

Aprendizado de máquina aplicado a previsão de retornos de ativos virtuais no Mercado da Comunidade Steam

Pedro Mendes André Ricardo Oliveira da Fonseca*

Introdução

Este projeto se propõe a usar técnicas de aprendizado de máquina para prever o retorno de ativos virtuais no Mercado da Comunidade Steam (MCS).¹ O MCS é um mercado virtual secundário onde usuários podem comprar e vender itens virtuais para jogos da plataforma Steam, uma loja de jogos online. O mercado é composto por itens de jogos como Counter-Strike: Global Offensive, Dota 2, Team Fortress 2, entre outros. Os itens são negociados na moeda corrente da conta utilizada e a plataforma retém 15% do valor da transação. O restante é creditado na carteira do usuário, que pode ser usado para adquirir outros itens ou jogos na plataforma.² O preço dos ativos é determinado pela oferta e demanda dos usuários, podendo variar ao longo do tempo. A análise de mercados virtuais pode ser uma importante ferramenta para testar hipóteses econômicas, pois modelar um sistema fechado com poucas variáveis a serem analisadas permite abstrair a complexidade de mercados reais.

Objetivos

Como objetivo geral, este trabalho planeja utilizar técnicas de aprendizado de máquina tanto para modelar o comportamento de ativos virtuais no MCS como também, para fins acadêmicos, criar um portfólio teórico lucrativo com tais informações.

Como objetivos específicos, este trabalho se propõe a:

*Professor orientador - Centro de Matemática Computação e Cognição (CMCC/UFABC).

¹Disponível em: <https://steamcommunity.com/market>. Acesso em: 19 nov 2023.

²Disponível em: <https://help.steampowered.com/en/faqs/view/61F0-72B7-9A18-C70B>. Acesso em: 19 nov 2023.

- Realizar uma análise qualitativa das regras da plataforma, a fim de entender o funcionamento do mercado e as possíveis limitações para a coleta de dados;
- Analisar os principais ativos negociáveis por meio de estatística descritiva;
- Desenvolver um modelo estatístico que se baseie em técnicas de aprendizado de máquina;
- Medir o desempenho do modelo utilizado.

Problema

O problema central deste projeto reside na capacidade de prever os retornos dos ativos virtuais no Mercado da Comunidade Steam (MCS) por meio de técnicas de aprendizado de máquina. O contexto inclui a investigação de padrões sazonais nos preços dos ativos, como influenciados por eventos sazonais e lançamentos de jogos, bem como a possibilidade de construir um portfólio de ativos com retornos superiores à retenção de 15% do valor da transação cobrado pela plataforma Steam.

Hipóteses

- Hipótese 1: O uso do modelo de aprendizado de máquina Long Short Term Memory (LSTM) é adequado para prever os rendimentos dos ativos virtuais no contexto do MCS.³
- Hipótese 2: Existem padrões sazonais no comportamento dos preços dos ativos virtuais no MCS, influenciados por eventos como promoções sazonais e lançamentos de jogos populares.
- Hipótese 3: É possível criar um portfólio de ativos tal que seu retorno seja superior à retenção de 15% do valor da transação cobrado pela plataforma Steam.

Justificativa

O MCS é uma plataforma única que permite aos jogadores comprar, vender e trocar itens virtuais recebidos em jogos. Houve um expressivo aumento nos preços e nas quantidades das transações desses ativos virtuais, refletindo a crescente popularidade dos videogames e o interesse em possuir e negociar tais itens. O valor estimado do mercado de mais de US\$ 15 bilhões ressalta sua importância econômica, tornando-o um objeto de estudo oportuno para a pesquisa acadêmica (Glaus, 2022).

Apesar de sua relevância, existem poucas pesquisas acadêmicas sobre a Steam e

³A escolha do modelo é detalhada na seção de metodologia.

seu mercado de ativos virtuais. Isso pode ser resultado de diversos elementos, como o conhecimento limitado sobre a indústria de jogos, que vai além da simples compra e venda de jogos; as dificuldades relativamente elevadas para ingressar na plataforma; e a natureza fechada do ambiente, onde os recursos não podem ser retirados ou transferidos para contas externas ou bancos (Glaus, 2022).

Portanto, o crescimento do MCS, sua importância econômica e a escassez de pesquisas acadêmicas sobre o assunto fazem dele um tema atraente para investigação. A possibilidade de testar modelos em um ambiente controlado com poucas variáveis a serem consideradas também se mostra interessante, ao contribuir para a redução da complexidade do ambiente de testes para pesquisas aplicadas. A realização de pesquisas nessa área pode fornecer informações valiosas sobre o comportamento do mercado, podendo ser benéfica tanto para pesquisadores como para a comunidade de jogos em geral.

Metodologia

Para a viabilidade do projeto, é necessário conhecer as regras de compra e venda de ativos virtuais no MCS e, para isso, este trabalho fará uma análise exploratória da página de perguntas mais frequentes do mercado.⁴ Apesar da plataforma Steam não fornecer em sua API oficial campos específicos para o MCS, os dados sobre os ativos virtuais a serem analisados podem ser acessados pelo site SteamApis⁵ e os dados sobre todas as principais promoções e eventos da loja pelo site SteamDB.⁶ Com o acesso a estes dados, será realizada uma análise descritiva dos principais ativos negociados na plataforma, com informações como volume de negociações, média e volatilidade dos retornos e demais estatísticas descritivas.

O modelo escolhido para o projeto é o LSTM, um tipo de arquitetura de rede neural recorrente (RNN) projetada especificamente para lidar com dependências de longo prazo e sequências de dados. As RNNs são uma classe de redes neurais que podem processar sequências de dados mantendo uma memória interna ou um estado que lhes permite reter informações de etapas de tempo anteriores, sendo natural a aplicação dele para séries temporais (James *et al.*, 2023). O modelo também considerará os resultados encontrados por Glaus (2022) no que diz respeito ao efeito sazonal das promoções anuais nos preços dos ativos da loja.

Para medir o desempenho do modelo, além do uso de medidas como erro quadrático

⁴Disponível em: <https://help.steampowered.com/en/faqs/view/61F0-72B7-9A18-C70B>. Acesso em: 2 ago 2023.

⁵Disponível em: <https://steamapis.com/>. Acesso em: 2 ago 2023.

⁶Disponível em: <https://steamdb.info/>. Acesso em: 3 ago 2023.

médio (RMSE)⁷, o projeto utilizará previsão conforme (CP)⁸, uma técnica de aprendizado de máquina e inferência estatística que fornece uma maneira de quantificar a incerteza associada às previsões utilizando informações anteriores. Diferentemente dos modelos de previsão tradicionais que geram estimativas pontuais, a previsão conforme produz regiões de previsão que contêm o valor verdadeiro com uma probabilidade especificada (Shafer e Vovk, 2008).

Revisão bibliográfica

Este trabalho se baseará fundamentalmente nos trabalhos de Glaus (2022) para descrever os efeitos de sazonalidade no MCS; Han *et al.* (2021), Yu *et al.* (2019) e James *et al.* (2023) para definir e estimar os modelos; Shafer e Vovk (2008) e Angelopoulos e Bates (2022) para aplicar a técnica de CP para medir os resultados do modelo.

O trabalho de Glaus (2022) investiga o fenômeno do efeito-promoção no MCS durante o período de dezembro de 2012 a outubro de 2021.⁹ Ao analisar portfólios compostos por itens de oito jogos populares frequentemente negociados no mercado, o estudo encontrou uma forte presença de sazonalidade nos retornos dos itens. Além disso, as medidas de liquidez são examinadas para entender os fatores que contribuem para o surgimento do efeito-promoção. O artigo também explora a eficácia das estratégias de investimento criadas para capitalizar no efeito-promoção, considerando os custos de negociação e seu impacto sobre a lucratividade. Os resultados revelam um efeito-promoção significativo em toda a amostra agregada e nos portfólios de jogos individuais. Esse efeito demonstra que os retornos durante os períodos de promoção exibem um padrão de retornos iniciais mais baixos, seguidos de retornos mais altos no final e após o período de promoção, em comparação com os dias sem promoção. Esse padrão se mantém para a maioria dos portfólios, e a força do efeito-promoção varia entre portfólios ponderados por igual e ponderados por preço, indicando que ele é influenciado pelo valor dos itens. O estudo estabelece que o efeito-promoção não é impulsionado apenas por promoções sazonais individuais específicas, mas sim pela combinação das promoções de verão, outono e inverno. Além disso, um exame das medidas de liquidez sugere que o efeito de promoção pode ser impulsionado pelo aumento do volume de negociação nos dias de promoção. Este artigo contribui para este trabalho por mostrar que tal efeito deve ser incorporado nas estimações de retornos do modelo escolhido. Porém, o estudo conclui que a implementação de estratégias de negociação com base no efeito-promoção é desafiadora devido aos custos substanciais de

⁷No sentido original, root mean square error (tradução nossa).

⁸No sentido original, conformal prediction (tradução nossa).

⁹Segundo Glaus (2022), o efeito-promoção consiste em uma anomalia de mercado onde os retornos dos itens dos jogos tendem a ser menores em dias no começo e maiores no fim de uma grande promoção na plataforma Steam em relação a dias onde não existem grandes promoções.

negociação, que impedem que essas estratégias superem o desempenho de uma estratégia simples de comprar e manter.

Os trabalhos de Shafer e Vovk (2008) e Angelopoulos e Bates (2022) definem a técnica de CP. Os chamados modelos de aprendizado de máquina de caixa-preta já estão difundidos nas suas mais diversas aplicações. Porém, medidas robustas para quantificar a incerteza dos modelos ainda não são tão comuns, apesar de poderem evitar falhas consideráveis na aplicação dos modelos. A técnica de predição conforme é um paradigma amigável para criar intervalos de incerteza estatisticamente robustos para a predição de tais tipos de modelos. Tal método independe de distribuições no sentido de que possuem garantias explícitas e não assintóticas, mesmo sem nenhuma suposição dos modelos utilizados. Assim, tal técnica se mostra bastante robusta para avaliar o desempenho dos modelos deste trabalho. Por fim, os trabalhos de Han *et al.* (2021), Yu *et al.* (2019) e James *et al.* (2023) trazem uma visão geral e alguns fundamentos sobre métodos de aprendizado de máquina, além de mostrarem a definição e a aplicação de modelos de RNN, a família de modelos a qual o modelo escolhido por esse trabalho, LSTM, faz parte.

Cronograma

2024												
	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez
Entrega do projeto de pesquisa												
Revisão bibliográfica												
Análise qualitativa do objeto de pesquisa												
Análise quantitativa do objeto de pesquisa												
Construção do modelo preditivo												
Verificação de desempenho do modelo												
Escrita da monografia												
Entrega												

Referências

- ANGELOPOULOS, A. N.; BATES, S. **A gentle introduction to conformal prediction and distribution-free uncertainty quantification**. arXiv, 7 dez. 2022. Disponível em: <<http://arxiv.org/abs/2107.07511>>. Acesso em: 29 ago. 2023
- GLAUS, T. **Seasonality of in-game item returns in the Steam Community Market**. Rochester, NY, 19 mar. 2022. Disponível em: <<https://papers.ssrn.com/abstract=4314409>>. Acesso em: 29 ago. 2023
- HAN, Z. *et al.* **A review of deep learning models for time series prediction**. **IEEE Sensors Journal**, v. 21, n. 6, p. 78337848, mar. 2021.
- JAMES, G. *et al.* **An introduction to statistical learning: with applications in Python**. Cham: Springer, 2023.
- KAKUSHADZE, Z.; SERUR, J. A. **151 trading strategies**. Rochester, NY, 17 ago. 2018. Disponível em: <<https://papers.ssrn.com/abstract=3247865>>. Acesso em: 21 out. 2023
- KELLY, B. T.; XIU, D. **Financial machine learning**. Rochester, NY, 24 jul. 2023. Disponível em: <<https://papers.ssrn.com/abstract=4520856>>. Acesso em: 29 ago. 2023
- SHAFFER, G.; VOVK, V. A tutorial on conformal prediction. **Journal of Machine Learning Research**, v. 9, n. 3, 2008.
- STANKEVICIUTE, K.; M. ALAA, A.; SCHAAR, M. VAN DER. **Conformal time-series forecasting**. Advances in Neural Information Processing Systems. **Anais...**Curran Associates, Inc., 2021Disponível em: <<https://proceedings.neurips.cc/paper/2021/hash/312f1ba2a72318edaaa995a67835fad5-Abstract.html>>. Acesso em: 29 ago. 2023
- XU, C.; XIE, Y. **Conformal prediction set for time-series**. arXiv, 15 jun. 2022. Disponível em: <<http://arxiv.org/abs/2206.07851>>. Acesso em: 29 ago. 2023
- YU, Y. *et al.* **A review of recurrent neural networks: LSTM cells and network architectures**. **Neural Computation**, v. 31, n. 7, p. 12351270, 1 jul. 2019.