University of Kurdistan Hewlêr (UKH)

Department of Computer Science and Engineering

Software Engineering UGIII

SE301-Special Topics in Software Engineering

AY 2024-2025 / Semester II

End-to-End Machine Learning Project with GUI, GitHub Copilot, and Clean Code Practices

Team Members:

Niyan Mahmoud, Sidra Ayoob, Yad Danar

Lecturer:

Prof. Ibrahim Ismael Hamarash

Submission Date:

June 14th, 2025

Contents

Introduction

This project explores how online learning has impacted students academic performance, focusing on mental health and lifestyle factors like stress, anxiety, and screen time. We used real survey data to train a machine learning model that can predict changes in academic performance based on these inputs. To make the model accessible, we also built a simple user interface where users can enter data and get predictions instantly.

Dataset Description

The dataset we used contains responses from students about their mental health experiences during the shift to online learning. The information was collected through surveys and focuses on different emotional, psychological, and behavioral changes that students went through while adapting to remote education. It looks at things like stress levels, anxiety before exams, screen time, and changes in academic performance.

This dataset is really useful for analyzing how online learning environments may have affected student well-being. It can be used for various purposes such as data visualization, exploratory data analysis (EDA), and machine learning projects. By studying the patterns and relationships in the data, we can gain insights into how factors like screen time or anxiety are linked to academic performance.

There are 1,000 entries in the dataset, each representing a student's response, and it includes 10 different columns. These columns cover demographic details, mental health indicators, and lifestyle habits. Overall, the dataset offers a good foundation for understanding how remote education has impacted students mental health and learning outcomes.

Source: Kaggle

Filename: *Student Mental Health Analysis During Online Learning.csv*

Columns:

- Name
- Gender
- Age
- Education Level
- Screen Time (hrs/day)
- Sleep Duration (hrs)
- Physical Activity (hrs/week)
- Stress Level
- Anxious Before Exams
- Academic Performance Change (target)

Tools and Requirements

Component	Tool Used
Language	Python
IDE	Jupyter Notebook, VS Code
ML Library	Scikit-learn
GUI Framework	Streamlit
Version Control	Git & GitHub
Environment Mgmt.	requirements.txt
Collaboration	GitHub Copilot

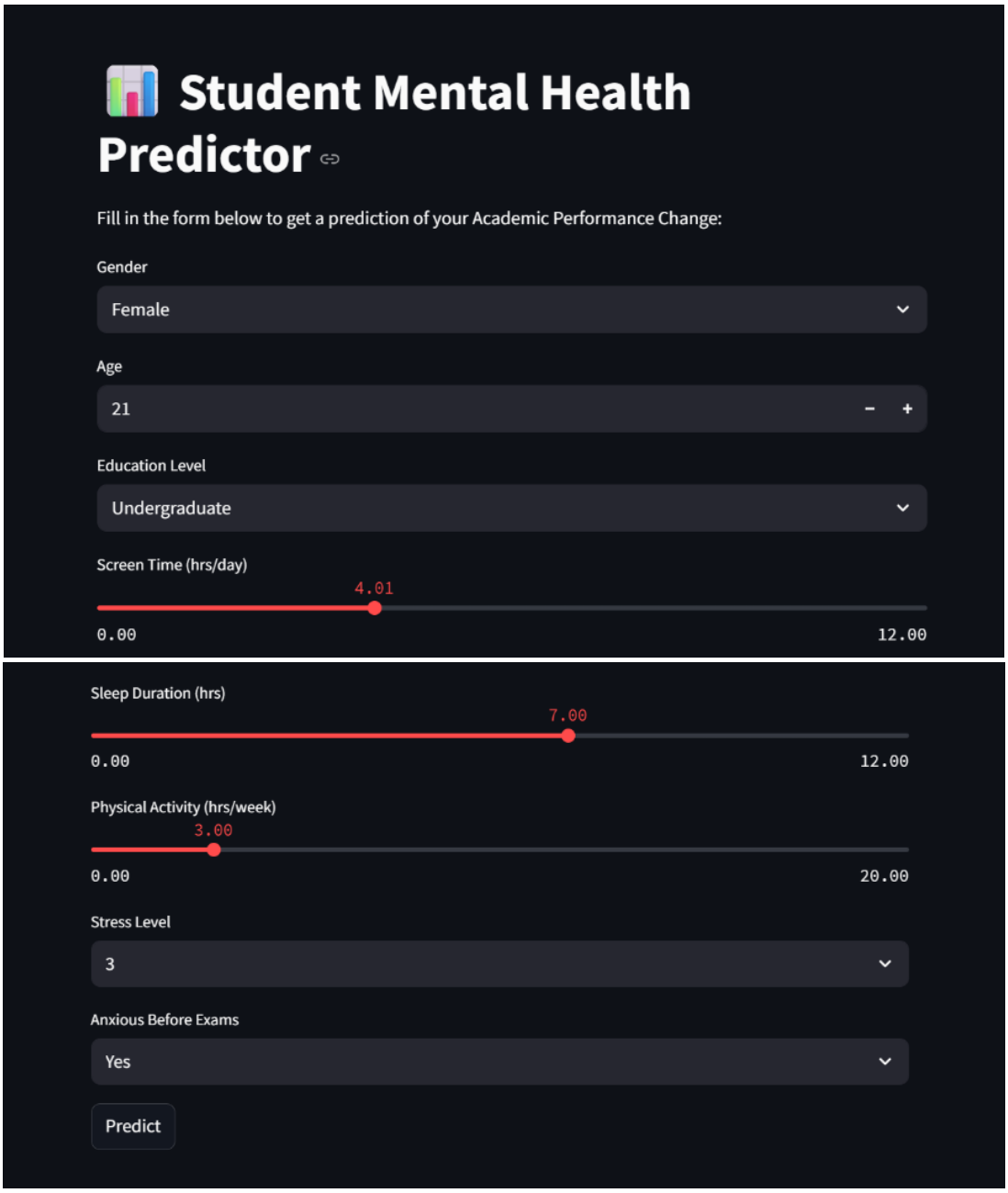
Model Training

In our *train_model.py* file, we first load and clean the dataset, then handle things like converting text-based (categorical) data into numbers and scaling the numerical values. After that, we split the data into training and testing sets. Once everything is ready, we train the model and save it as a .pkl file called *student_model.pkl*, so it can be used later in the GUI.

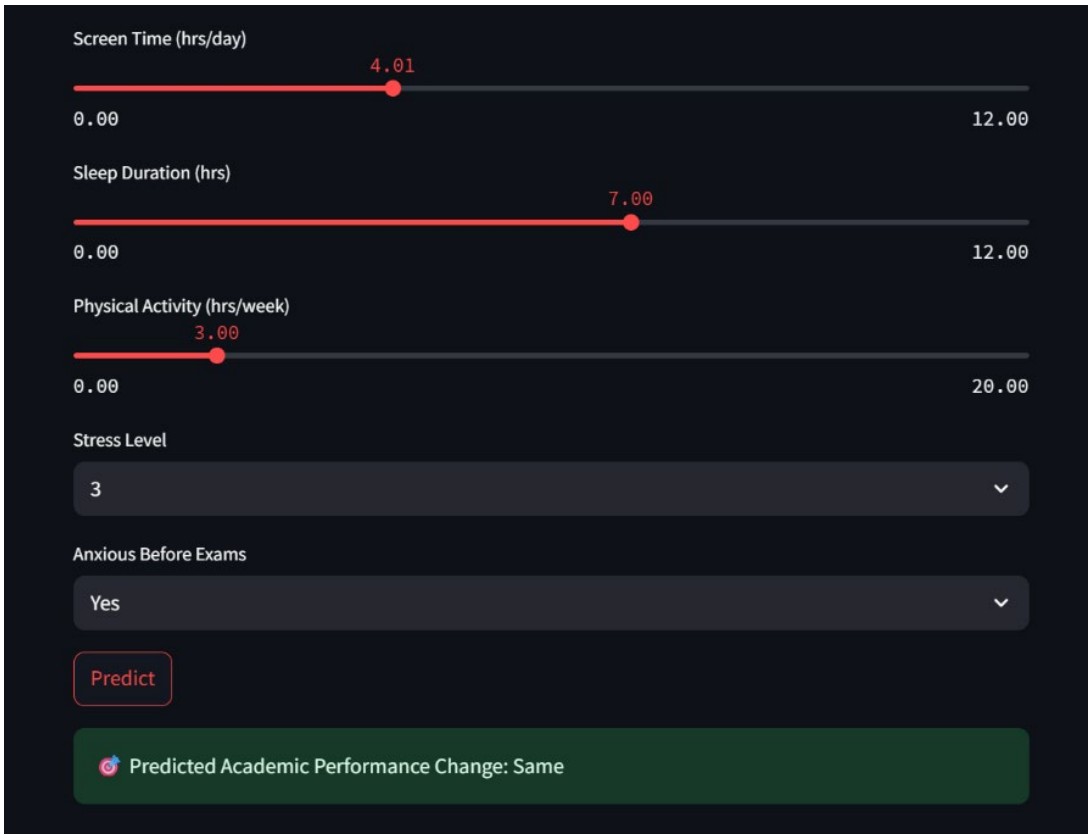
Streamlit App

We built a simple web app using Streamlit (in the *gui_app.py* file) that lets users input student information, like screen time, stress level, and anxiety before exams. Once the user fills in the form, the app uses our trained machine learning model to predict whether the student's academic performance is likely to improve, stay the same, or decline. Behind the scenes, the app loads the saved model file (*student_model.pkl*) and applies the same preprocessing steps that we used during training, like encoding and scaling, to make sure the prediction is accurate and consistent with how the model was trained. This makes the app easy to use and gives quick, real-time results.

GUI Screenshots



Result:



Conclusion

Overall, we were able to complete a full end-to-end machine learning project that ties together everything we've learned. We started by working with real-world data about students mental health during online learning, which we cleaned and preprocessed to make it ready for analysis. After that, we trained a machine learning model that performed well in predicting changes in academic performance. Finally, we created a simple and interactive GUI using Streamlit, so users can input student data and instantly get a prediction. This project helped us understand how all the different steps from data preprocessing to model training and deployment come together to form a complete ML application.