

Brandon Mather

Project Proposal

## **Predicting Global Sales of Video Games**

### **Topic:**

This project aims to develop a predictive model to estimate the global sales of video games based on factors such as platform, genre, publisher, and release year.

### **Business Problem:**

The video game industry is highly competitive, and accurately predicting sales performance of new video games can help business decisions for developers and publishers. This project seeks to address this challenge by building a model that can forecast sales.

### **Datasets:**

The data for this project will be sourced from a dataset obtained by scraping vgchartz.com. This dataset contains information about video games with sales greater than 100,000 copies, including details such as platform, genre, publisher, release year, and sales in different regions (NA, Europe, Japan, other regions)

### **Methods:**

The methods that I will use will include:

- Data Cleaning – Handling missing values and feature engineering.
- Exploratory Data Analysis – Understand and identify trends and correlations.
- Feature Selection – Determine which features to include in the model.
- Modeling – Create a recommendation model to suggest movies based on user preferences.
- Model Evaluation – Evaluate model performance using metrics such as F1 score, precision, and recall.
- Analysis – Analyze model results.

### **Ethical Considerations:**

Potential ethical concerns for this project would include ensuring fairness and avoiding biases, respecting privacy and data protection, and being transparent about how the model's prediction is used.

### **Challenges/Issues:**

The types of challenges that could be faced include:

- Data Quality – Dealing with incomplete data, making sure it's reliable for making accurate recommendations.

- Bias – Addressing potential biases in the data or predictions.

**References:**

To validate my results, I can use academic research papers, publications, and online resources on topics such as video game market trends and performance metrics.