

Precificação de NFT's com Random Forest Regressor e Bagging Regressor: um estudo de caso

Isaac Miranda Camargos¹

UFTM, Uberaba, MG(Campus ICTE)

Leandro Cruvinel Lemes²

UFTM, Uberaba, MG(Campus ICTE)

A rede Blockchain consiste em um banco de dados representados por blocos de informações criptografadas, ligadas em cadeia, em que cada bloco compreende múltiplas transações [5]. Tal tecnologia foi proposta por Nakamoto [4] para possibilitar o livre comércio digital, diretamente entre pessoas, livrando-se da necessidade de uma terceira entidade mediadora e, conseqüentemente, escapando de taxas pelos seus serviços. Com base nisso, juntamente com a criação desse banco de dados descentralizado, surgiu a primeira criptomoeda, o Bitcoin, que após sua criação levou alguns anos até sua popularização, sendo marcado por uma expressiva valorização no início de 2021. Neste mesmo ano, no dia 10 de novembro, o Bitcoin atingiu o histórico valor de 69.044,77 dólares por uma única moeda [1].

Embora não ligados a uma instituição financeira, as criptomoedas seguem o padrão de tokens ERC-20, o que às conferem as mesmas características de uma moeda convencional, por exemplo, moedas de um mesmo tipo possuem o mesmo valor [2]. Por outro lado, os NFT's (non-fungible tokens) são um tipo de criptomoeda derivados de contratos inteligentes da rede Blockchain. Os NFT's são cunhados no padrão de tokens ERC-721 [7], onde cada token é único, podendo ser distinguido e rastreado por seu identificador. Em prática, os NFT's podem ser qualquer tipo de bem digital, sendo os mais comuns trabalhos de arte colecionáveis, objetos virtuais, músicas ou itens dentro de jogos.

Os NFT's, diferentemente das criptomoedas, são únicos tanto em identificação como em características, o que torna sua precificação uma tarefa árdua e subjetiva. Neste presente trabalho, motivados pela crescente valorização desses tokens e popularização dessa tecnologia, temos como objetivo avaliar o modelos Random Forest Regressor e Bagging Regressor, algoritmos de aprendizado de máquina (AM), na estimativa do preço de um NFT considerando seus atributos. Como objeto de estudo, utilizaremos o jogo Gods Unchained, cujo objetivo é derrotar o seu oponente em uma batalha de cartas. Neste caso, os NFT's são as cartas, essas possuem características únicas bem definidas, como mana, vida, ataque e descrição. Tais individualidades visam favorecer a avaliação de preços, considerando a ausência de literatura científica nessa área.

O trabalho foi dividido em três etapas: coleta, pré-processamento e modelagem dos dados. As bibliotecas Selenium, Pandas, Numpy, Scikit-learn e Natural Language Toolkit (NLTK), escritas na linguagem de programação Python, foram utilizadas em cada etapa.

No que diz respeito à coleta, os dados foram carregados de interfaces de programação de aplicações (API's) e de páginas web fornecidas pelo próprio jogo em estudo [3]. Nome, raridade, coleção, tipo, mana, ataque, defesa, descrição, preço, porcentagem de vitória e porcentagem de baralhos em que cada carta está presente, foram as características coletadas. Ademais, o preço da carta foi obtido em Ethereum e GODS, essa última é uma criptomoeda desenvolvida para transações

¹isaacmirandacamargos@gmail.com

²leandro.lemes@uftm.edu.br

dentro do jogo. Em síntese, o banco de dados a ser trabalhado foi construído combinando todas essas informações.

Já no pré-processamento dos dados, as cartas não NFT's foram excluídas, valores extremos foram removidos utilizando-se boxplot, as variáveis categóricas foram discretizadas e as variáveis numéricas foram normalizadas subtraindo-se a média e dividindo pelo desvio padrão. Além disso, ainda na etapa de pré-processamento, para a descrição da carta (dado textual), técnicas de remoção de caracteres especiais, links, tags e stopwords (palavras de alta frequência e pouca significância), extração de radical e vetorização foram aplicadas.

Por fim, na etapa de modelagem, se fez uso do Random Forest Regressor e Bagging Regressor [6], modelos de aprendizagem de máquina. As métricas utilizadas foram erro médio absoluto (EMA), erro médio quadrático (EMQ) e R^2 . Os resultados obtidos na precificação das cartas em Ethereum e GODS, podem ser vistos na tabela 1 e tabela 2, respectivamente.

Tabela 1: Resultados da precificação em Ethereum.

Modelo aplicado	EMA médio	EMQ médio	R^2
Random Forest Regressor	0.1435 ± 0.0115	0.0544 ± 0.0101	0.7994 ± 0.0322
Bagging Regressor	0.1500 ± 0.0124	0.0598 ± 0.0109	0.7793 ± 0.0350

Tabela 2: Resultados da precificação em GODS.

Modelo aplicado	EMA médio	EMQ médio	R^2
Random Forest Regressor	0.1441 ± 0.0118	0.0553 ± 0.0105	0.7996 ± 0.0322
Bagging Regressor	0.1509 ± 0.0128	0.0610 ± 0.0115	0.7791 ± 0.0350

Os resultados da precificação em Ethereum e GODS foram semelhantes. Com relação aos modelos, o Random Forest Regressor obteve erros inferiores nas duas precificações em comparação com o Bagging Regressor.

Referências

- [1] CoinGecko. **Site oficial do CoinGecko**. Online. Acessado em 14/03/2022, <https://www.coingecko.com/en/coins/bitcoin>.
- [2] V. Fabian e B. Vitalik. **Erc-20 token standard**. Online. Acessado em 15/03/2022, <https://eips.ethereum.org/EIPS/eip-20>.
- [3] Immutable. **GUDecks**. Online. Acessado em 27/03/2022, <https://gudecks.com/meta/card-rankings>.
- [4] Satoshi Nakamoto. "Bitcoin: A peer-to-peer electronic cash system". Em: **Decentralized Business Review** (2008), p. 21260.
- [5] Nofer, M., Gomber, P., Hinz, O., and Schiereck, D. "Blockchain". Em: **Business and Information Systems Engineering** 59(3) (2017), 183–187. DOI: {10.1007/s12599-017-0467-3}.
- [6] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". Em: **Journal of Machine Learning Research** 12 (2011), pp. 2825–2830.
- [7] E. William et al. **Erc-721 non-fungible token standard**. Online. Acessado em 15/03/2022, <https://eips.ethereum.org/EIPS/eip-721>.