



University of New Haven

TAGLIATELA COLLEGE OF ENGINEERING

Electrical & Computer Engineering and Computer Science

Electrical & Computer Engineering & Computer Science (ECECS)

# PerformaPRED

A Student Academic Performance Predictor



SPRING 24

# CONTENTS

Project Name .....2

Executive Summary .....2

Technical Report.....3

Highlights of Project.....3

Submitted on .....3

Abstract .....4

Methodology .....5

Results Section .....6

Discussion.....7

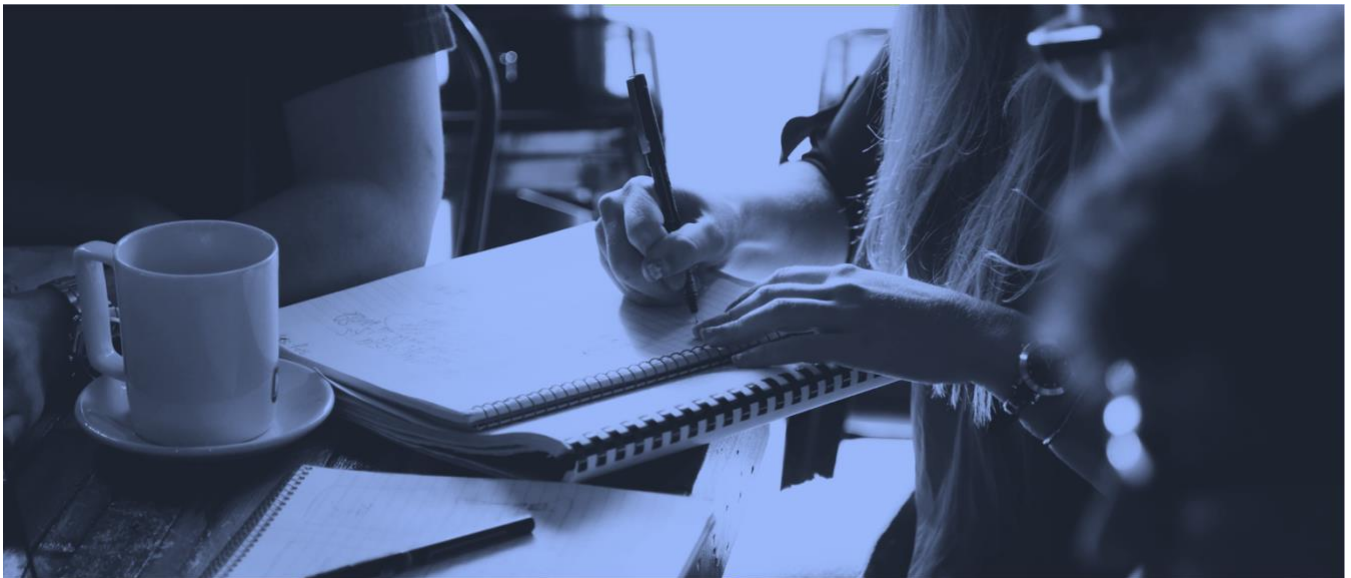
Conclusion .....8

Contributions/References.....8

## Executive Summary

The goal of this research project is to forecast a student's likelihood of passing the final exam using Machine Learning techniques. The UCI Machine Learning Repository provided the data set that was utilized.

**Dataset:** StudentsPerformance.csv



### Team Members:

Name 1: VEERA SWAMY G

Name 2: Revanth Myneni

Name 3: Satya Sai Sri Nikhil N

### Questions?

[Vgatt6@unh.newhaven.edu](mailto:Vgatt6@unh.newhaven.edu)

[Smyne2@unh.newhaven.edu](mailto:Smyne2@unh.newhaven.edu)

[Snamu1@unh.newhaven.edu](mailto:Snamu1@unh.newhaven.edu)

## *PerformaPRED*

*A Student Academic Performance  
Predictor*

### Highlights of Project

Right now, the educational systems require creative approaches to raise the standard of instruction in order to maximize outcomes and lower the failure rate.

We wanted to work on this difficult issue because, as IT department students who have learned a little bit about machine learning over the past month, we know that deep analysis of learners' past records can play a crucial role in an institute's ability to give quality education to learners.



## Abstract

The effect of the global COVID-19 epidemic on educational systems and student outcomes is the subject of this study. After the outbreak, academic attainment has declined, and it is critical to identify the contributing elements and investigate workable solutions. This research is driven by the desire to improve education and lower failure rates by introducing novel approaches, especially by utilizing machine learning techniques.

The project intends to show the value of in-depth examination of students' prior records in promoting better educational outcomes, building on the understandings obtained from the IT department's recent machine learning experimentation. Our aim is to provide significant perspectives and remedies for the contemporary issues that affect educational establishments across the globe, by means of this research.

## Cover Page

This study explores how educational systems and student performance are affected by the global COVID-19 pandemic. It makes novel solutions, especially by utilizing machine learning techniques, and aims to comprehend the complex problems that the epidemic presents. The research emphasizes the vital role that in-depth examination of students' past records plays in raising the standard of instruction, taking inspiration from previous IT department studies. Through this study, the authors hope to add important new information to the current conversation on educational reform in the aftermath of extraordinary upheavals around the world.

## Introductory Section

Education is an important element of the society, every government and country in the world work so hard to improve this sector. With the corona-virus outbreak that has disrupted life around the globe in 2020, the educational systems have been affected in many ways; studies show that student's performance has decreased since then, which highlights the need to deal with this problem more seriously and try to find effective solutions, as well as the influencing factors.

## Review of available research

Existing research on the repercussions of the COVID-19 pandemic on education underscores significant disruptions, particularly noting a decline in student performance. The literature emphasizes the pressing requirement for creative solutions to improve the quality of education and diminish failure rates. Importantly, a growing body of evidence indicates that machine learning holds potential in achieving these goals. The examination of learners' historical records through sophisticated data-driven techniques emerges as a promising approach to tackle the evolving challenges in education. In the face of unprecedented changes, this literature review advocates for a comprehensive and forward-thinking strategy, highlighting the incorporation of technology, data-centric policies, and inventive methodologies to adeptly navigate the complexities of the post-pandemic educational landscape.

## Methodology

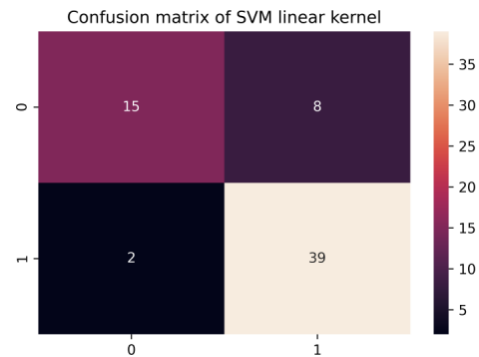
The approach utilizes Support Vector Machines (SVM) and Logistic Regression algorithms for the analysis and prediction of educational outcomes using learners' historical data. SVM is applied to recognize classification patterns, and Logistic Regression is utilized to model the likelihood of student success. This method provides a holistic framework for evaluating and improving educational performance.



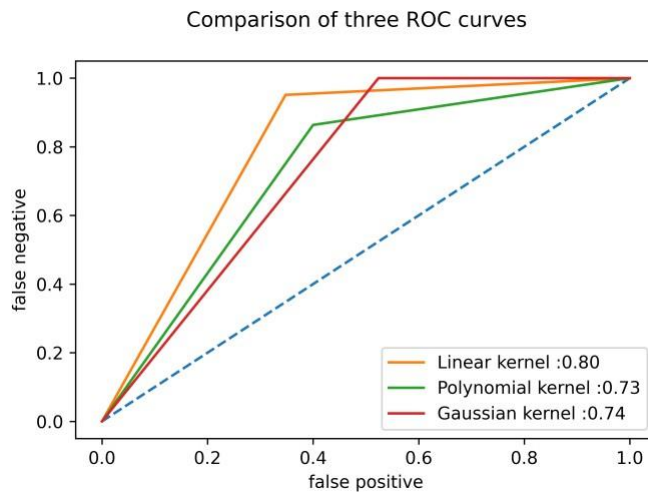
## Results Section

After Training our first svm model and confusion matrix, here are our results:

Metric	Value
training time	10ms
accuracy	84.0 %
f1 score	0.82
The roc_auc_score	0.8

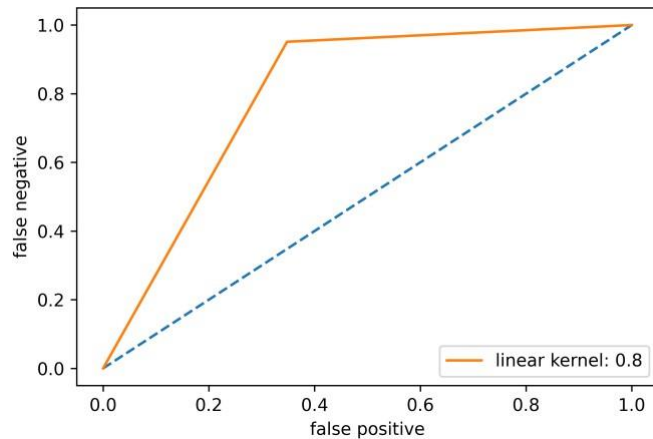


Now after we train our three models, here the ROC plot:



### The more accurate svm classifier:

- **The training time:** 11ms
- **The accuracy:** 84.375 %
- **The f1 score:** 0.82
- **The roc\_auc\_score:** 0.8



## Discussion

The results section is followed by the **discussion section**, where you craft your main arguments by building on the results you have presented earlier.

The "discussion section" is where you rely on the power of narrative to enable numbers to communicate your thesis to your readers. You refer the reader to the research question and the knowledge gaps you identified earlier. You highlight how your findings provide the ultimate missing piece to the puzzle.

Of course, not all analytics return a smoking gun. At times, more frequently than I would like to acknowledge, the results provide only a partial answer to the question and that, too, with a long list of caveats.



## Conclusion

The improvement of the educational system is a major issue. As engineering students, we can contribute to this goal by utilizing resources and technologies, such as machine learning materials, to develop creative solutions that will benefit underprivileged students, particularly those who live in challenging environments. circumstances (social, educational, and demographic concerns). Our objective for this project was to develop a model that uses many features to forecast students' state. Our main goals were to determine the optimal classification method and the most important variables influencing pupils' academic standing to give them an overview, or valedictorian, of the ideal circumstances for success.

In order to predict student performance and difficulty for this project, we used a variety of classification techniques, including logistic regression and SVM. We also evaluated this model using a variety of metrics, including the confusion matrix, roc curve, and f1 score. In the end, SVM emerged as the best method with an 80% precision rate when compared to other algorithms. Before diving into our primary problems, there were a few things to consider: three algorithm comparisons, model implementation, data processing, and data visualization.

## Contributions/References

**Algorithmic Framework:** Presenting a strategy that provides a reliable way for probability modeling and classification by combining Support Vector Machines (SVM) and Logistic Regression algorithms for student performance prediction analysis.

**Improvement of Education:** By offering insights into how machine learning might be used to evaluate past data in order to forecast and improve educational outcomes, it will add to the current conversation about creative approaches to raising academic standards.

<https://ieeexplore.ieee.org/abstract/document/8079967>

**Our GitHub:** [nikhilnamuduri/dsci6007project \(github.com\)](https://github.com/nikhilnamuduri/dsci6007project)

-----Thank You-----