

Preparação e Pré-processamento de Dados para Análise Preditiva entre Jogos Indie e Multi-A

Gabriel de França Marques (RA: 10395270)¹, Henrique Magno dos Santos(RA: 10335286)¹,
Pedro Machado Gomes Caixeta (RA: 10314309)¹

¹Ciência da Computação (CC) – Faculdade de Ciência e Informação (FCI) –
Universidade Presbiteriana Mackenzie (UPM)

10395270@, 10335286@, 10314309@mackenzista.com.br

Abstract. *This project aims to perform a predictive analysis using as a main basis a dataset of 1500 games from the Steam sales platform. The main objective is to predict and find a pattern related to revenue, reviews, other data, and the price of games in the store.*

Resumo. *Este projeto apresenta como objetivo realizar uma análise preditiva utilizando como base principal um dataset de 1500 jogos da plataforma de vendas Steam. O objetivo majoritário é prever e buscar um padrão relacionado à receita, avaliações, outros dados e o preço dos jogos na loja.*

1. Introdução

1.1. Contextualização

O mercado de jogos digitais tem se mostrado um campo fértil para a aplicação de técnicas de análise de dados e inteligência artificial. A capacidade de prever métricas-chave, como receita, avaliações e preço de jogos, pode trazer insights valiosos para desenvolvedores e publishers, auxiliando-os na tomada de decisões estratégicas.

Neste projeto, o objetivo é realizar uma análise preditiva utilizando um dataset de 1500 jogos da plataforma Steam. A escolha desse dataset se justifica pela relevância e disponibilidade de informações detalhadas sobre o mercado de jogos digitais. O estudo e previsão dessas métricas relevantes para o sucesso de um jogo podem contribuir significativamente para o setor.

1.2. Justificativa

O estudo e a previsão de métricas relevantes para o sucesso de um jogo, como receita, avaliações e preço, podem trazer insights valiosos para o setor.

1.3. Objetivo

O objetivo deste projeto é realizar uma análise preditiva utilizando um dataset de jogos independentes (indies) da plataforma Steam, buscando encontrar padrões e prever informações importantes para o sucesso de um jogo indie.

1.4. Opção do projeto

A escolha deste dataset de jogos indie da Steam se justifica pela relevância e disponibilidade de informações detalhadas sobre o mercado de jogos digitais independentes.

2. Descrição do Problema

O principal problema a ser abordado neste projeto é a identificação de fatores-chave que influenciam a receita dos jogos indies na plataforma Steam, com foco na análise da relação entre a pontuação de avaliação (review) e a receita desses jogos. Pretende-se desenvolver modelos preditivos capazes de estimar a receita dos jogos indies com base nas suas características, especialmente a pontuação de revisão.

3. Dataset

O dataset utilizado neste projeto contém informações de jogos independentes (indies) da plataforma Steam. Serão realizadas etapas de limpeza, tratamento de dados ausentes, transformação de variáveis e análise exploratória para preparar o dataset para a modelagem preditiva, com foco na análise da relação entre a pontuação de avaliação (review) e a receita dos jogos indies.

4. Metodologia e Resultados Esperados

Serão aplicadas técnicas de aprendizado de máquina supervisionado, como regressão linear e árvores de decisão, para prever as métricas de receita, avaliações e preço dos jogos. Espera-se obter modelos preditivos com bom desempenho, capazes de fornecer insights valiosos para o mercado de jogos digitais.

5. Bibliografia

O dataset foi obtido em [Topcu 2024]. Os slides de "Preparação e Pré-processamento dos dados" utilizados como base de estudo foram ministrados pelo Prof. Dr. Ivan Carlos Alcântara de Oliveira, sendo disponibilizados nas seguintes partes: [de Oliveira 2024a], [de Oliveira 2024b] e [de Oliveira 2024c].

Referências

- de Oliveira, P. D. I. C. A. (2024a). Preparação e pré-processamento dos dados parte i. Slide de aula. Acesso em: 2024-09-21.
- de Oliveira, P. D. I. C. A. (2024b). Preparação e pré-processamento dos dados parte ii. Slide de aula. Acesso em: 2024-09-21.
- de Oliveira, P. D. I. C. A. (2024c). Preparação e pré-processamento dos dados parte iii. Slide de aula. Acesso em: 2024-09-21.
- Topcu, A. C. (2024). Top 1500 games on steam by revenue 09-09-2024. <https://www.kaggle.com/datasets/alicemtopcu/top-1500-games-on-steam-by-revenue-09-09-2024>. Acesso em: 2024-09-21.