Progress Report 1

Capstone Project: Gaming Addiction and Mental Health

Course Instructor: Dr. Humera Noor Minhas

Submission By: Frank Stalin Dsouza

Submission Date: May 6, 2025

1. Research and Background Study

Based on existing research and insights from The World Health Organization (WHO), the following factors were identified as indicators of Gaming Addiction:

- Playing more than 5 hours a day
- Getting less than 6 hours of sleep
- Avoiding social interactions
- Drop in school/work performance
- High stress or anxiety levels

2. Data Collection (Survey & Existing Data)

To build my gaming addiction prediction model, I'm gathering data in two ways: First, I created a Google Forms survey with 15 questions (columns) about gaming habits (like daily playtime, favorite games es), mental health (stress, sleep, mood), and lifestyle (school/work, social life). So far, people have filled it out mostly teens and young adults and early responses show people play over 4 hours daily, with many reporting sleep issues or anxiety. Second, I'm using existing datasets from Kaggle that track similar behaviors, which I'll merge with my survey data later.

Here is the snapshot of the google form survey:

What is your age?	What is your gender?	What is your occupation?	How many hours do you play video games p	How many days per week do you play game	What type of games do you mostly play?	Do you often lose track of time while gaming
23	Prefer not to say	Student	1	4	Other	Yes
22	Female	Intern sdet	1	1	Other	Yes
23	Male	Student	1	3	Simulation	Yes
23	Male	Student	1	1	Action/Adventurers	No
23	Male	Manager at a Gaming Cafe	1	4	Other	No

On a scale of 1 to 10, how stressed do you f	On a scale of 1 to 10, how anxious do you fe	On average, how many hours of sleep do yo	How often do you feel socially withdrawn or	Have you ever felt guilty or depressed after l	Do you think gaming helps you cope with str	Do you consent to your responses being use
3	2	7 hours	Rarely	No	Yes	Yes
4	4	8	Never	No	No	Yes
5	5	5	Sometimes	Yes	Yes	Yes
1	1	8	Never	No	Yes	Yes
7	3	7-8 hrs	Sometimes	No	Yes	Yes

3. Initial Data Check

After collecting survey responses and downloading existing datasets, I did a quick check to see if everything looked right. I noticed that most gamers in my survey play 2-4 hours daily, and about 1 in 9 play more than 4 hours, which matches what other studies say. The Kaggle datasets also looked clean, with clear numbers on gaming time, sleep, and mood. I have observed that people who play competitive online games (like Fortnite or League of Legends) reported more stress than those who play single-player games.

4. Tools and setup

Python in Jupyter Notebook (through Google Colab) with all needed libraries:

- Pandas and NumPy (For handling numbers and tables)
- Matplotlib and Seaborn (For creating chart)
- Scikit-learn and XGBoost (For prediction model)
- GitHub repository (Keep everything safely)
- Streamlit (For interactive dashboard)
- SQLite (tiny database)

5. Observations and Improvements

Initial work revealed opportunities to strengthen the project:

- Data Collection Boost: Sharing the survey in platforms like Reddit, Discord gaming groups etc doubled response rates.
- Improved Comparisons: Analysis became more understandable when all datasets were aligned with WHO's daily screen-time thresholds.

6. Next Steps

I'll now focus on getting more survey responses by sharing the form in gaming Discord groups and Reddit and merge with Kaggle dataset. Then I'll clean the data by fixing any odd answers and standardize measurement also check for duplicate responses or outliers. After that, I will create simple charts to spot early trends as Exploratory Data Analysis (EDA) like whether people who game more sleep less.I'll also get ready for the machine learning part by picking the most important factors (like gaming time and stress levels) and splitting the data into practice and test sets. Finally, I'll organize all my work on GitHub, streamlit cloud and start writing the methods section for the final report. By the next update, I should have clean data, clear charts showing the first trends, and a plan for building the prediction model.