RomeC-137 / ML-phase-3

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

my first machine learning application to non-dummy data

0 stars 0 forks 0 watching Branches Activity
Tags

Public repository

1 Branch 0 Tags Go to file t Go to file Add file Code

RomeC-137 Create .gitignore 48df7b9 · 9 minutes ago

.gitignore	Create .gitignore	9 minutes ago
ML project.ipynb	finalnotebook	9 minutes ago
README.md	Create README.md	2 days ago
Student_Performance_Classifi...	Finalised	yesterday
svm.ipynb	awaiting comparison	2 days ago

README

🎓 Student Performance Classification Project

🧠 Project Overview

This project addresses a **classification problem** using a dataset of student performance. The main goal is to identify students who are at risk of failing, enabling school administrators to take early action and provide support.

📌 Business Problem

School administrators need a reliable method to identify students likely to fail their final grade (defined as G3 < 10). Early identification allows for targeted academic intervention, reducing dropout rates and improving overall performance.

👥 Stakeholders

The primary stakeholders are:

- School management teams
- Academic support staff and counselors
- Teachers and intervention specialists

📊 Dataset Summary

The dataset includes the following features:

- **Demographics**: age, family background, parents' education
- **Academic Behavior**: study time, absences, failures
- **Grades**: first (G1), second (G2), and final (G3) period grades

⚙️ Modeling Process

The classification task was addressed using the following models:

- Logistic Regression
- Support Vector Machine (SVM)
- Random Forest

Preprocessing Steps:

- Feature selection and encoding
- Train/test split
- Data normalization where required

Model Evaluation:

Evaluated models using metrics suitable for classification:

- Accuracy
- Precision
- Recall
- F1 Score

✅ Results and Recommendation

Among the models tested, the **Random Forest** model showed the best performance overall with strong accuracy and interpretability. Important predictors included:

- Previous grades (G1 and G2)
- Study time
- Absences

🚀 Recommendations

- Implement the Random Forest model in school monitoring systems.
- Alert educators and counselors about at-risk students early.
- Design personalized interventions based on the predictions.

📁 Project Files

- `Student_Performance_Classification_Presentation.pptx` : Non-technical presentation for stakeholders.
- `Final_Chronological_ML_Notebook.ipynb` : Full code from data loading to model evaluation.
- `Student_Performance_Classification_Presentation.pdf` *(optional for export)*

🙏 Acknowledgments

Thanks to the school administrators and educators supporting this initiative to improve student success through data-driven insights.

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

- **Jupyter Notebook** 100.0%