

Classification of Video Game Genre

Darek Charlet

NWMSU Data Science Student

Question

Can we use a decision tree to accurately predict the genre of a video game based on platform, sales, release date, and other factors?

The decision tree was **not significantly accurate** in predicting the genre of the games in my test data. However, it showed that the most important factor in predicting genre was the **platform**.

The decision tree was too large to reasonably visualize. Instead, the most common genres for two of the platforms:

After trying a few different tunings, the best tree model was correct **56.7%** of the time.

This is not high enough to consider using the results for any sort of real application.

However, it did reveal the importance of platform in the classification. This insight could be used in combination with genre metrics to determine what platform a future game should be developed on.

Ex. Role playing games are more common on PS4 consoles than X Box, so an RPG developer should consider PS4 as their platform of choice.

Method

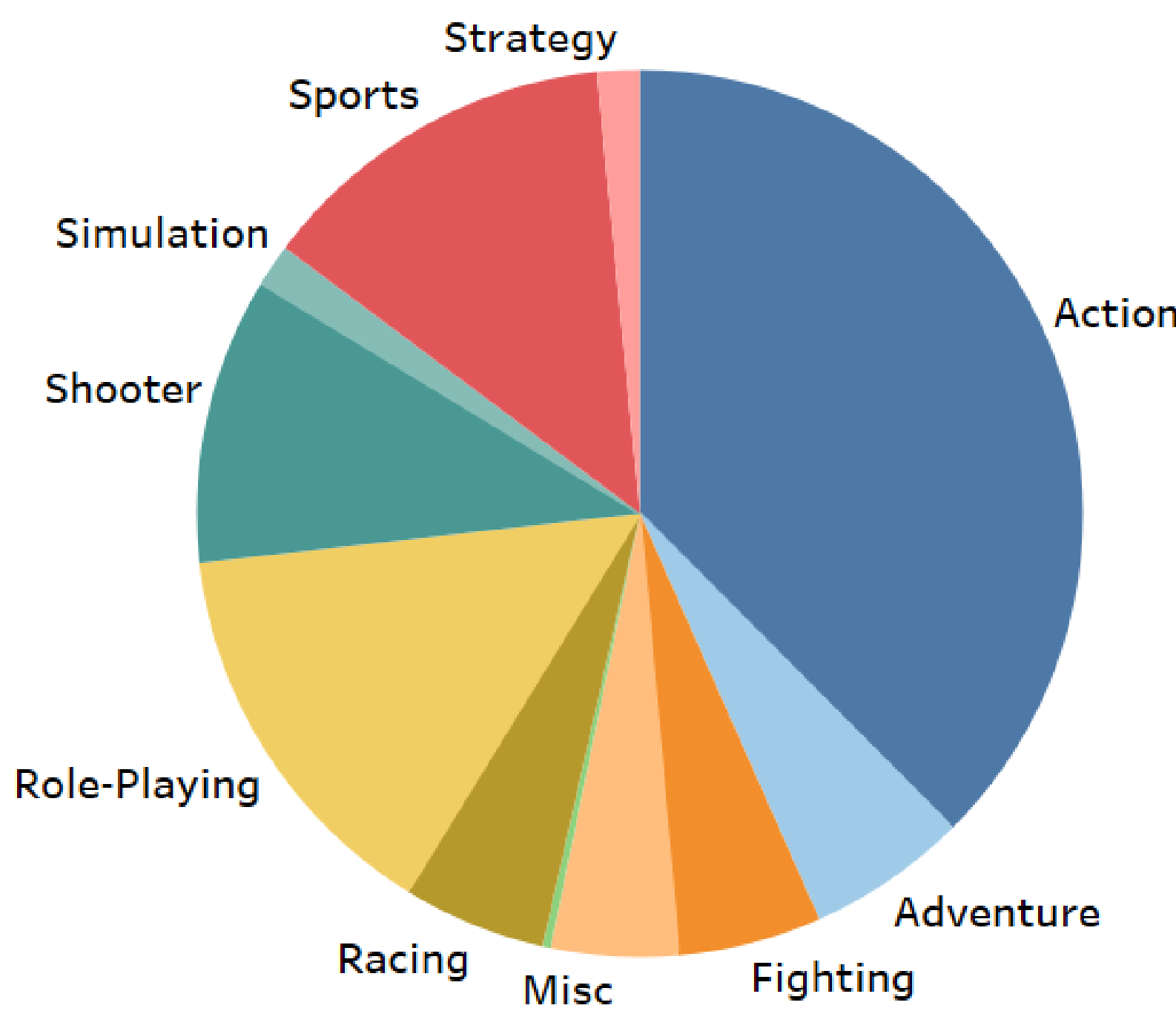
Excel and Tableau were used for data cleaning and preparation. Some fields had to be removed, some had to be filtered, and some were converted to another type to fit the model I was trying to build.

A J48 decision tree was generated in Weka with the remaining values.

Xbox One



PlayStation 4



Resources

Data Source - <https://www.kaggle.com/gregorut/videogamesales>

Project Repo - <https://github.com/17dchar/VideoGame-Genres-Research>

Dr. Eloë – Data Science Professor