PAPER NAME

**Project Synopsis_Group_10_Final.docx**

| | |
|---|---|
| WORD COUNT | CHARACTER COUNT |
| **4213 Words** | **24809 Characters** |
| PAGE COUNT | FILE SIZE |
| **39 Pages** | **3.8MB** |
| SUBMISSION DATE | REPORT DATE |
| **May 29, 2024 10:59 AM GMT+5:30** | **May 29, 2024 11:00 AM GMT+5:30** |

## 9% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 7% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 7% Submitted Works database

## Excluded from Similarity Report

- Quoted material
- Cited material
- Small Matches (Less then 10 words)
- Manually excluded sources
- Manually excluded text blocks

# A Project Report

# On

# STUDENT DROPOUT ANALYSIS

## END TERM VII SEMESTER SYNOPSIS REPORT

***Submitted in partial fulfillment of the requirement of the degree of***

***of***

**Bachelors of Technology**
**in Computer Science Engineering and Information Technology**

***by***

**Aditya Aggarwal(20CSU003)**
**Aditya Arora(20CSU004)**
**Aman Vohra(20CSU010)**
**Saakshi Rana(20CSU091)**

**Under the supervision of**
**Dr. Akanksha Kaushik**
**Assistant Professor, Department of Computer Science and Engineering**

**Department of CSE & IT**
**The NorthCap University, Gurugram**
**May 2024**

# CERTIFICATE

**This is to certify that the Project Synopsis entitled, "Student Dropout Analysis" submitted by "Aditya Aggarwal, Aditya Arora, Aman Vohra and Saakshi Rana" to The NorthCap University, Gurugram, India, is a record of bona fide synopsis work carried out by them under my supervision and guidance and is worthy of consideration for the partial fulfilment of the degree of Bachelor of Technology in Computer Science and Engineering of the University.**

**Dr. Akanksha Kaushik**

**Date: …22/05/2024……**

# ACKNOWLEDGEMENT

We would like to express our sincere gratitude to everyone who supported and contributed to our group project on student dropout rate analysis. Special thanks to our mentor Dr. Akanksha Kaushik, for their invaluable guidance and constructive feedback throughout the project. We also extend our appreciation to the faculty and staff of The Northcap University for providing the resources and support necessary for our research.

We are grateful to the students, educators, and administrators who participated in our surveys and interviews, sharing their insights and experiences, which greatly enriched our work. Additionally, we acknowledge the collaboration and dedication of our group members: Aditya Aggarwal, Aditya Arora, Aman Vohra and Saakshi Rana. Your teamwork and commitment were essential to the success of this project.

Finally, we thank our families and friends for their encouragement and understanding. Your collective efforts and contributions have been crucial to the completion and success of this project.

# ABSTRACT

The issue of student dropout rates in India poses a significant challenge to the education system, affecting not only individual students but also societal progress. Understanding the factors contributing to student attrition is crucial for implementing effective interventions. This project aims to conduct a comprehensive analysis of student dropout rates in India using machine learning models.

The dataset comprises a diverse range of variables including demographic information, academic performance, socio-economic status, and other relevant factors. Through exploratory data analysis, key patterns and trends influencing dropout rates will be identified. Feature engineering techniques will be employed to preprocess and extract meaningful insights from the dataset.

Various machine learning algorithms such as logistic regression, decision trees, random forests, and gradient boosting will be applied to build predictive models. The models will be trained on historical data to predict the likelihood of student dropout based on the identified features. Evaluation metrics such as accuracy, precision, recall, and F1-score will be utilized to assess the performance of each model.

The outcomes of this project have the potential to inform policymakers, educators, and stakeholders about effective strategies for reducing student dropout rates in India. By leveraging machine learning models, this analysis aims to contribute towards creating a more inclusive and equitable education system that ensures better retention and success for all students.

# TABLE OF CONTENT

# LIST OF FIGURES

# 1. INTRODUCTION

Student dropout is a critical issue in educational institutions worldwide, significantly impacting individuals and society. It refers to students discontinuing their education before completing their degree or program. Understanding the factors behind student dropout and developing strategies to address it is crucial for improving educational outcomes and helping students reach their potential.

Dropout rates vary across educational levels, regions, and socio-economic backgrounds. For individuals, dropping out can lead to reduced job opportunities, lower earnings, and increased risk of engaging in criminal activities. For society, high dropout rates result in a less educated workforce, increased social service costs, and slower economic growth.

Several factors contribute to student dropout, including academic difficulties, financial constraints, lack of engagement, personal issues, and institutional shortcomings. Academic challenges can lead to frustration, while financial problems can make education unaffordable. Lack of engagement and support from the school community can leave students feeling isolated. Personal issues, like health problems or family responsibilities, can also interfere with education. Institutional factors, such as inadequate support services and rigid curricula, can exacerbate the problem.

Addressing student dropout requires a comprehensive approach involving educators, policymakers, parents, and students. Interventions may include academic support, financial aid, mentorship, counseling, and initiatives to create a more inclusive and engaging school environment. Early identification of at-risk students and targeted support can help reduce dropout rates and promote student success.

In conclusion, student dropout is a complex issue needing coordinated efforts from various sectors. Understanding its causes and implementing effective strategies can improve educational outcomes and benefit both individuals and communities. This report will analyze the factors behind student dropout, explore potential interventions, and provide actionable solutions.

## 1.1 BACKGROUND STUDY

Understanding and addressing student dropout rates in India is critical due to the multifaceted nature of the issue. Factors such as socio-economic background, academic performance, and psychosocial dynamics play pivotal roles in determining whether a student stays enrolled or leaves their educational journey prematurely. Despite numerous efforts aimed at improving retention rates, dropout rates persist as a significant challenge within the Indian education system.

Machine learning models present a promising avenue for gaining deeper insights into dropout patterns and predicting the likelihood of students dropping out based on historical data. By leveraging these advanced algorithms, we can analyze vast datasets to uncover nuanced correlations and predictive indicators that may not be readily apparent through traditional statistical methods.

19 The objective of this project is to harness the power of machine learning to conduct a comprehensive analysis of historical dropout data from various educational institutions in India. Through this analysis, we aim to identify patterns, trends, and underlying factors contributing to dropout rates across different demographic groups and geographical regions. By gaining a deeper understanding of these dynamics, we can develop more targeted and effective interventions to prevent dropout and improve student retention.

Ultimately, the insights generated from this project have the potential to inform policymakers, educators, and stakeholders about the most effective strategies for reducing dropout rates and promoting educational equity and inclusivity in India. By deploying machine learning models in this endeavor, we can enhance the efficiency and accuracy of our analyses, leading to more informed decision-making and impactful interventions in the realm of education.

# 1.2 EXPANDING THE DATASET

**Mental Health Metrics:**

Include columns that capture information about students' mental health, such as self-reported stress levels, anxiety, depression, or access to counseling services. These metrics can help identify students who may be struggling emotionally and at higher risk of dropping out.

**Attendance Patterns:**

Track students' attendance patterns, including the number of missed classes, late arrivals, or consistent absenteeism. Irregular attendance can be indicative of academic or personal challenges.

**Extracurricular Involvement:**

Include data on students' participation in extracurricular activities, clubs, or sports. Engaged students often have a stronger sense of belonging to the college community and may be less likely to drop out.

**Satisfaction Surveys:**

Administer satisfaction surveys to students periodically, collecting data on their overall satisfaction with the college, specific programs, and support services.

**Internship/Job Placement Status:**

Include information on students' success in securing internships or job placements related to their field of study. Lack of career prospects may influence dropout decisions.

**Library and Resource Center Usage:**

Record how frequently students use on-campus resources like libraries, resource centers, and academic labs. Utilization of these facilities can indicate students' commitment to their studies.

# 2. LITERATURE REVIEW

The study conducted under Core utilized data extracted from the Economics Information System and implemented a Decision Tree analysis. The findings from the training sample indicated that 98 students, constituting 34.3%, discontinued their studies, while 188 students, accounting for 65.7%, continued into the second year.

Among the students who dropped out, 79 of them received a failing grade in Mathematics. Nineteen students discontinued for other reasons, with seven attributing their dropout to a failing mark in Statistics, despite having a passing grade in Mathematics. Twelve students dropped out despite passing both Mathematics and Statistics. Among these, eight cited a dropout reason as having a mark of 2 (sufficient) or lower in Introduction to Economics, while in four students, the mark exceeded 2.

The study, published in the SHS Web of Conferences, reported that the logistic regression model achieved 94.44% accuracy, 95% precision, 98% recall, and 97% F1 score in predicting student paper churn. Of the 90 data used for testing, the model predicted 71 student failures. However, there are four cases of students who were mistakenly thought to have failed**.** Additionally, 14 students who dropped out of school were identified in the sample.

In a study conducted under IEOM, the authors investigated and reported on the effectiveness of three supervised learning approaches in predicting student graduation. Random Forest emerged as the most successful model, scoring a maximum of 95.00% accuracy, a 97.78% precision, and a 96.70% F1 Score. This outcome demonstrated the program's success in fulfilling both the primary and secondary objectives of the study, as it effectively differentiated the performance metrics of various supervised learning techniques and accurately predicted student outcomes.

In a study published in IJATCSE, the authors employed Decision Tree Classifier and Naive Bayes Classifier to identify key factors influencing student dropouts. Research shows that the main determinants of student failure relate to academics, including academic performance, quarterly test scores, attendance, and writing level**.** Additionally, the research highlighted the significant impact of parents' income on student dropout rates.

# 2.1 USE CASES/OBJECTIVES FOR INSTITUTIONS

I launch my product in the form of a website, it can be engineered according to your needs whatever it may be, you may be a coaching institute, college, high secondary school etc.

Whatever features you want to add according to your needs will be catered to but the final outcome of the project will remain the same, whether or not the student will drop out or no.

**Parents' Portal:**

Create a user-friendly web interface where parents can input their child's details, such as academic performance, attendance, and extracurricular activities. The model can then provide a predictive assessment of the likelihood of their child dropping out. This would help parents proactively engage with their child's education and potentially seek additional support.

**Teacher's Toolkit:**

Offer a toolkit for teachers in sorting the students who can be future dropouts. Teachers can input students' academic performance, attendance to receive predictions. This can enable educators to provide targeted assistance and mentoring to at-risk students.

**Student Self-Assessment:**

Create a student-facing portal where students can assess their own risk of dropping out. This can promote self-awareness and encourage students to seek help or engage with support services if they identify as being at risk.

**High School Transition Support:**

Extend the model's application to high schools to help students make informed decisions about college readiness and course selection. High school counselors and students can use the tool to plan educational paths that minimize dropout risks.

**Coaching Institutes:**

Adapt the model for coaching institutes to identify students at risk of discontinuing their test preparation courses. Institutes can provide additional support and resources to help students stay on track and achieve their goals.

**Career Counselors:**

Offer a tool for career counselors to assess students' dropout risks in the context of their chosen career paths. This can help tailor counseling sessions and guidance to address potential challenges.

## 2.2 USE CASES/OBJECTIVES FOR GOVERNMENT

Policymakers in government can use your product in the following ways:

### 2.2.1 Targeted Resource Allocation:

Government officials can use the predictive model to identify regions, schools, or districts with high dropout rates. This information can guide the allocation of resources, such as funding, additional support services, and infrastructure improvements, to areas where they are needed most.

### 2.2.2 Early Intervention Planning:

The model can help policy makers identify the specific factors contributing to dropout rates, whether they are academic, socioeconomic, or related to other circumstances. This information can inform the design of early intervention programs and policies that target the root causes of dropout.

### 2.2.3 Monitoring Policy Effectiveness:

Government agencies can regularly assess the impact of their educational policies and initiatives by comparing dropout rates before and after policy implementation. Your product can serve as a valuable tool for measuring the success of various policies and making necessary adjustments.

### 2.2.4 Identifying At-Risk Schools and Districts:

Government officials can use the model to identify schools or districts with a higher risk of dropout, allowing for targeted support and policy implementation in these areas

# 3. PROBLEM DEFINITION AND REQUIREMENT ANALYSIS

Creating a machine learning system to forecast student dropout rates by analyzing historical data on various factors such as demographics, academic performance, and socio-economic backgrounds. The system aims to identify students at risk of dropping out and intervene promptly with targeted support strategies. Through accurate predictions and timely interventions, the goal is to improve student retention rates and enhance overall educational outcomes.

The feasibility of this project is highly promising, primarily due to the availability of a comprehensive dataset and the potential benefits it offers. Here are key considerations regarding the feasibility:

**Data Availability:** The dataset contains a wide array of student-related attributes, which is essential for conducting a thorough analysis. This data has already been collected and is readily accessible for analysis.

**Analytical Tools**: We have access to a variety of advanced analytical tools and statistical techniques to explore the dataset and develop predictive models. Machine learning algorithms, statistical regression, and data visualization tools will be utilized to extract valuable insights.

**Ethical Considerations**: We will ensure the project adheres to ethical standards and data privacy regulations.

**Outcome Expectations**: The expected outcomes of this project are significant. These include the identification of key factors contributing to student dropout, the development of a predictive model, and actionable recommendations for early intervention. Successful implementation of these findings has the potential to substantially reduce the dropout rate in our college, leading to improved student retention and academic success.

# 4. OUTCOMES

**4.1 Cost Savings**: Reducing dropout rates can result in cost savings for both students and institutions. Students who complete their degrees are likely to earn higher incomes, while institutions can save on the costs associated with recruiting and enrolling new students.

**4.2 Improved Graduation Rates**: Implementing the findings from dropout analysis can lead to higher graduation rates, which can positively impact an institution's reputation and alumni success.

**4.3 Student Engagement**: Through the identification of factors contributing to dropout, institutions can focus on improving student engagement strategies, such as active learning methods, co-curricular activities, and community involvement.

**4.4 Community Impact**: Reducing college dropout rates can have a positive impact on the local and regional community by producing a more educated and skilled workforce, which can attract businesses and drive economic development.

# 5. DESIGN AND IMPLEMENTATION

## 5.1 EXPLORATORY DATA ANALYSIS

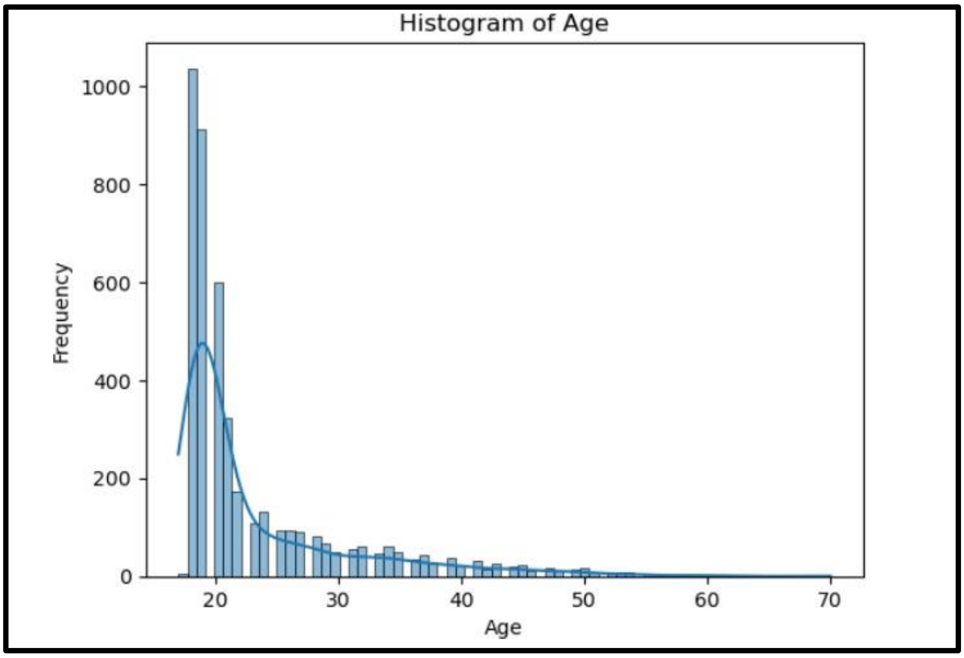



**Fig 5.1.1 This diagram shows the histogram of Age plotted against the Frequency of its appearance.**

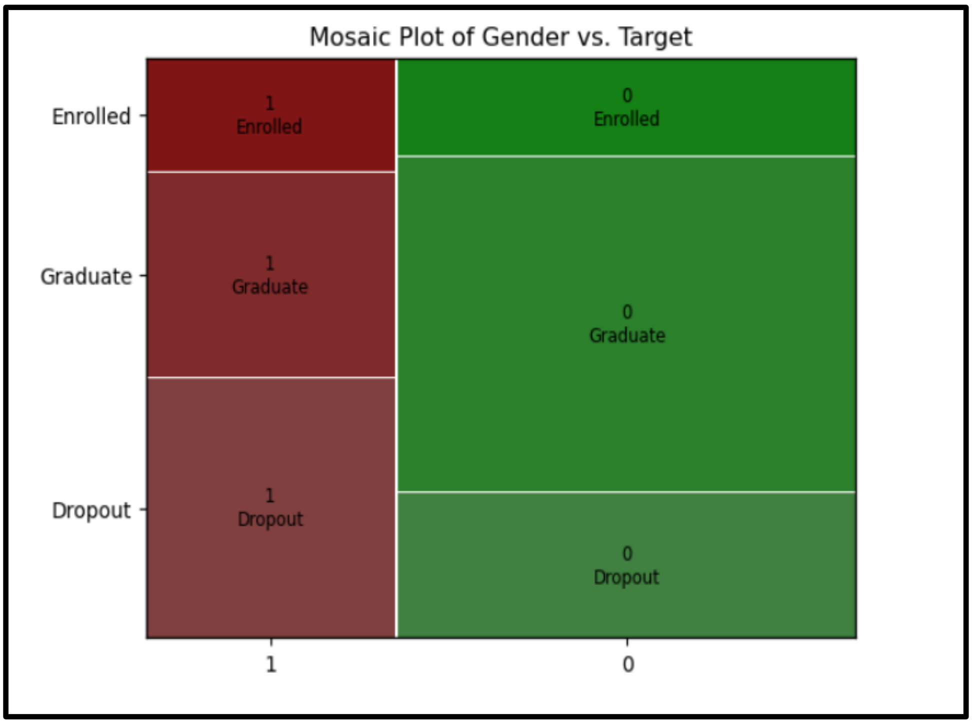



**Fig 5.1.2 This diagram shows the mosaic plot for Gender vs Target.**

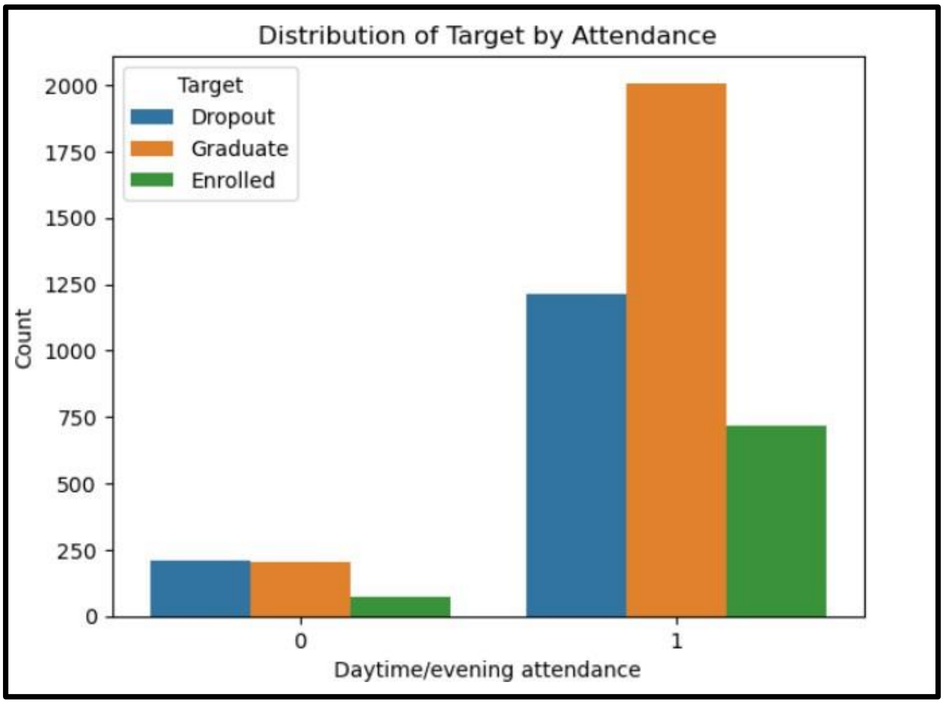



**Fig 5.1.3 This diagram shows the Distribution of Target by the Attendance for each category**

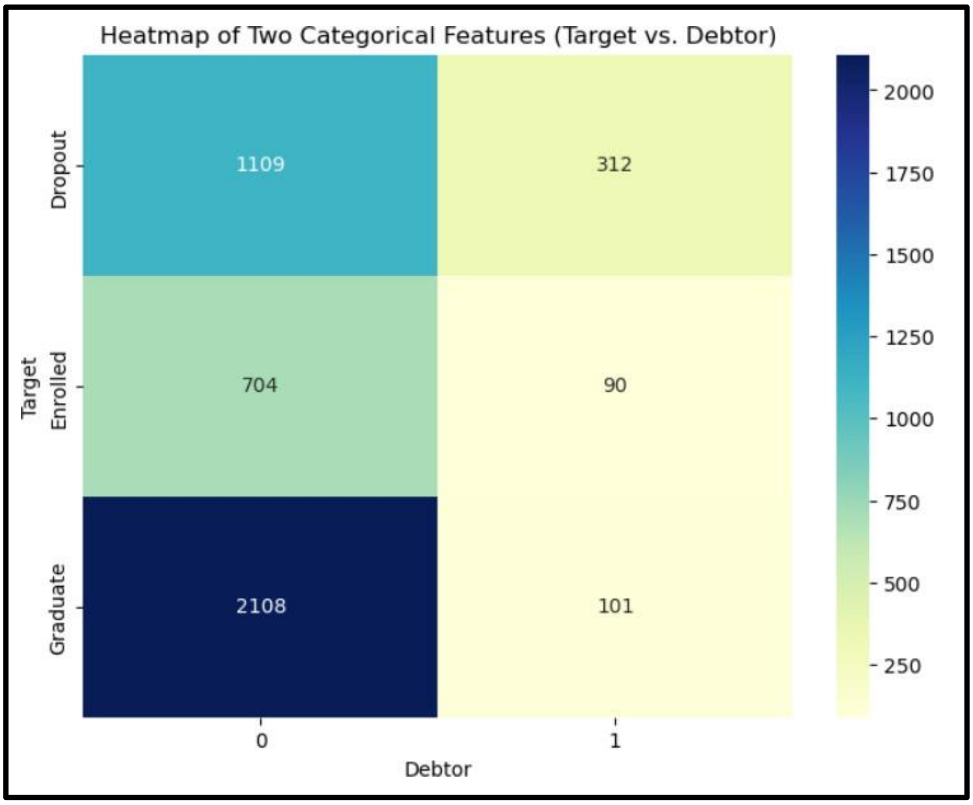



**Fig 5.1.4 This heat map shows two categorical features Target vs. Debtor.**

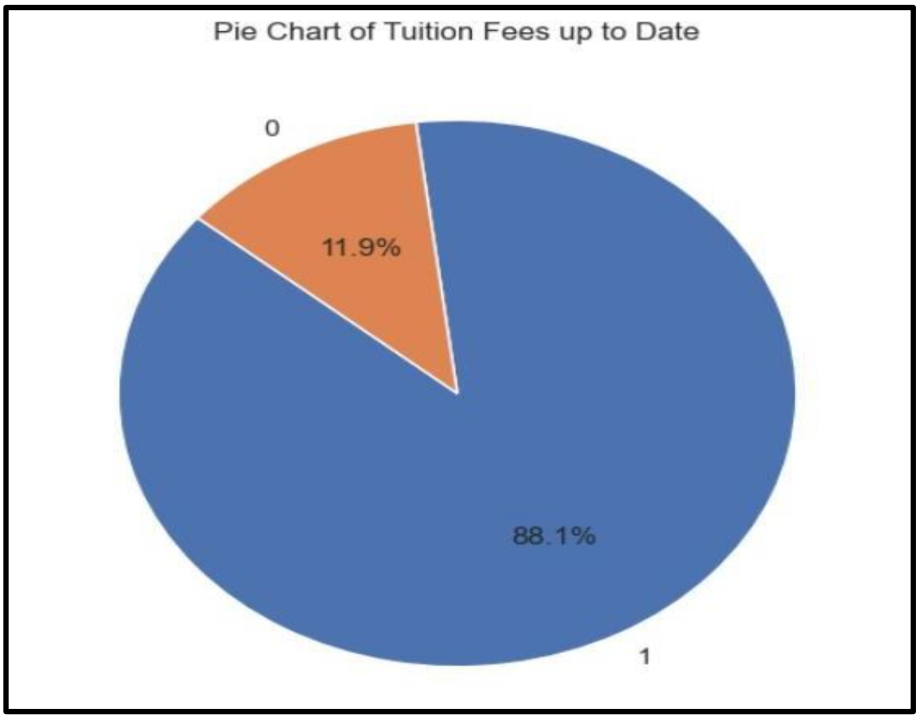



**Fig 5.1.5 This pie chart shows the percentage of students who have paid their tuition fees.**

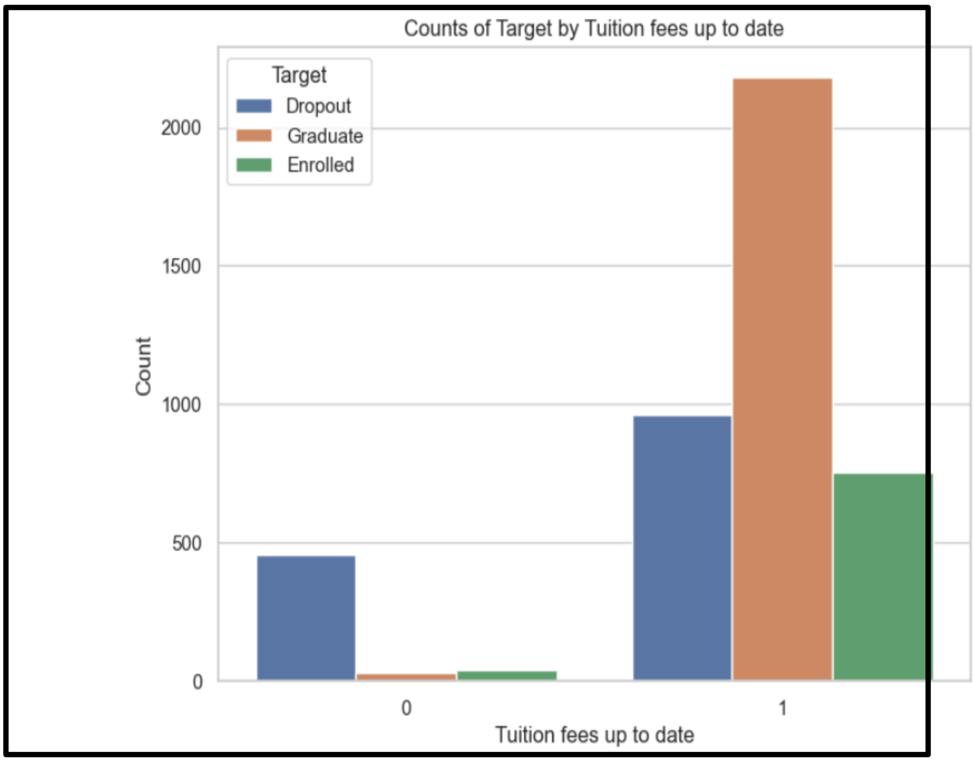



**Fig 5.1.6 This count plot shows the count of the number of students who have paid their tuition fees up to date.**

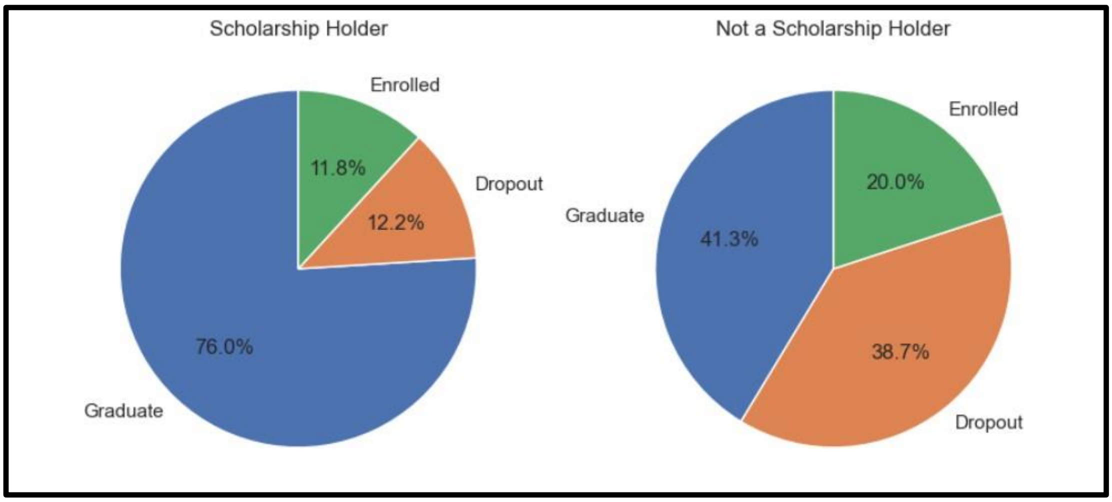



**Fig 5.1.7 This pie chart shows the percentage of students who have received scholarship and those who haven't.**

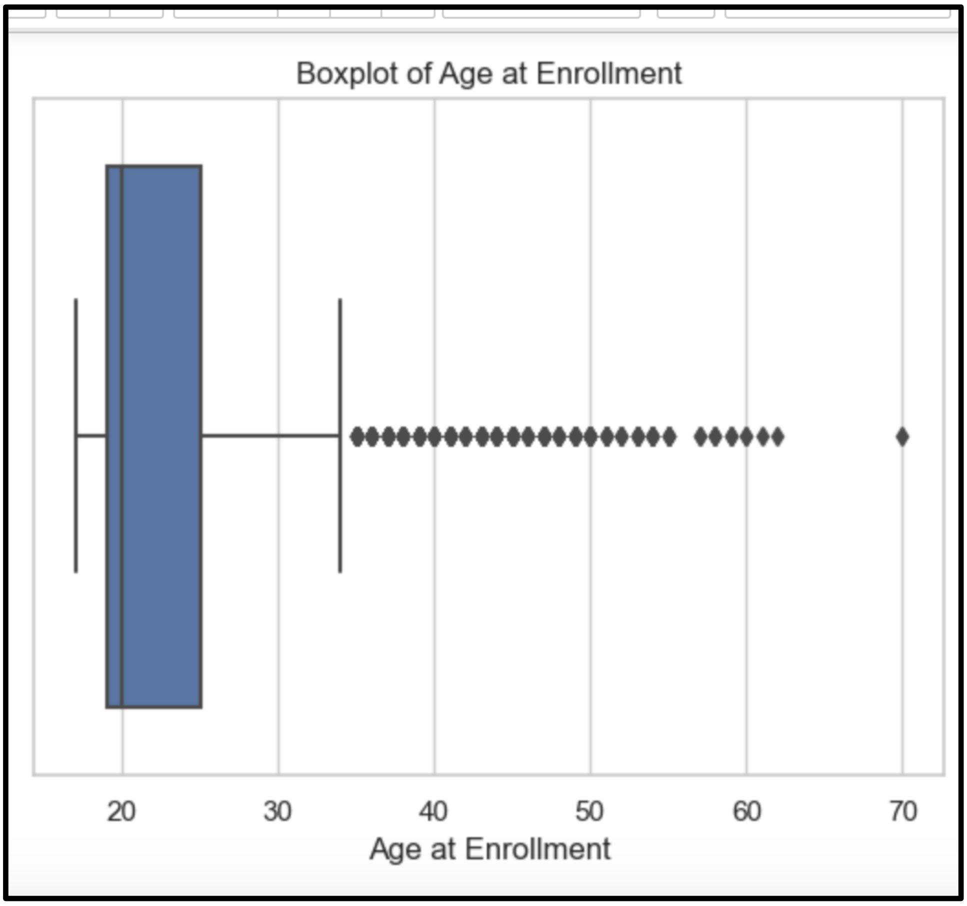



**Fig 5.1.8 This Box Plot shows the Age wise Distribution of the dataset along with outlier at Age 70.**

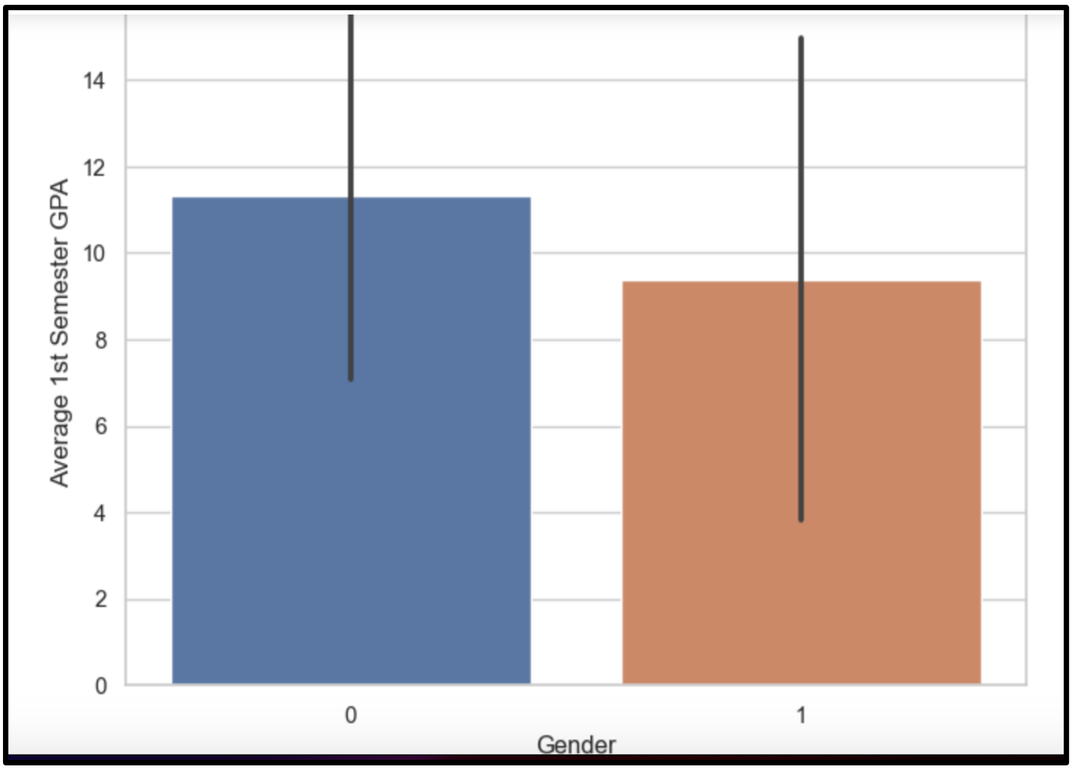



**Fig 5.1.9 The plot shows Avg. GPA by Gender**

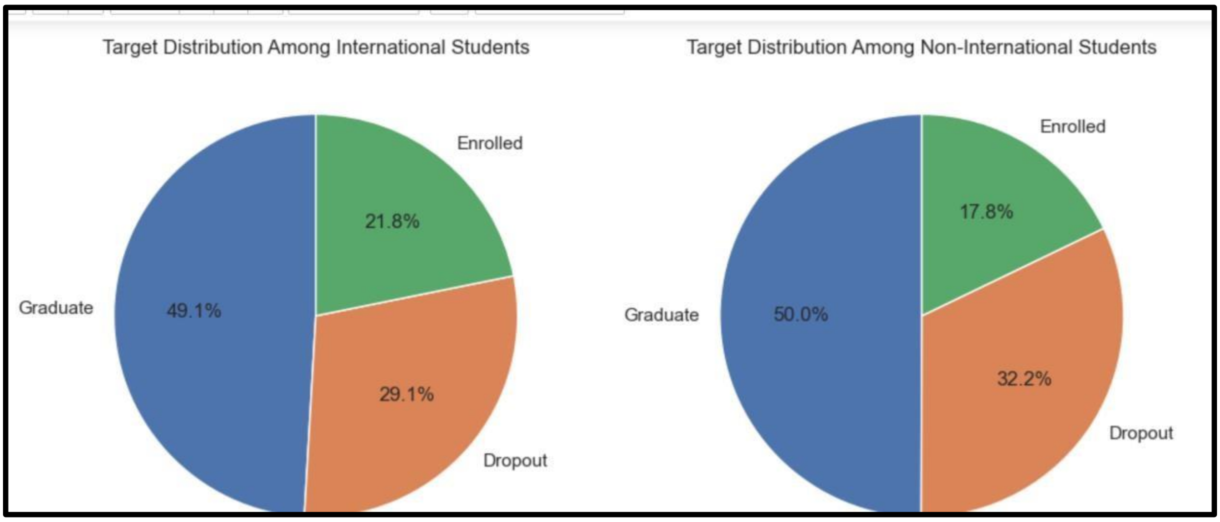



**Fig 5.1.10 This pie chart shows the target distribution among International and Non International Students.**

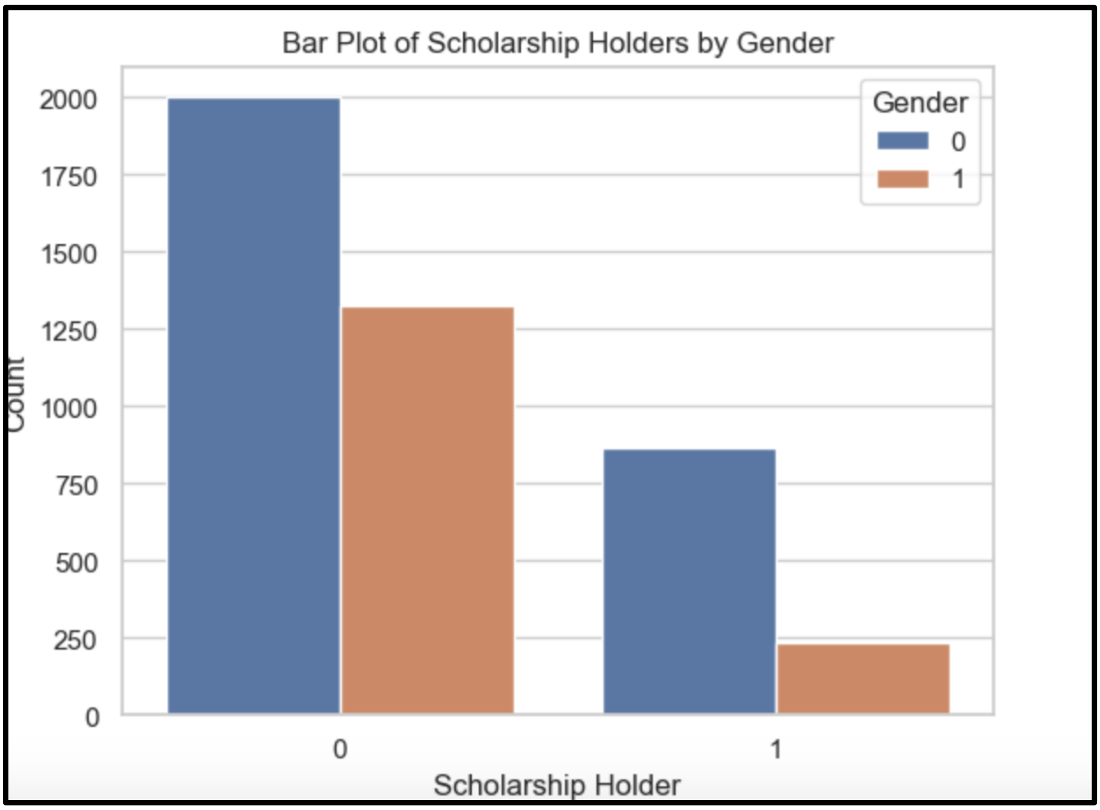



**Fig 5.1.11 This Bar chart shows scholarship holders by Gender**

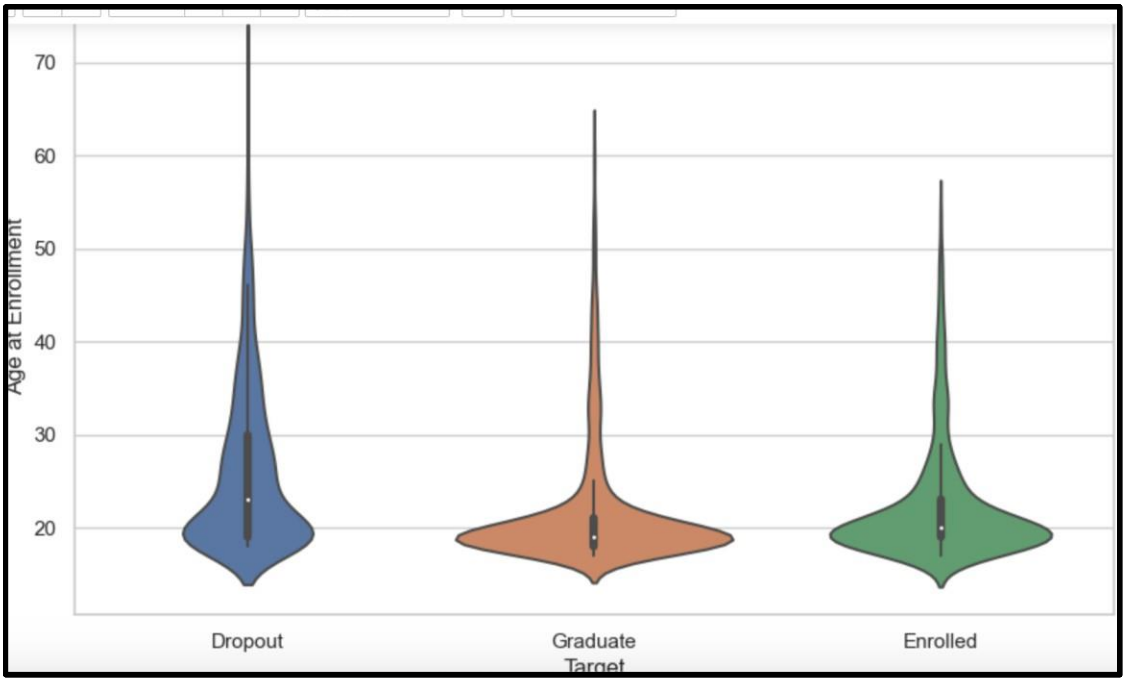



**Fig 5.1.12 This Violin plot shows Age at enrollment by target category**

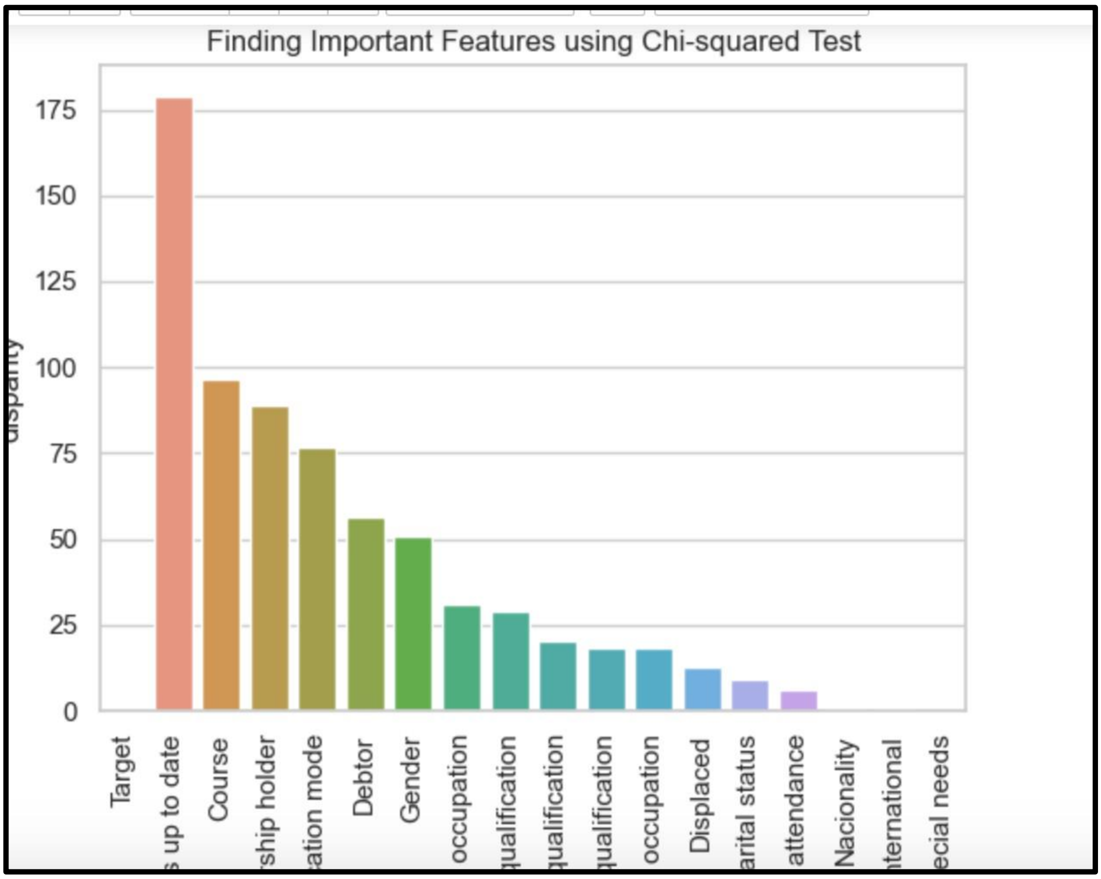



**Fig 5.1.13 This test shows the important features in the dataset**

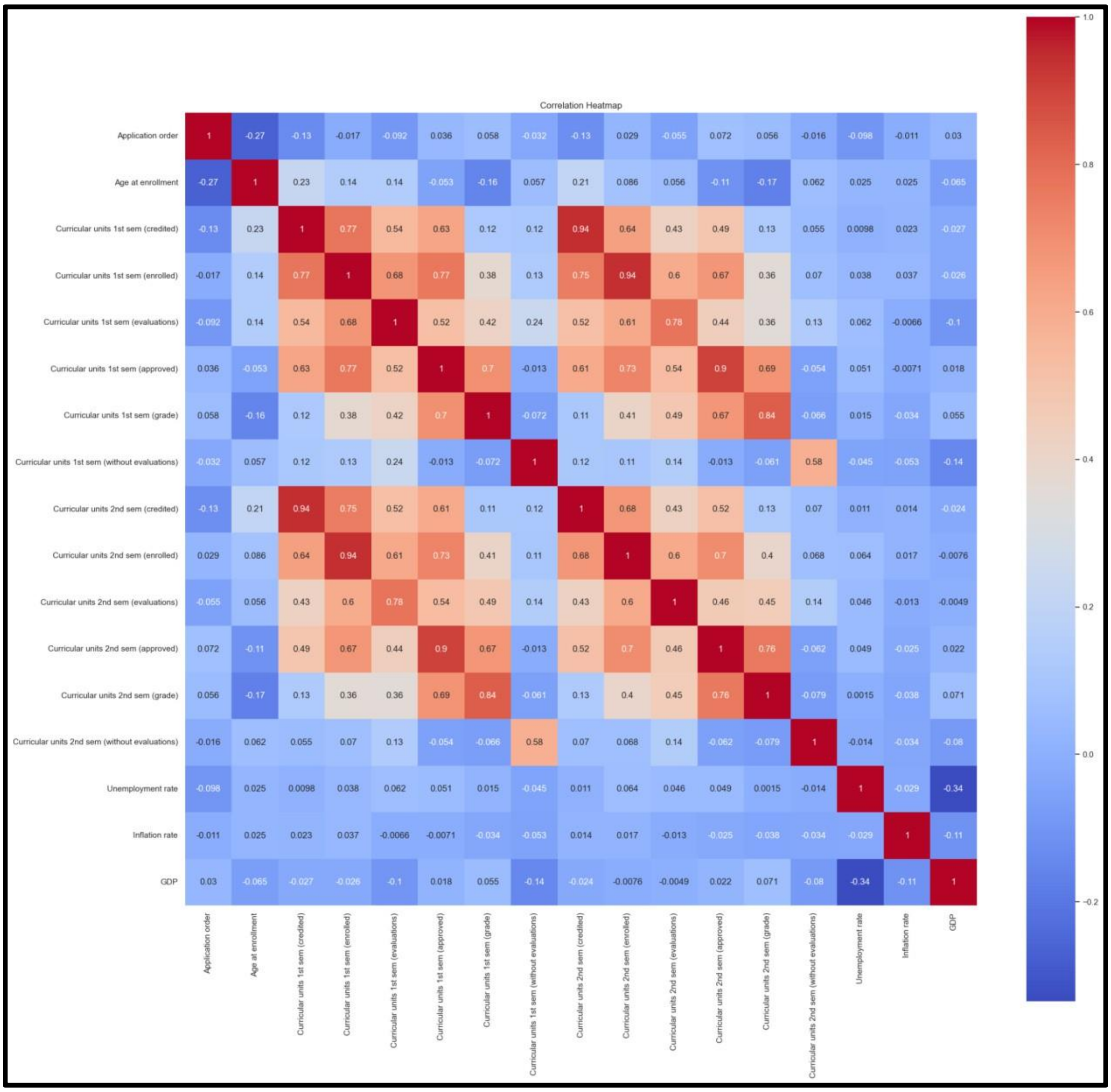



**Fig 5.1.14 This map shows the correlation between the features and we find curricular units 1st semester (credited) is correlated with 2nd semester(credited).**

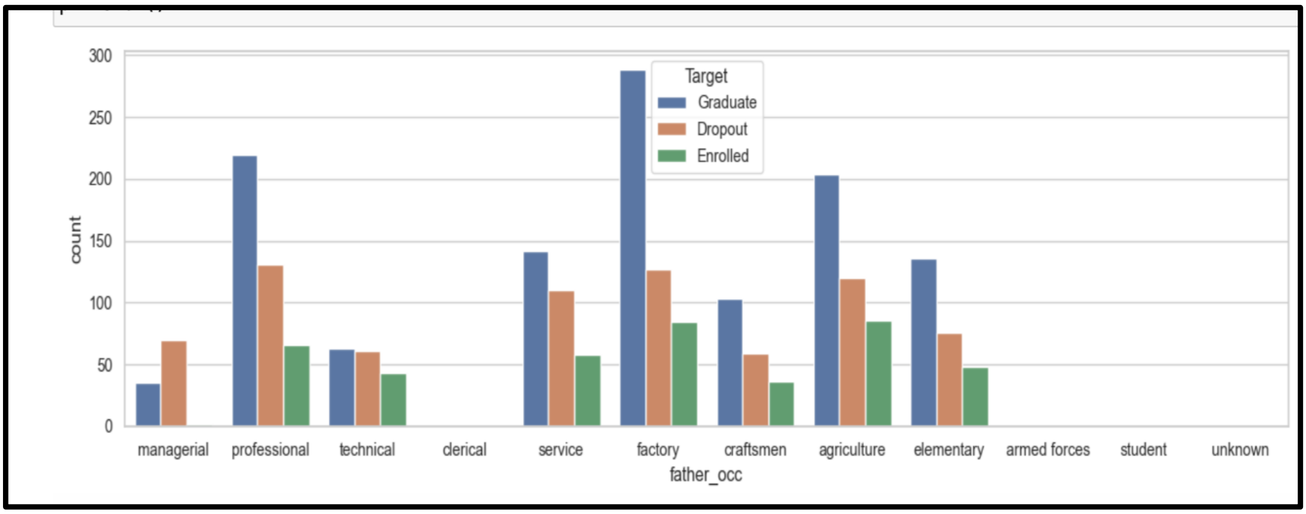
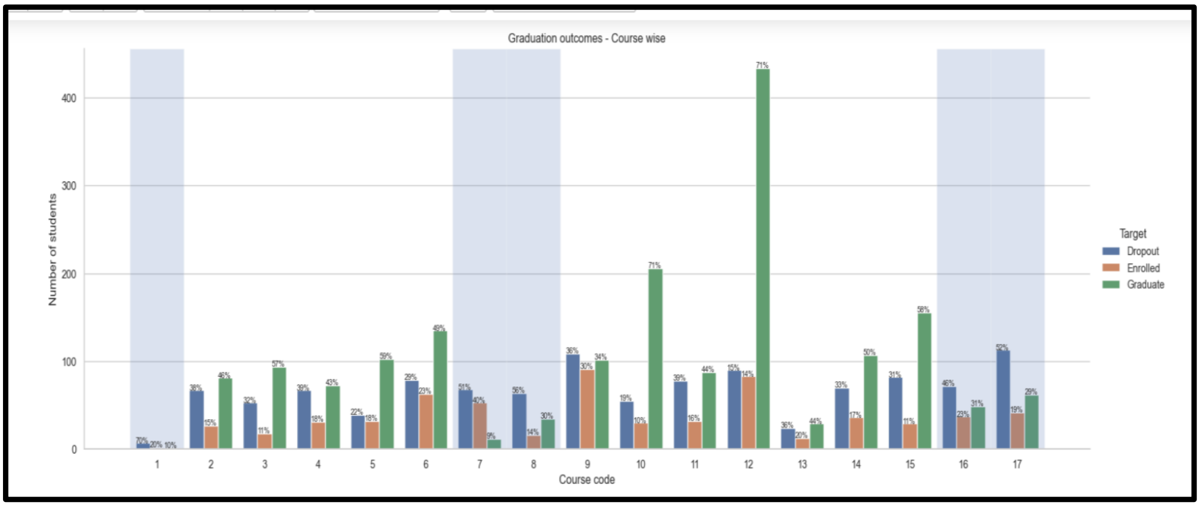



**Fig 5.1.15 This Count Plot shows the Father's Qualification by target category.**




The following courses have a much higher dropout rate(45% or above) than the university average

33 - 70% - Biofuel production technologies

9119 - 51% - Informatics engineering

9130 - 56% - Equinculture

9853 - 46% - Basic Education

# 5.2 HYPOTHESIS TESTING

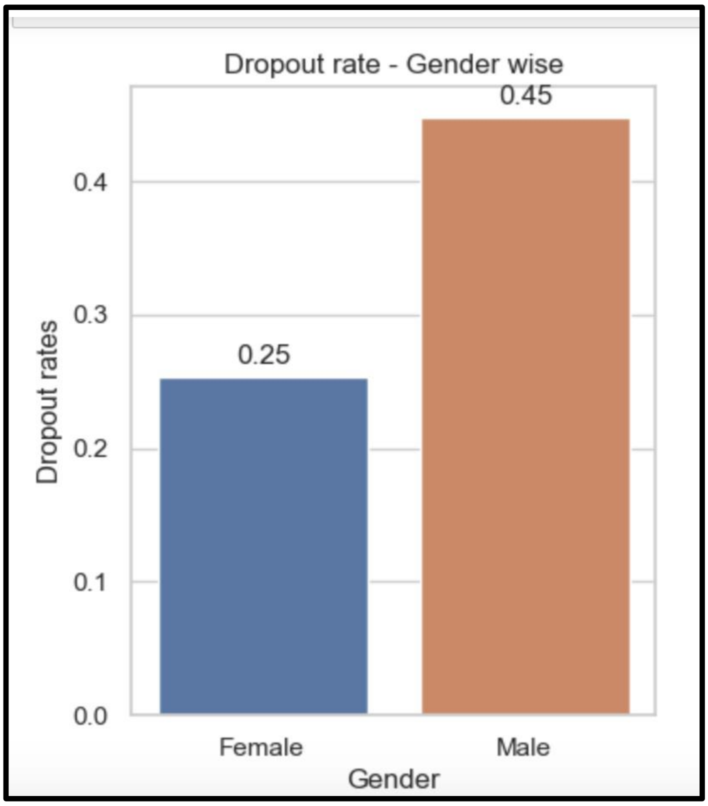



**Fig 5.2.1 Bar chart showing Dropout Rates w.r.t Gender**

**Drop out rates among male students are significantly different from female students. The rate for male students are almost double as that of female students.**

```
Chi-square statistic: 165.23501832801566
P-value: 1.3172604770142044e-36
There is a significant association between age group and dropout status.
```

### Father's Qualification Hypothesis

```
from scipy.stats import chi2_contingency
father_cross_tab = pd.crosstab(train["father_qual"],  train["Target"])

stat, p, dof, expected = chi2_contingency(father_cross_tab)

## interpret p value
alpha = 0.05 # 95% confidence level
if p < alpha:
    print(f"P value is {p}")
    print("Null hypothesis is rejected.")
else:
    print(f"P value is {p}")
    print("Failed to reject the Null hypothesis.")
```

```
P value is 0.00036598005260889 21
Null hypothesis is rejected.
```

**Here, the null hypothesis has been rejected which indicates that the dropout rate is dependent on Father's qualification.**

### Scholarship Holder

```
from scipy.stats import chi2_contingency
father_cross_tab = pd.crosstab(train["scholarship_holder"],  train["Target"])

stat, p, dof, expected = chi2_contingency(father_cross_tab)

## interpret p value
alpha = 0.05 # 95% confidence level
if p < alpha:
    print(f"P value is {p}")
    print("Null hypothesis is rejected.")
else:
    print(f"P value is {p}")
    print("Failed to reject the Null hypothesis.")
```

```
P value is 2.876831146824691e-59
Null hypothesis is rejected.
```

**Here, the null hypothesis has been rejected which indicates that it is dependent on whether the student is a scholarship holder or not.**

# 5.3 MODEL BUILDING

## NORMALIZATION USING MIN-MAX SCALING

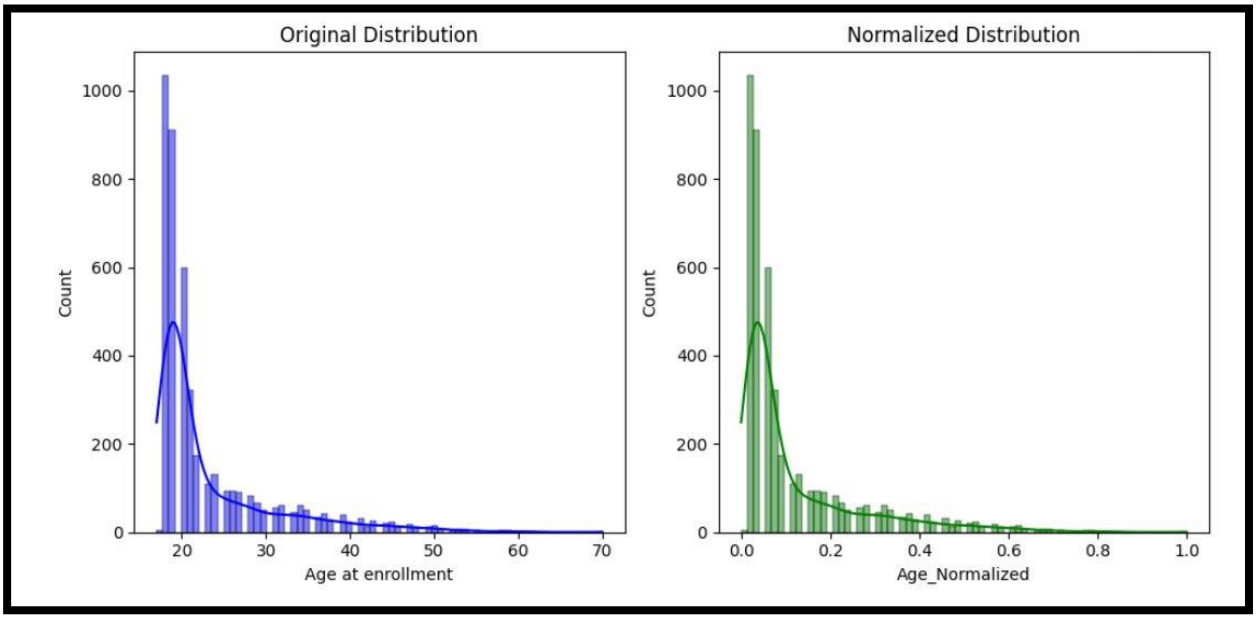



**Fig 5.3.1 Normalization of ages at enrollment**

**NORMALIZATION:** Normalization is a preprocessing technique used in machine learning to scale the features of a dataset within a certain range, typically between 0 and 1 or -1 and 1. It ensures that all features contribute equally to the learning process, preventing domination by features with larger scales.

# LOGISTIC CLASSIFICATION

```
F1 score on test set:  0.9132231404958677

              precision    recall  f1-score   support

     Dropout       0.93      0.84      0.88       284
    Graduate       0.90      0.96      0.93       442

    accuracy                           0.91       726
   macro avg       0.92      0.90      0.91       726
weighted avg       0.91      0.91      0.91       726
```

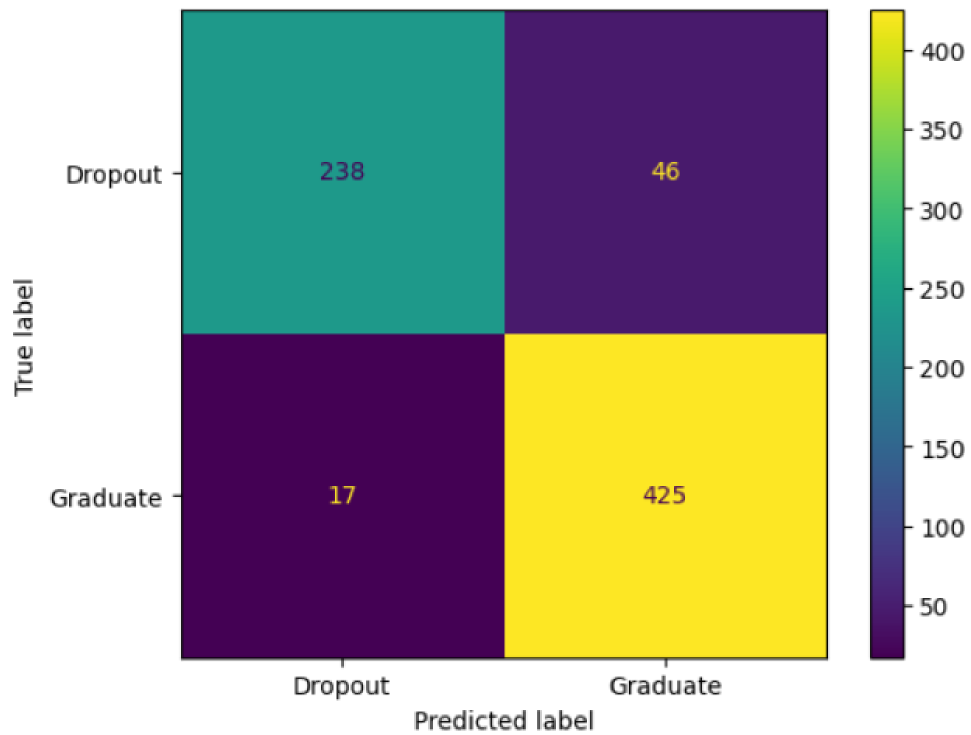



**Fig 5.3.2 Logistic Classification Model**

**LOGISTIC CLASSIFICATION:** Logistic classification is a statistical method used for binary classification tasks, predicting the probability of an event occurring based on input features. It employs the logistic function to map input variables onto a probability scale, making it a fundamental tool in machine learning.

**Here, we have applied a Logistic classification on the dataset. We have calculated the precision, recall, f1-score and the support as depicted in the figure. The F1-score on the test set is 0.9132**

# RIDGE CLASSIFICATION

```
F1 score on test set:  0.90633608815427

              precision    recall  f1-score   support

     Dropout       0.97      0.79      0.87       284
    Graduate       0.88      0.98      0.93       442

    accuracy                           0.91       726
   macro avg       0.92      0.88      0.90       726
weighted avg       0.91      0.91      0.90       726
```

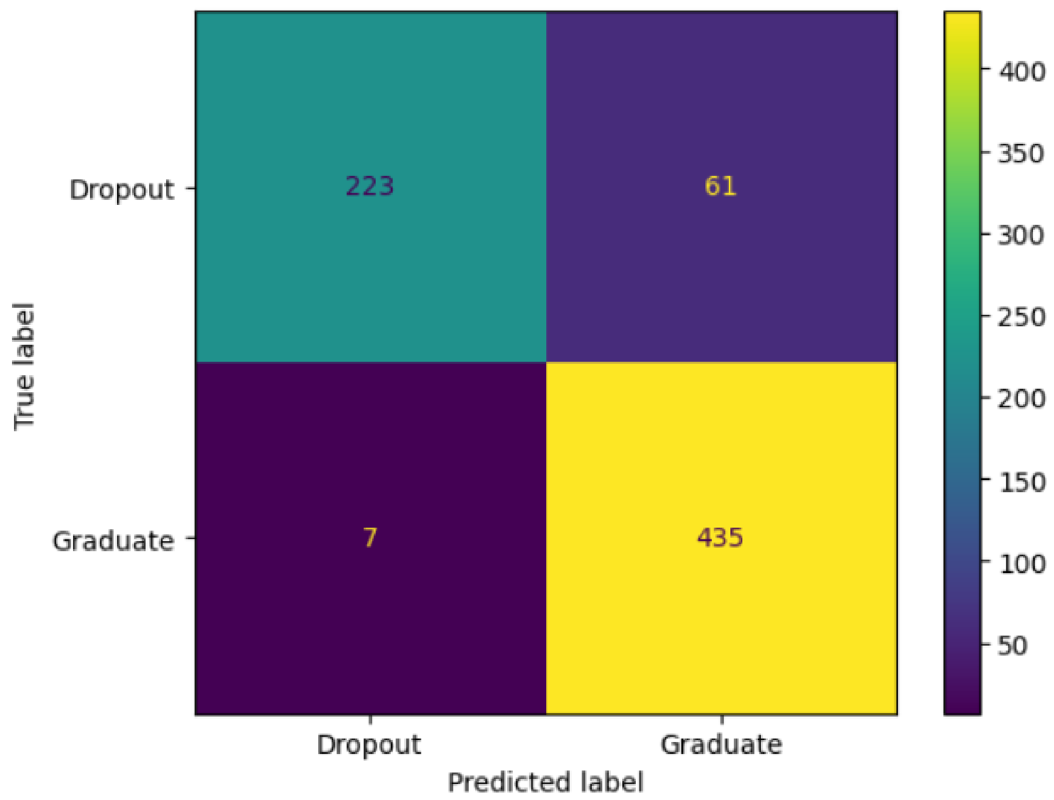



**Fig 5.3.3 Ridge Classification Model**

**RIDGE CLASSIFICATION:** Ridge classification is a variant of logistic regression that introduces regularization through the L2 penalty term, mitigating overfitting by shrinking the coefficients. This technique is particularly useful when dealing with multicollinearity in high-dimensional datasets, promoting stable and interpretable models.

**Here, we have applied a Ridge classification on the dataset. We have calculated the precision, recall, f1-score and the support as depicted in the figure. The F1-score on the test set is 0.9063**

## SUPPORT VECTOR MACHINE CLASSIFICATION

```
F1 score on test set:  0.9173553719008265

              precision    recall  f1-score   support

     Dropout       0.96      0.82      0.89       284
    Graduate       0.90      0.98      0.94       442

    accuracy                           0.92       726
   macro avg       0.93      0.90      0.91       726
weighted avg       0.92      0.92      0.92       726
```

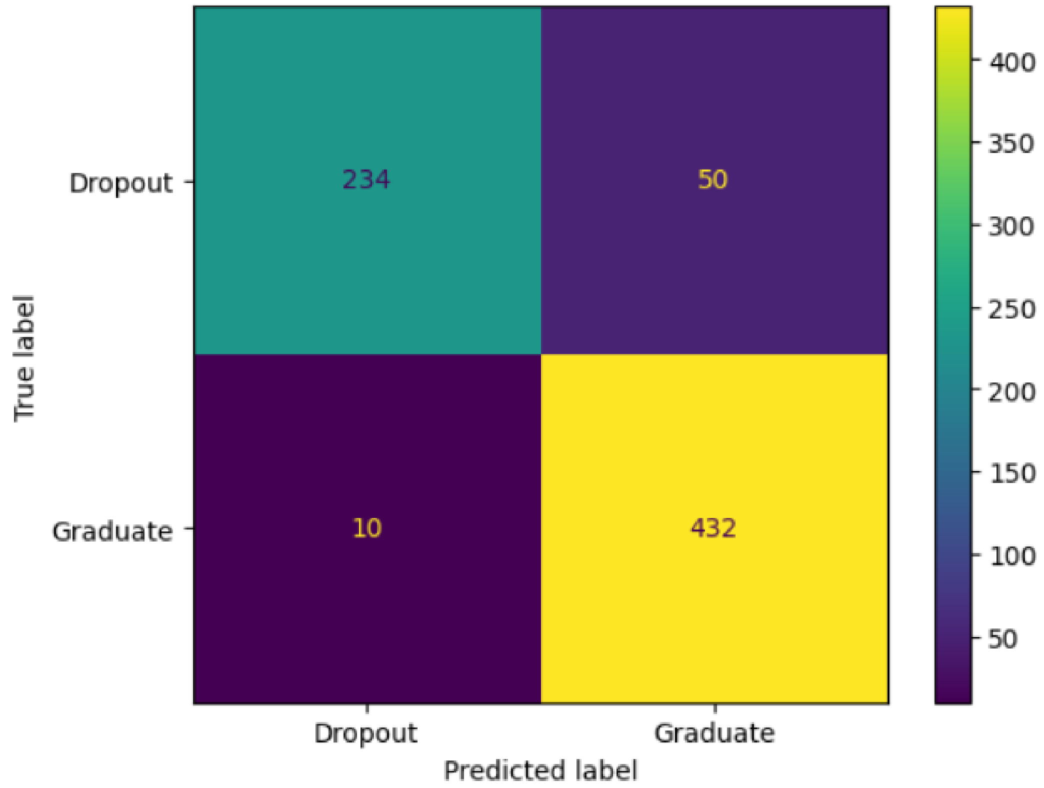



**Fig 5.3.4 Support Vector Machine Model**

**SUPPORT VECTOR MACHINE:** Support Vector Machines (SVM) are powerful supervised learning models capable of performing classification, regression, and outlier detection tasks. They work by finding the hyperplane that best separates data points into different classes, maximizing the margin between classes while minimizing classification errors.

**Here, we have applied a SVM classification on the dataset. We have calculated the precision, recall, f1-score and the support as depicted in the figure. The F1-score on the test set is 0.9173.**

## TREE CLASSIFIER

```
F1 score on test set:  0.8388429752066116

              precision    recall  f1-score   support

     Dropout       0.84      0.73      0.78       284
    Graduate       0.84      0.91      0.87       442

    accuracy                           0.84       726
   macro avg       0.84      0.82      0.83       726
weighted avg       0.84      0.84      0.84       726
```

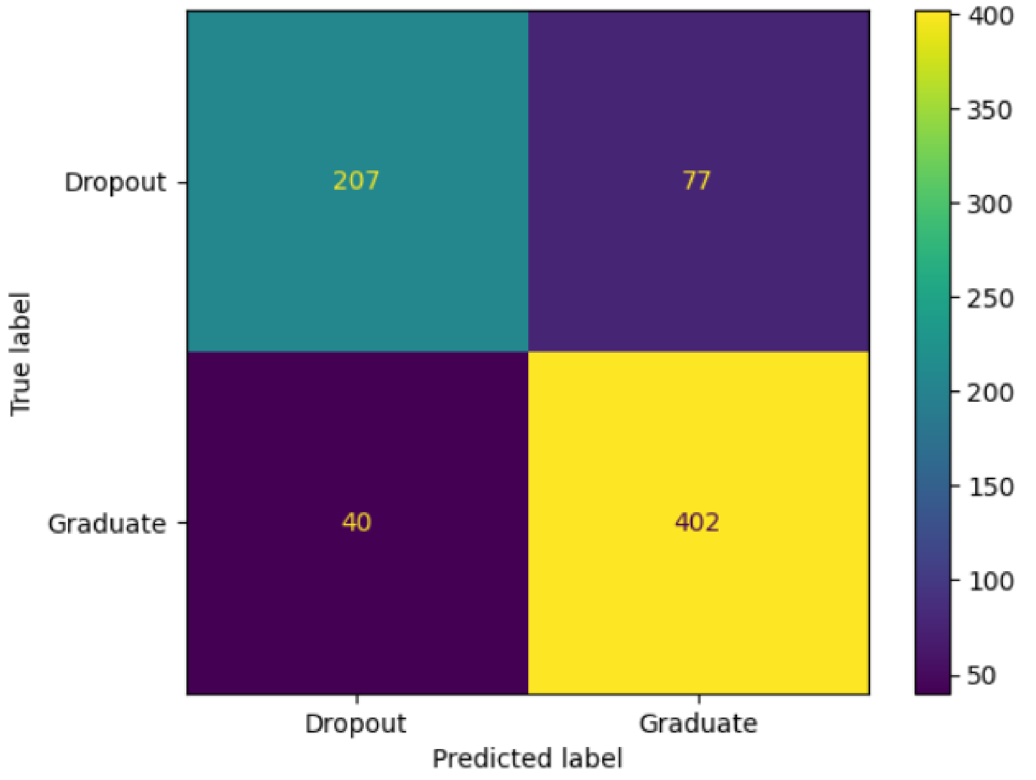



**Fig 5.3.5 Tree Classifier Model**

**TREE CLASSIFIER MODEL:** A tree classifier is a type of machine learning model that uses a hierarchical structure resembling a decision tree to classify instances. It recursively partitions the feature space based on certain criteria until it reaches leaf nodes representing the class labels.

**Here, we have applied a tree classifier on the dataset. We have calculated the precision, recall, f1-score and the support as depicted in the figure. The F1-score on the test set is 0.8388**

# RANDOM FOREST CLASSIFIER

F1 score on test set: 0.9118457300275482

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Dropout | 0.94 | 0.83 | 0.88 | 284 |
| Graduate | 0.90 | 0.97 | 0.93 | 442 |
| accuracy | | | 0.91 | 726 |
| macro avg | 0.92 | 0.90 | 0.91 | 726 |
| weighted avg | 0.91 | 0.91 | 0.91 | 726 |

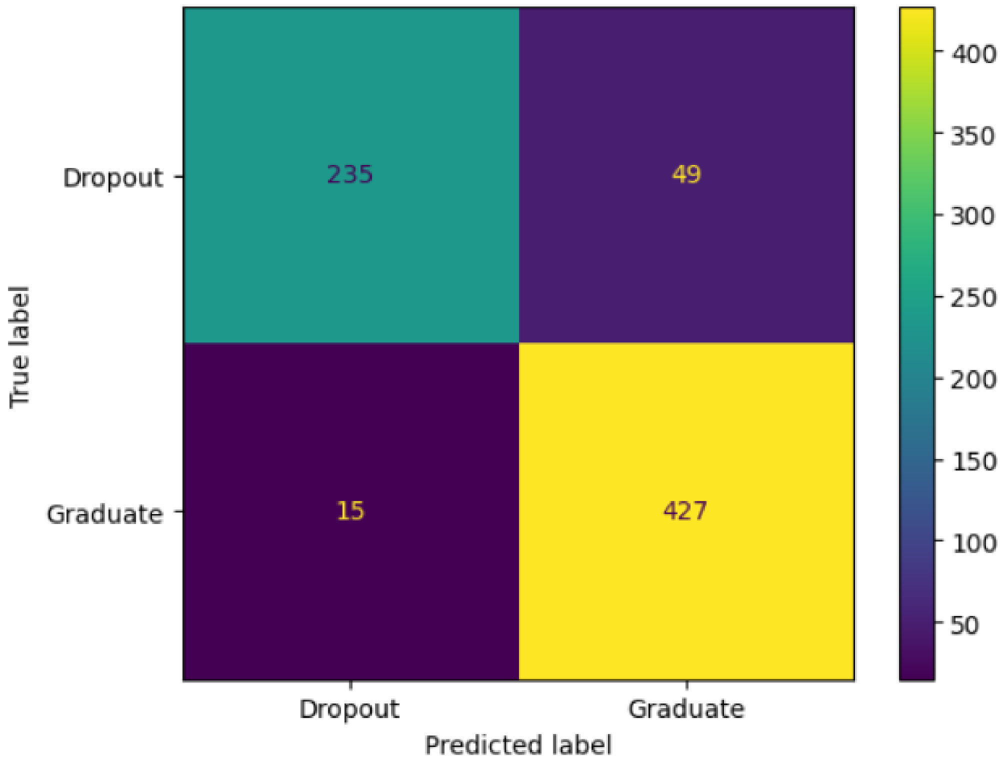
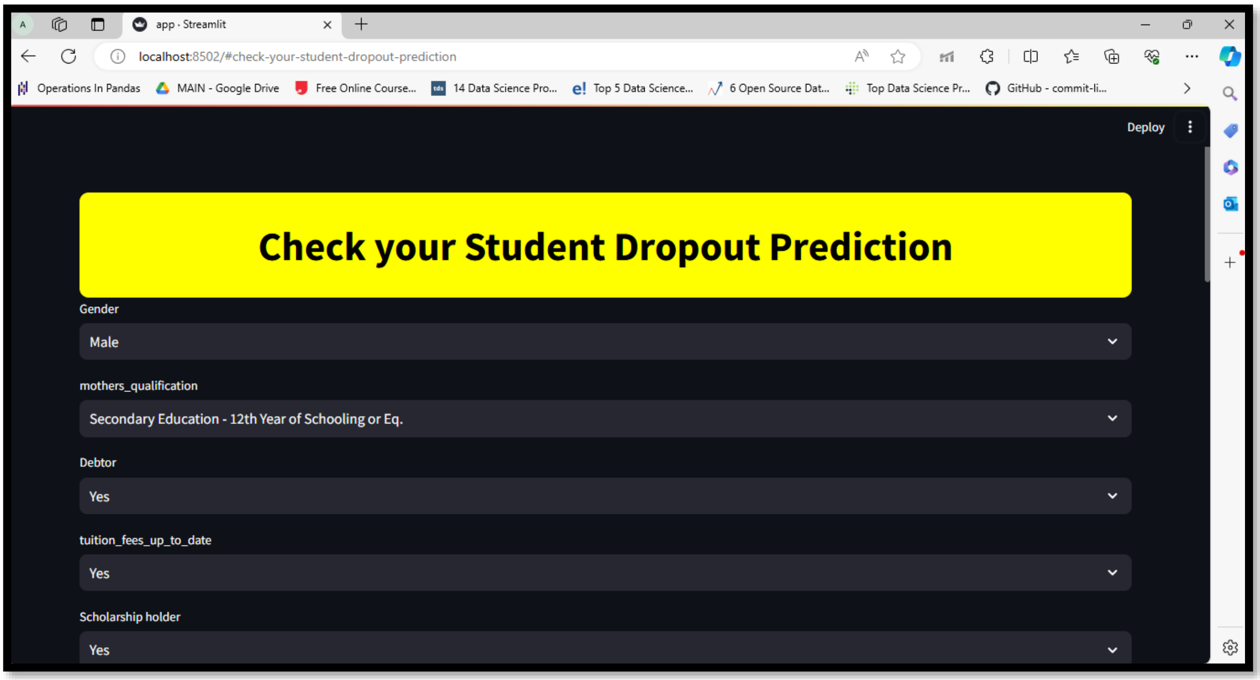


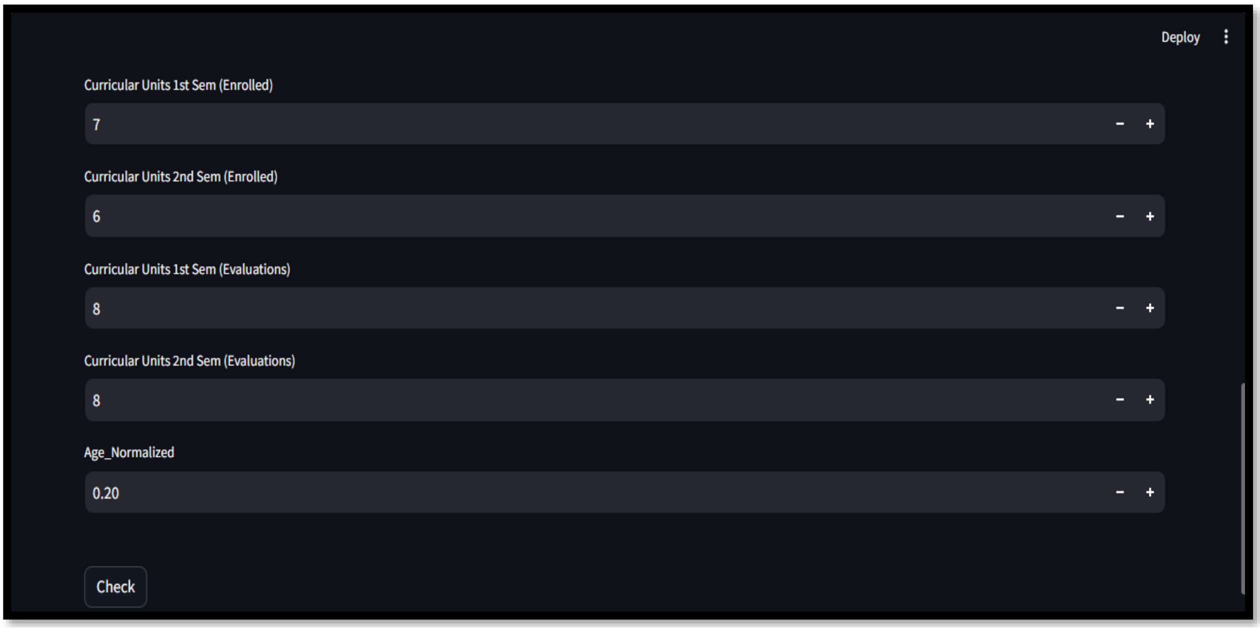

Fig 5.3.6 Random Forest Classifier

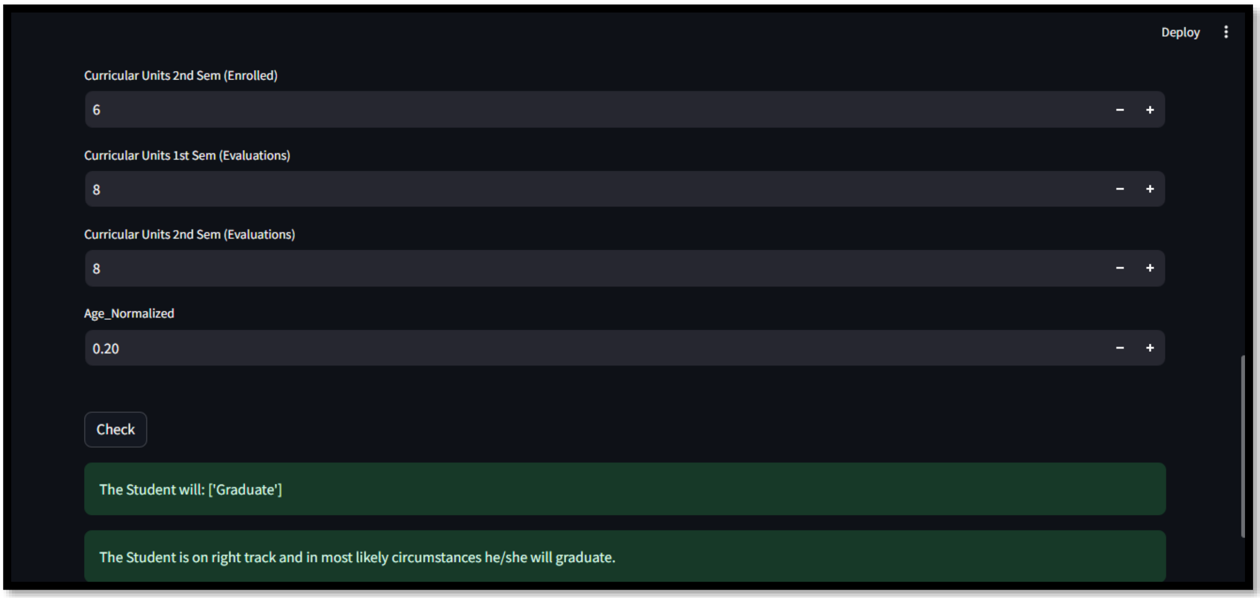
**RANDOM FOREST CLASSIFIER:** A random forest classifier is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes as the prediction. It introduces randomness in the tree-building process to improve performance and reduce overfitting.

**Here, we have applied a random forest classifier on the dataset. We have calculated the precision, recall, f1-score and the support as depicted in the figure. The F1-score on the test set is 0.9118**

# 6. TESTING AND DEPLOYMENT





Deploy
Curricular Units 1st Sem (Enrolled)
7
Curricular Units 2nd Sem (Enrolled)
6
Curricular Units 1st Sem (Evaluations)
8
Curricular Units 2nd Sem (Evaluations)
8
Age_Normalized
0.20
Check


Deploy
Curricular Units 2nd Sem (Enrolled)
6
Curricular Units 1st Sem (Evaluations)
8
Curricular Units 2nd Sem (Evaluations)
8
Age_Normalized
0.20
Check
The Student will: ['Graduate']
The Student is on right track and in most likely circumstances he/she will graduate.

# 7. VISUALIZATION USING POWER BI

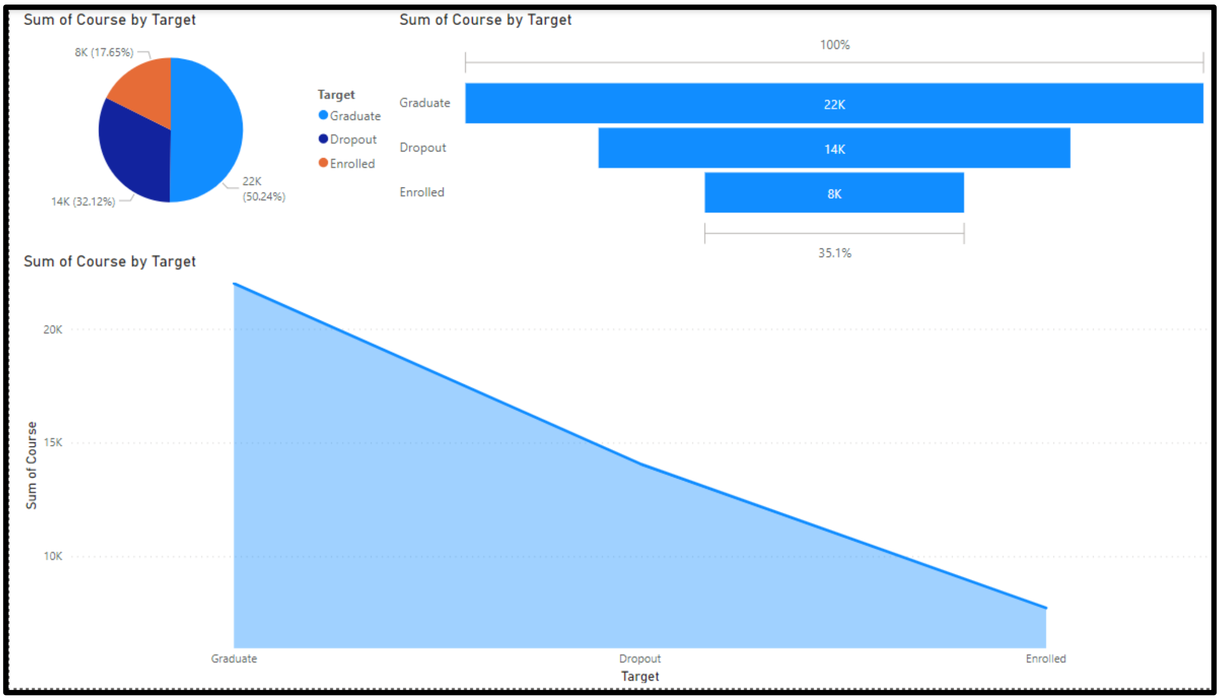



**Fig 7.1.1 Pie Chart and Area Chart**

**This diagram shows the number of students that have graduated, dropped out and are currently enrolled with the help of three different charts.**

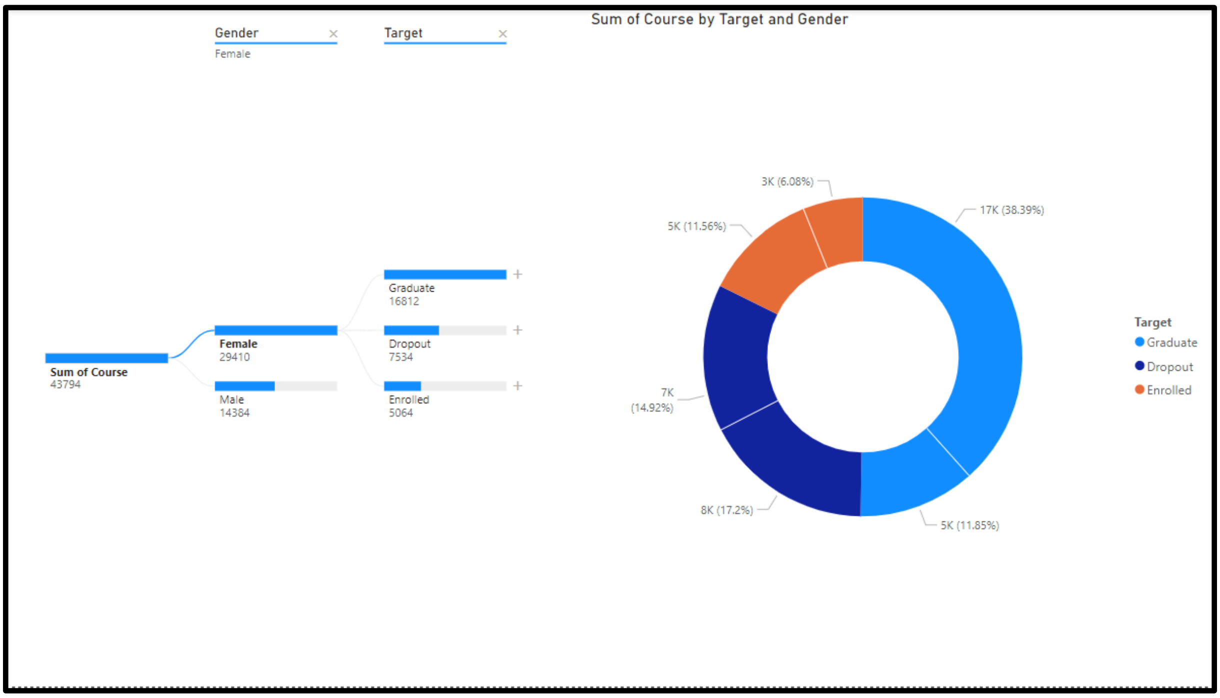



**Fig 7.1.2 Doughnut Chart**

**The chart shows male and female students who have graduated, dropped out, and are currently enrolled in courses**

| Gender | Dropout | Enrolled | Graduate | Total |
|---|---|---|---|---|
| **Female** | **7534** | **5064** | **16812** | **29410** |
| No | 6597 | 4087 | 9621 | **20305** |
| Yes | 937 | 977 | 7191 | **9105** |
| **Male** | **6532** | **2664** | **5188** | **14384** |
| No | 6119 | 2323 | 3740 | **12182** |
| Yes | 413 | 341 | 1448 | **2202** |
| **Total** | **14066** | **7728** | **22000** | **43794** |

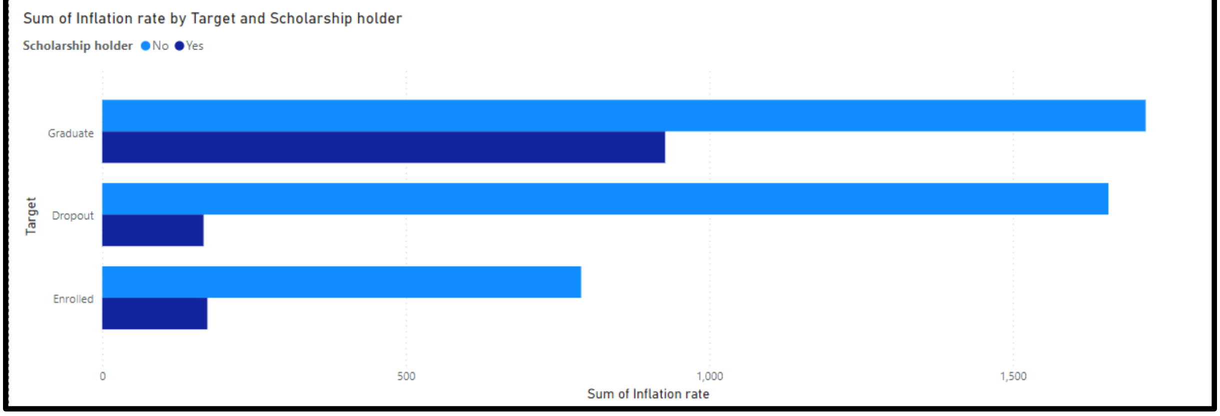



**Fig 7.1.3 Bar Chart**

**This diagram shows the scholarship holders in the category of graduate, dropout and enrolled.**

# 8. ANALYSIS AND RESULTS

## Key Insights

•32% of the students have dropped out and another 18% of students graduated late.

•Most of the students who drop out, do so before completing their first or second semester. Those who drop out later have a lower grade compared to others.

•The following 4 courses have very high drop out rates and require immediate attention:

1) Biofuel Production Technologies
2) Informatics Engineering
3) Equinculture
4) Basic Education

•Male students are at higher risk of dropping out than female students.

•Older students are more likely to dropout.

•Parents' backgrounds have an impact on graduation outcomes.

•The dropout rate among indebted students is very high(64%)

•Students joining after completing secondary school are the most successful. High dropout rates are observed among students who have joined after completing high school or a bachelor's degree.

# 9. CONCLUSION AND FUTURE ENHANCEMENTS

Student dropout is a complex and multifaceted issue that poses significant challenges to educational institutions and society at large. Through this analysis, we have identified key factors contributing to dropout rates, including academic difficulties, financial constraints, lack of engagement, personal issues, and institutional shortcomings. Addressing these factors requires a comprehensive and coordinated approach involving educators, policymakers, parents, and students.

Effective interventions, such as academic support programs, financial aid, mentorship, counseling services, and initiatives to foster an inclusive and engaging school environment, have been shown to reduce dropout rates and improve student success. Early identification of at-risk students and the provision of targeted support are crucial strategies in mitigating dropout rates and promoting positive educational outcomes.

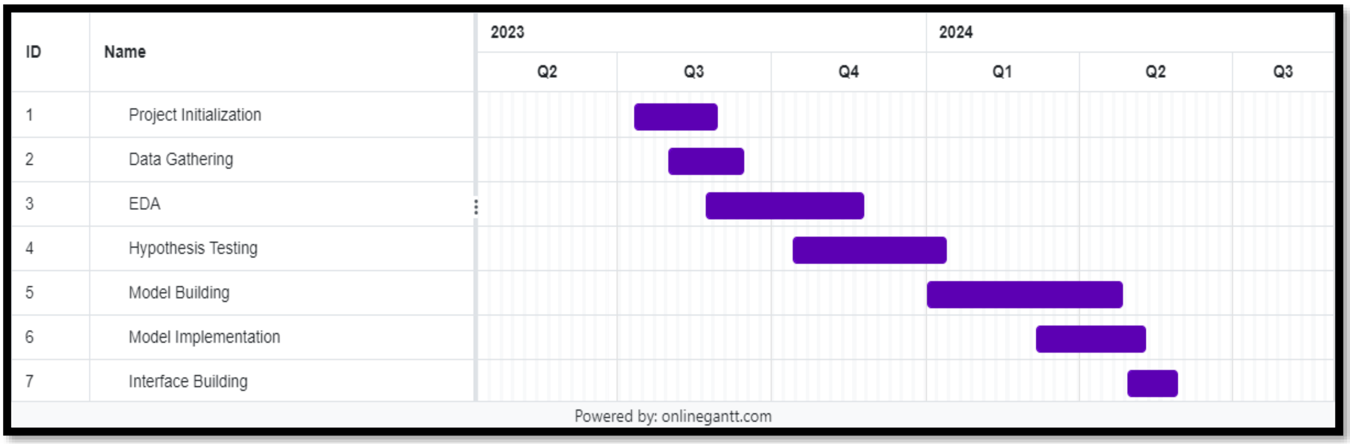
**Advanced Data Analytics**: Utilizing big data and machine learning techniques to predict at-risk students more accurately. This can help in providing timely interventions tailored to individual needs.

**Personalized Learning Plans:** Developing personalized learning plans that cater to the unique strengths and weaknesses of each student, thus keeping them engaged and motivated.

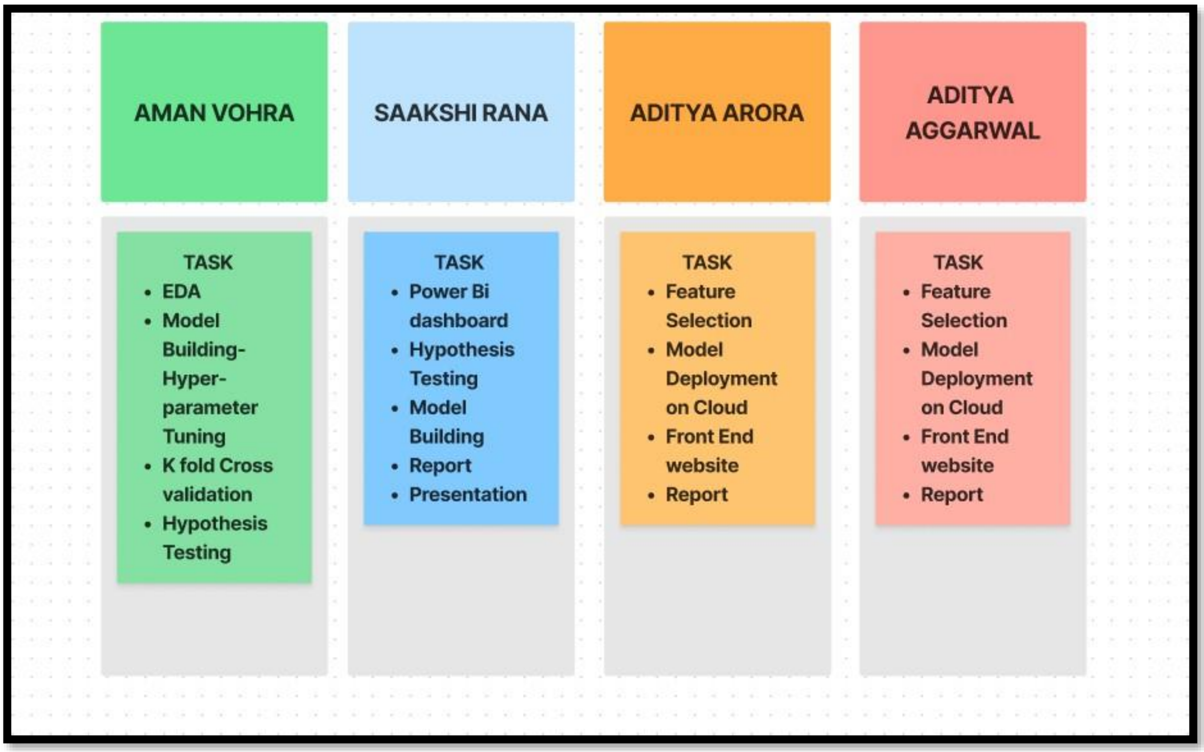
**Increased Financial Support**: Expanding scholarship programs, grants, and other financial aid options to ensure that financial constraints do not hinder students' ability to complete their education.

**Enhanced Support Services**: Strengthening mental health and counseling services within educational institutions to help students navigate personal and academic challenges more effectively.

# GANTT CHART




# RESPONSIBILITY CHART



# 10.REFERENCES

[1] Altonji, J.G. 1993. "The Demand for and Return to Education when Education Outcomes are Uncertain." Journal of Labour Economics, vol. 11, Part I, 48-83.

[2] Chuang, H.L. 1997. "High School Youths' Dropout and Re-Enrollment Behavior." Economics of Education Review, vol. 16, no. 2, 171-186.

[3] Chaplin, D.D., M.D. Turner and A.D. Pape. 2003. "Minimum Wages and School enrolment of Teenagers: A Look at the 1990's." Economics of Education Review, vol. 22, 11-21.

[4] Eckstein, Z. and K.I. Wolpin. 1999. "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities." Econometrica, vol. 67, no. 6, 1295-1339.

[5] Ferrer, A. and D. Lauzon. 2005. The Role of Employment Opportunities and Earnings in the Decline of the High School Dropout Rate Over the 1990s. Manuscript. UBC.

[6] Doll, J.J., Eslami, Z., & Walters, L. (2013). Understanding why students drop out of high school, according to their own reports. SAGE Open, 3 (4), 1-15. doi: 10.1177/2158244013503834

[7] BAYER, A. The college drop-out: factors affecting senior college completion. Sociology of Education, 1968,41, 305–316.

[8] BROWN, F. G. Identifying college drop-outs with Minnesota Counselling Inventory.Personnel and Guidance Journal, 1960,39, 280–282.

[9] DANIEL, K. L. B.A study of dropouts at the University of Alabama with respect to certain academic and personality variables. (Doctoral dissertation, University of Alabama) Ann Arbor, Mich.: University Microfilms, 1963. No. 64-9118.

## 9% Overall Similarity

Top sources found in the following databases:

- 7% Internet database
- 5% Publications database
- Crossref database
- Crossref Posted Content database
- 7% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

| | | |
|---|---|---|
| 1 | dsp-psd.tpsgc.gc.ca<br>Internet | 1% |
| 2 | fdocuments.in<br>Internet | 1% |
| 3 | scribd.com<br>Internet | <1% |
| 4 | Tilburg University on 2024-05-26<br>Submitted works | <1% |
| 5 | Taylor's Education Group on 2024-02-17<br>Submitted works | <1% |
| 6 | UT, Dallas on 2023-12-02<br>Submitted works | <1% |
| 7 | researchgate.net<br>Internet | <1% |
| 8 | pdfs.semanticscholar.org<br>Internet | <1% |

| | | |
|---|---|---|
| 9 | ijisae.org<br>Internet | <1% |
| 10 | dokumen.tips<br>Internet | <1% |
| 11 | lib.dr.iastate.edu<br>Internet | <1% |
| 12 | University of Hertfordshire on 2023-05-11<br>Submitted works | <1% |
| 13 | "Machine Learning for Cyber Physical System: Advances and Challeng...<br>Crossref | <1% |
| 14 | American University of the Middle East on 2024-01-07<br>Submitted works | <1% |
| 15 | Liverpool John Moores University on 2023-04-21<br>Submitted works | <1% |
| 16 | Tilburg University on 2024-05-20<br>Submitted works | <1% |
| 17 | Volunteer State Community College on 2024-03-27<br>Submitted works | <1% |
| 18 | camudigitalcampus.com<br>Internet | <1% |
| 19 | 1login.easychair.org<br>Internet | <1% |

## Excluded from Similarity Report

- Quoted material
- Cited material
- Small Matches (Less then 10 words)
- Manually excluded sources
- Manually excluded text blocks

EXCLUDED SOURCES

| | |
|---|---|
| **www150.statcan.gc.ca**<br>Internet | 2% |
| **statcan.gc.ca**<br>Internet | 2% |
| **statcan.ca**<br>Internet | 2% |
| **link.springer.com**<br>Internet | 2% |
| **core.ac.uk**<br>Internet | 1% |
| **scholarworks.montana.edu**<br>Internet | <1% |

EXCLUDED TEXT BLOCKS

**END TERM VII SEMESTER SYNOPSIS REPORTSubmitted in partial fulfillment of th...**
northcap on 2024-03-06

**Under the supervision ofDr**
The NorthCap University, Gurugram on 2021-10-22

**2024**

The NorthCap University, Gurugram on 2024-03-13

**to The NorthCap University, Gurugram, India, is a record ofbona fide synopsis work...**

northcap on 2024-03-06

**appreciation to the faculty and staff of The**

University of Bolton on 2023-01-25

**Various machine learning algorithms such as logistic regression, decision trees, ra...**

University of Economics & Law on 2024-03-24