**Correlação entre as variáveis de um filme e a Nota de Críticos e do Público**

Cleiton Pereira da Silva – São Paulo – SP cleitonps.mkt@hotmail.com;

Leila Rabello de Oliveira

**Correlação entre as variáveis de um filme e a Nota de Críticos e do Público**

**Resumo**

O estudo avaliará a correlação entre as diversas variáveis de um filme e as notas atribuídas pelos críticos especializados e pelo público em geral. A avaliação técnica de um filme é correlacionada com diversos fatores, como elenco, direção, roteiro, gênero, orçamento e duração, entre outros. Com o avanço das tecnologias de coleta e análise de dados, é possível explorar amplamente as relações entre essas variáveis e as notas atribuídas pelos críticos e o publico.

Aqui utilizaremos técnicas de data analytics e machine learning, que nos permitirão examinar um conjunto de dados abrangente de uma ampla lista de filmes, contendo informações como elenco, diretor, gênero, orçamento e duração, juntamente com as notas atribuídas pelos críticos e pelo público. Faremos uso de métodos estatísticos e algoritmos de regressão para identificar quais variáveis têm maior influência na determinação da nota média.

Este estudo tentará contribuir para uma melhor compreensão das relações entre as variáveis de um filme e as avaliações de críticos e público, fornecendo possíveis insights para a indústria cinematográfica. Além disso, os resultados obtidos podem auxiliar na tomada de decisões, como a seleção de elenco, direção e roteiro, visando alcançar uma recepção mais positiva tanto pelos críticos especializados quanto pelo público em geral.

Palavras-chave: Correlação, Nota de Críticos, Nota do Público, Variáveis de um Filme, Data Analytics, Machine Learning.

**Introdução**

A indústria cinematográfica tem despertado um interesse crescente no estudo das relações entre as características de um filme e as avaliações feitas pelos críticos especializados e pelo público em geral. Diversos fatores, como elenco, direção, roteiro, gênero, orçamento e duração, são considerados na avaliação técnica de um filme. Com o avanço das tecnologias de coleta e análise de dados, é possível explorar de forma abrangente as relações entre essas variáveis e as notas atribuídas pelos críticos e pelo público.

Nesta monografia, utilizaremos técnicas de data analytics e machine learning para examinar um conjunto de dados abrangente de uma ampla lista de filmes. Esse conjunto de dados contém informações como elenco, diretor, gênero, orçamento e duração, além das notas atribuídas pelos críticos e pelo público. Por meio de métodos estatísticos e algoritmos de regressão, buscaremos identificar quais variáveis possuem maior influência na determinação da nota média.

O objetivo deste estudo é contribuir para uma melhor compreensão das relações entre as características de um filme e as avaliações realizadas por críticos e público. Ao fornecer possíveis insights para a indústria cinematográfica, os resultados obtidos poderão auxiliar na tomada de decisões, como a seleção de elenco, direção e roteiro, visando alcançar uma recepção mais positiva tanto pelos críticos especializados quanto pelo público em geral.

Por meio da análise aprofundada dessas relações, esperamos contribuir para o avanço do conhecimento no campo do cinema, possibilitando uma melhor compreensão dos fatores que influenciam a recepção e o sucesso de um filme.

**Material e Métodos**

Os materiais e métodos utilizados podem ser resumidos nos tópicos abaixo.

Um conjunto de dados de filmes foi obtido no sistema Data.World contendo informações abrangente sobre os filmes, como elenco, direção, roteiro, gênero, orçamento e duração, além das notas atribuídas pelos críticos especializados (Indicada pelo site Metacritics) e pelo público em geral (Indicada pelo site IMDB).

Técnicas de Data Analytics foram aplicadas para explorar os dados disponíveis, identificar padrões, realizar análises estatísticas e extrair insights relevantes. A atividade envolveu a utilização de ferramentas e linguagens de programação como a linguagem R, para manipulação e análise dos dados.

Tecnicas de Machine Learning foram aplicadas, como regressão linear, regressão logística ou algoritmos de árvore de decisão, para modelar a relação entre as variáveis do filme e as notas atribuídas pelos críticos e pelo público. Será necessário treinar e ajustar os modelos com base nos dados disponíveis.

Análises Estatísticas foram realizadas para identificar a correlação entre as variáveis do filme e as notas atribuídas. A atividade envolver cálculo de coeficientes de correlação, testes de significância estatística e análise de regressão para determinar a influência das variáveis na nota média.

A avaliação dos Resultados ocorrerá para termos a relevância de cada variável no objetivo do estudo e significância estatística destas variáveis. A atividade irá interpretar os resultados e identificar as variáveis que apresentam maior correlação com as notas atribuídas pelos críticos e pelo público.

**Resultados Preliminares**

Em análise.

**Conclusão(ões) ou Considerações Finais**

Pendente.

**Agradecimentos**

Gostaria de agradecer a Cleverson Silva, meu irmão, graduando em Ciências da computação, que me apresentou ferramentas que possibilitaram meu aprendizado técnico para utilização de técnicas de machine learning. Sem o apoio dele, não teria me mantido engajado para completar este passo de minha pós-graduação.

**Referências**

Introduction to Econometrics with R - Christoph Hanck, Martin Arnold, Alexander Gerber, and Martin Schmelzer

Data Science for Business and Decision Making - Por Favero, Luiz Paulo - Belfiore, Patricia - 2019

Handbook of Regression Modeling in People Analytics - With Examples in R and Python - Por McNulty, Keith· 2021

**Apêndice ou Anexo** (opcional)

Pendente.