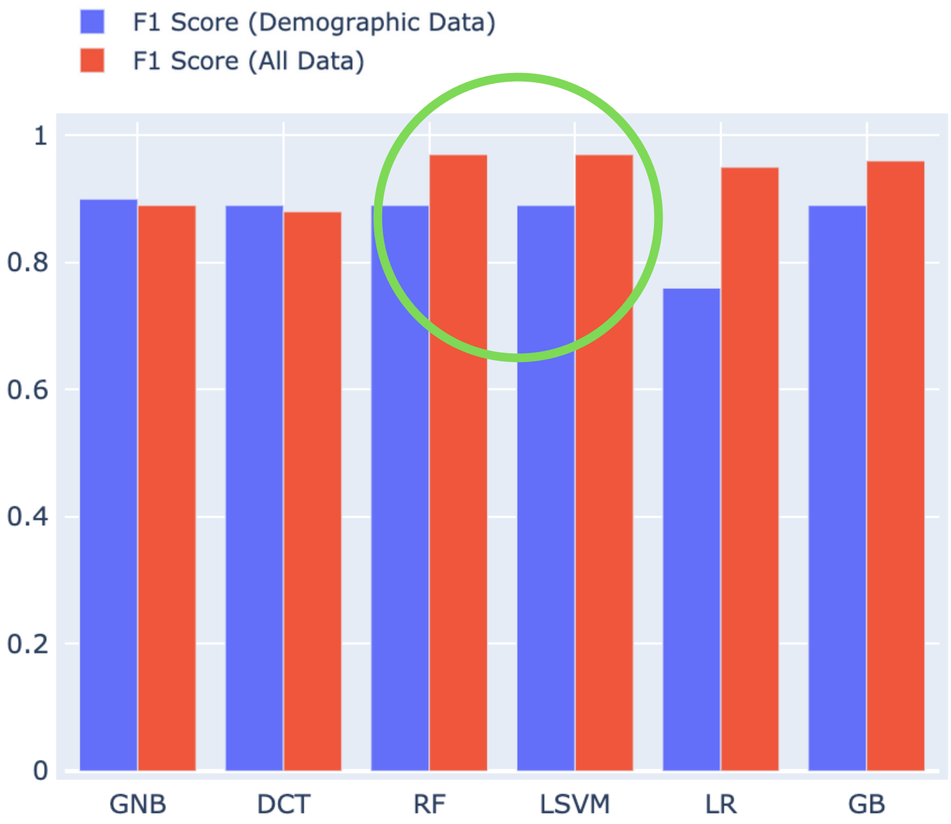
After comparing the results with original data, the extrapolated data gave better results and thus the extrapolated data was used for further analysis
(Performed for maths dataset also)

RESULTS OF PREDICTION FROM EXTRAPOLATED DATA

Prediction of G1 Score (Portuguese)
using Binary Classification Method



Accuracy

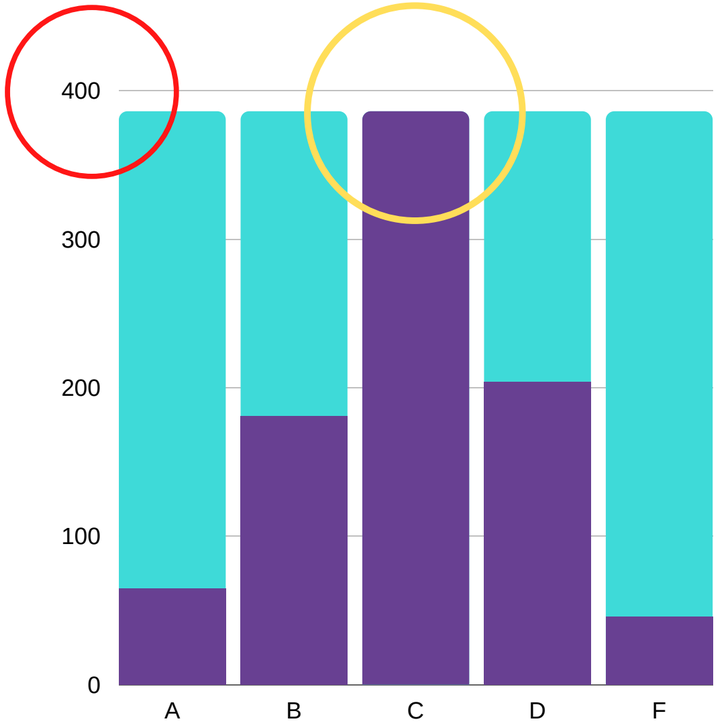


F1 Score

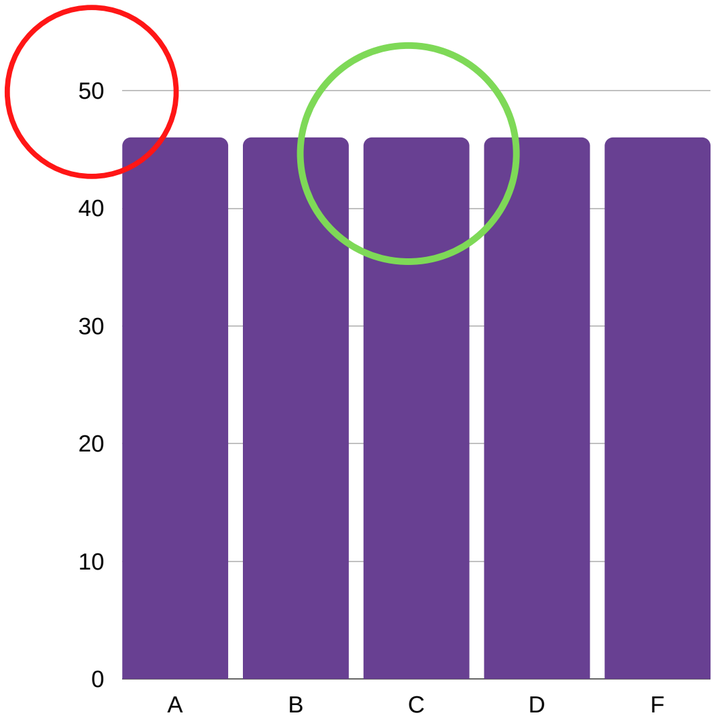
Results were obtained for Maths using Binary classification and for both the subjects, using 5 class classification techniqiue

DATA UPSAMPLING & DOWNSAMPLING (HANDLE DATA IMBALANCE)

Extrapolated Distribution of the grade G3 for Portuguese data



Upsampled Data (Portuguese)

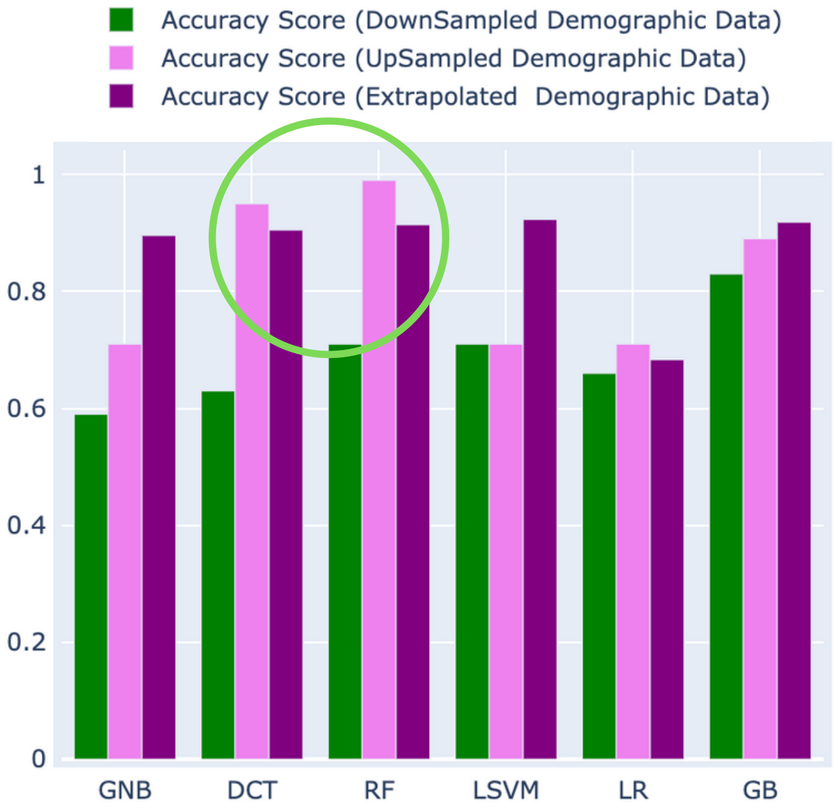


Downsampled Data (Portuguese)

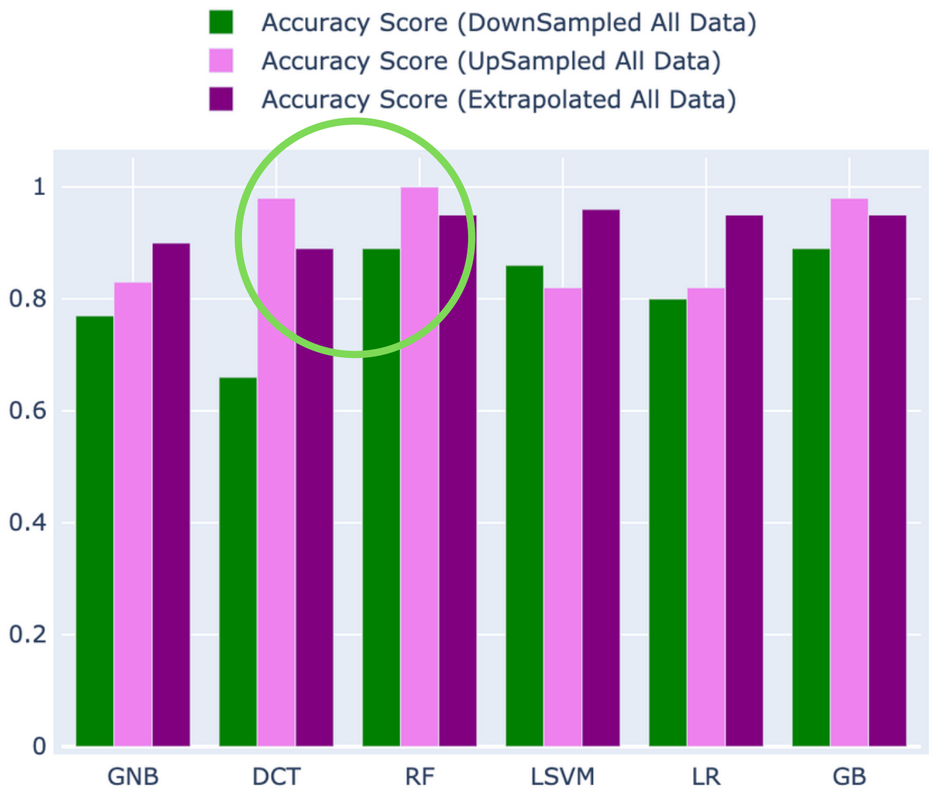
Similar Sampling techniques were performed on the data used for Binary Classification (Both Maths and Portuguese) and also for Maths data in 5 class classification

RESULTS OF PREDICTION USING SAMPLED DATA

Prediction of G1 Score (Portuguese) using Binary Classification Method



Comparison between accuracy obtained by 6 models



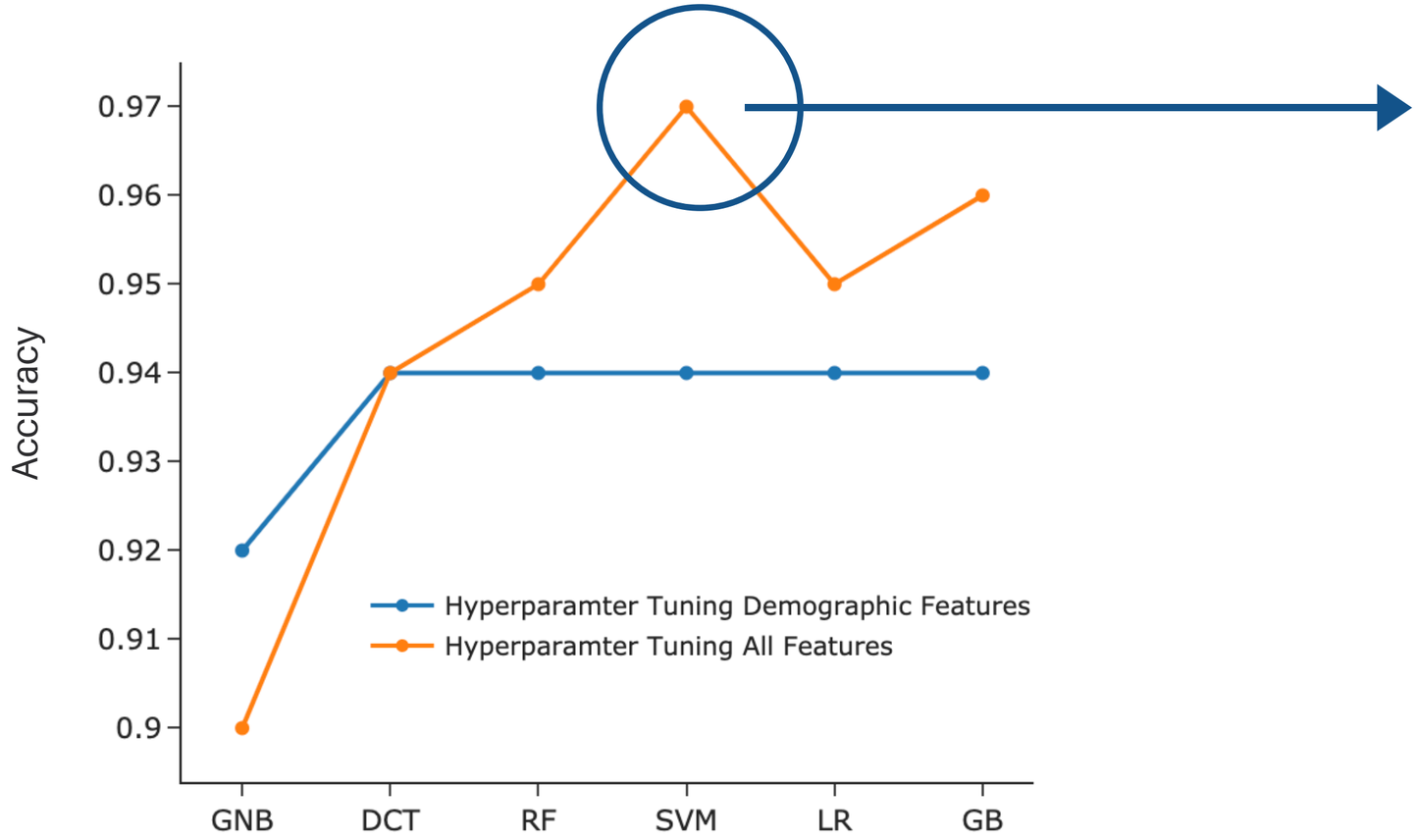
Comparison between accuracy obtained by 6 models

Results were obtained for Maths using Binary classification and for both the subjects, using 5 class classification techniqiue

HYPERPARAMETER TUNING

GridSearchCV and RandomizedSearchCV

Prediction of G1 Score (Portuguese) using Binary Classification Method



Best Parameters obtained:

C = 1
Kernel = 'Linear'
degree = 3
gamma = 1

Results were obtained for Maths using Binary classification and for both the subjects, using 5 class classification techniqiue
(Due to large computation time and memory issues, RandomizedSearchCV was preferred (Not much difference in results obtained by using GridSearchCV))

MOST IMPORTANT FEATURES

(Commonly found from all the above techniques)

Demographic Data

Mother's Education
Father's Job Status
Family Size
Age
School

Entire Dataset

Study Time
Father's Job
Go Out
Number of Absences
Number of Failures

PROACTIVE STUDENT SUPPORT AND NEXT BEST ACTION

SCENARIO MODELLING

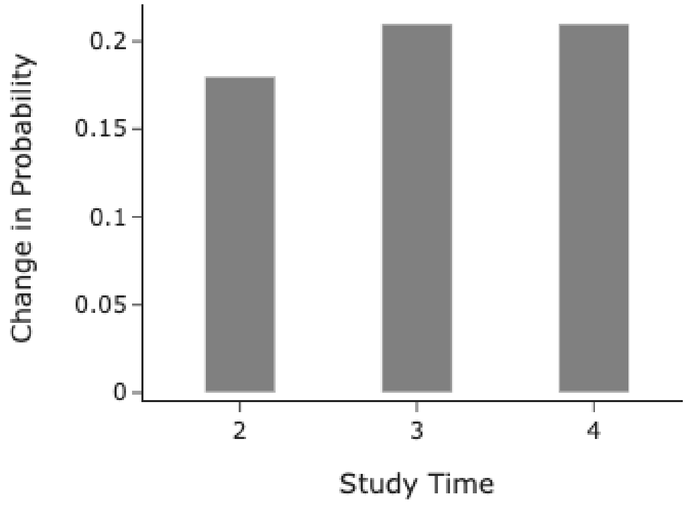
Everything looks good with a story!

SCENARIO 1

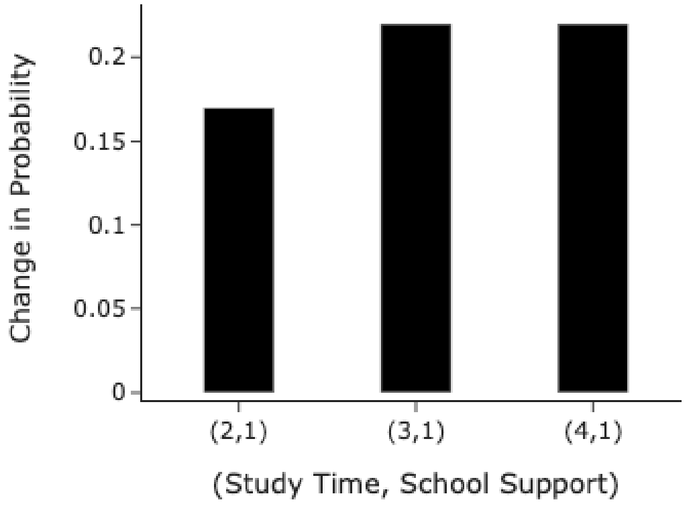
Student ID - 342

Initial probability of passing in G1 exam of Portuguese subject is 0.26
Initial study time value is 1 (studying less than 2 hours per week)

Scenario with Study Time



Scenario with Study Time & School Support



Similarly, more scenarios were explored using multiple features for both the subjects

"Strive for Progress, not Perfection"

"Strive for Progress, not Perfection"

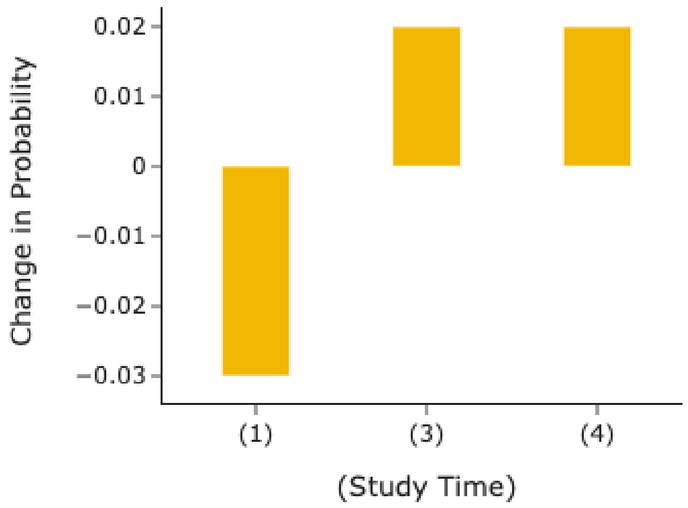
SCENARIO 2

Student ID - 45386

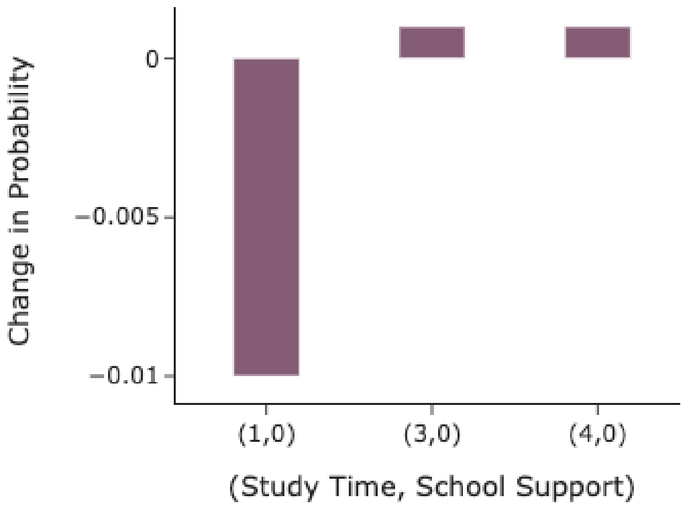
Initial probability of passing in G1 exam of Portuguese subject is 0.95

Initial study time value is 2 (studying between 2 & 5 hours per week)

Scenario with Study Time



Scenario with Study Time & School Support



Similarly, more scenarios were explored using multiple features for both the subjects

FINAL OUTCOMES

EXCEED CLIENT EXPECTATIONS

Client mentioned that the team has achieved all goals

REAL WORLD PROJECT EXPERIENCE

DEVELOPED MODEL FOR PREDICTING STUDENT'S GRADE

MAKES AN IMPACT

Developed Scenario Modeling System to provide suggestions to students

Thank You!

GROUP 25
STUDENT TEAM
THE UNIVERSITY OF MELBOURNE