UE20CS312

Data Analytics – Mini-Project

Project Literature Review

Team no.: 18

Members:

SRN: PES2UG20CS279 **Name:** Robin Roy
SRN: PES2UG20CS284 **Name:** Ronit Bhathija
SRN: PES2UG20CS291 **Name:** Sachin Vishwamitra

BEST GENRES AND PLATFORMS FOR VIDEO GAMES USING EDA AND MACHINE LEARNING

Abstract:

Playing video games for many years has led to a large volume of gaming data. Most of the video gaming business organizations highly depends on knowledge of prediction of sales trend.

Machine learning techniques are very effective tools in obtaining hidden knowledge from an enormous dataset to enhance predictions.

In this project, we will find out which genres and platforms of gaming are growing the most and becoming more popular based on sales.

Introduction:

The main purpose of this model is to help game developers and business organizations to effectively choose the type of games that they want to make.

In this paper, we are concerned with which genres and platforms of gaming are growing the most and becoming more popular based on sales. In order to solve this problem, we will be using historical time series sales data.

Our dataset consists of 11 variables and about 16000 samples with a combination of categorical and numeric variables.

We will first create visualisations of the dataset based on different attributes such as platform and genre. We will also plot their sales over time. Using this we will get a rough estimate of which genres and platforms have grown over the years.

After doing visualisations, we will create a prediction model for the growth of different genres of gaming. We will display all of the forecasts simultaneously for different genres and platforms. This will give a good view of which genres and platforms will grow in the future.

How others have solved this problem?

Bodduru Keerthana & Dr. K. Venkata Rao are the authors of “Sales Prediction on Video Games using Machine Learning” [1]. In this paper, they predicted the overall sales of video games over the years. They made use of the random forest machine learning model to do the predictions.

We will be using a similar approach, but using EDA and predicting the sales for different genres and platforms.

Approach:

Our main approach to solve this problem is to make visualisations of the dataset based on different attributes such as platform and genre. We will also plot their sales over time. Using this we will get a rough estimate of which genres and platforms have grown over the years.

After doing visualisations, we will create a prediction model for the growth of different genres of gaming.

We will test the data using many regression models such as Linear regression, decision trees, random forests, extra trees, and others. We will find out which would give the best predictions and use that model.

Challenges we faced:

The years had to be converted to integer form as they were in decimal form.

Few columns had null values, and we had to drop the respective rows since only a small percentage of column values were null.

References:

[1] Bodduru Keerthana & Dr. K. Venkata Rao are the authors of “Sales Prediction on Video Games using Machine Learning”

Links:

Our Project GitHub link: https://github.com/Obsarian/Video-Game-Sales-EDA.git