Informatics Institute of Technology
In Collaboration With

University of Westminster, UK

University of Westminster, Coat of Arms

Trading Recommendations System for Non-fungible Tokens

A dissertation by
Mr. Dinuka Ravijaya Piyadigama
w1742104 / 2018373

Supervised by
Mr. Guhanathan Poravi

May 2022

Submitted in partial fulfilment of the requirements for
the BSc (Hons) Computer Science degree at
the University of Westminster.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

LSTM Long short-term memory. 5, 6

MAE Mean Absolute Error. 4
ML Machine Learning. 2
MLP Multilayer Perceptron. 4
MSE Mean Squared Error. 4

NFT Non-fungible Token. 2, 5

CHAPTER 1: INTRODUCTION

CHAPTER 2: LITERATURE REVIEW

2.1 Chapter Overview

2.2 Problem Domain

The ERC-721 standard implements functionalities to transfer tokens from Blockchain accounts, to get the current token balance of an account, to get the owner of a specific token, the total supply of tokens available on the network, etc. Apart from the item itself, the creator can include metadata such as their signature in the NFT. What began on the Ethereum Blockchain with the ERC-721 standard has since been adopted by other Blockchains.

NFTs have a feature to allow a creator to make a certain percentage as royalty whenever the NFT is transferred to a new buyer. Since the items can be verified on the Blockchain, it also ensures that the original creator of the NFT can be tracked down and given due credit, any date in the future, no matter how many wallets it gets passed through (Chevet, 2018). Apart from the fact that a buyer can claim the right of ownership of the original item, they also get to financially support the creator. Ultimately, NFTs may gain value over time due to their scarcity. This gives collectors an additional advantage of being able to sell it for a higher price later on.

Creators of NFTs can also create "shares" for their NFT. This allows investors and fans to own a portion of an NFT without having to purchase the entire thing (*ERC-721 Non-Fungible Token Standard* 2021).

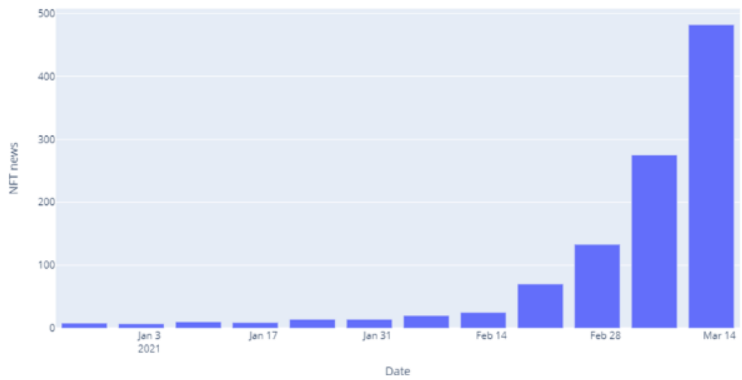

Figure 2.1: News trend in 2021 related to NFTs (Dowling, 2021a)

The above figure shows the increase in news trends related to NFTs since the start of 2021. It has been exponentially increasing and hitting headlines around the globe on a daily basis.

2.3 Concept Map

2.4 Existing work

There is only one study previously done with related to recommending Non-fungible Token (NFT)s and that study also comes in the form of a blog article on *OpenSea* (*What are you missing? Using basic machine learning to predict and recommend NFTs with OpenSea data - OpenSea Blog* 2020). The article considers the use of a basic Machine Learning (ML) technique called Multiple Regression with data gathered from OpenSea.

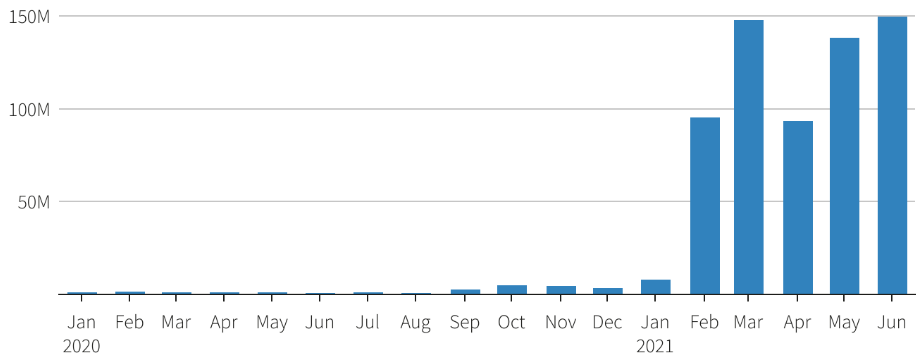

Figure 2.2: Monthly Ethereum-based NFT token sales volume on the OpenSea marketplace, in USD (Howcroft, 2021)

> *"Crypto has a founding tradition of emphasizing freedom and privacy. Maybe because of this prevailing cultural trend, the NFT space does not have many recommender systems."*
> (*What are you missing? Using basic machine learning to predict and recommend NFTs with OpenSea data - OpenSea Blog* 2020)

A hybrid Recommendations System (Cheng and Lin, 2020) which utilizes opinion & sentiment extraction techniques from user reviews to create preference profiles for movie recommendations, to enhance the quality of recommendations regardless of the rich or sparse nature of the dataset has been identified as one of the recent researches done towards pushing the limits of baseline recommendation models. The framework that has been designed here uses Collaborative Filtering as the base Recommendations model. The contribution of this research is applicable to the feature engineering stage of the system.

Sentiment analysis is applied on user-reviews to detect user-opinions about movies that were watched and reviewed by users. This data is used to create a user's preference profile, similar to what's created in Content-based filtering. The user's sentiment is identified as a step beyond traditional preference ratings.

Due to its capability of dealing with insufficient data, the framework is able to produce recommendations that are more accurate and efficient than existing baseline methods. This proves that using public opinion in the feature engineering stage can enhance the quality of recommendations.

Due to the fact that the semantic strategy of opinion extraction being generic, it is understood that it may not be ideal to identify different aspects in varied genres. Examples mentioned are, quality of sound may be of greater interest in action movies, while the story-line in dramas. Slang, irony & sarcasm haven't been taken into consideration when extracting user opinion. A major limitation identified in most systems that rely on similar opinion mining systems is that they are very dependant on the text mining technique used. Another identified drawback in this research by the author is that, to establish a preference profile, a person must have posted reviews on previous movies. If not, those users won't be able to get recommendations. This can be identified as a concern in systems that are dealing with user's who care about their privacy.

A Deep Belief Network and Sentiment Analysis (DBNSA) has been introduced to achieve data learning for recommendations (Chen and Hendry, 2019) to enhance recommendations produced by baseline-recommendation techniques. This deep learning model processes user comments to generate a possible user rating for user recommendations.

> *"Users usually transmit their decisions together with emotions."*
>
> (ibid.)

This research paper emphasizes the necessity of using user comments for recommendation systems since these comments contain a variety of emotional information that can influence the correctness and precision of recommendations.

Once applying sentiment analysis, a feature vector is created for the input nodes. A noise reduction procedure has been integrated into the system that deletes short comments, comments with no expression and false rating comments. This is used to improve the classification of user ratings. Finally, the DBNSA accomplishes data learning for the recommendations.

The paper published claims to outperform baseline models in training loss, precision and recall when tested on Yelp & Amazon datasets. When tested on the Trip-Advisor dataset, DBNSA had the best Mean Squared Error (MSE) training loss value & recall. The research also mentions that DBNSA saves more time, while producing results with better accuracy compared to other baseline models.

The main drawback that this paper points out is that the proposed system is not suitable & ready for real-time testing. The authors of the paper have also shown interest in testing the proposed method with a faster Deep Learning algorithm. Similar to the previously mentioned system, sarcastic user-comments have not been taken into consideration here as well. Out of the two recommendations models that were tested, *libSVM* was identified to have higher accuracy value, Mean Absolute Error (MAE) and F-score, while the Multilayer Perceptron (MLP) had the highest precision value.

Since user relationships and timeline comments also affect the user's decision making, these can be used to find information from relatable timelines to solve the cold start problem.

A hybrid approach that combines techniques from content-based filtering, user-to-user collaborative filtering and personalize recommendations (Ayushi and Prasad, 2018) has been introduced to address the limitation of dingle domain analysis. Data sparsity and cold start problem have been pointed out as the addressed limitations. Movie domain knowledge has been used to generate recommendations for books & music. After considering an array of supervised learning algorithms, the authors came to a conclusion that the Decision Tree classifier was found to give the highest accuracy.

The use of data from multiple domains allows the system to generate higher accuracy in suggestions. Twitter sentiment has been used to present the user with an analysis of the recommendations produced, to help users in their decision making process.

The drawback identified in the Recommendations System developed here is that Twitter sentiment is analysed, calculated and displayed only after showing the user recommendations.The author's suggestion is that only the items with positive sentiment could've been presented, at least results could've been bias towards positive sentiment.

Since NFTs have a distant relationship with crypto assets, it is expected to be of help to understand how crypto assets are evaluated when opted for selection to comprehend how NFT assets could be evaluated. A study done related to a modelling framework that exposes this area of research (Bartolucci and Kirilenko, 2020) assumes that two main features, namely security and stability can be used to determine the user-desire to own a specific crypto asset.

Investor's attitudes towards assets' features, information about the adoption trends, and expected future economic benefits of adoption have been simulated in order to predict the features of the assets that will most likely be adopted. The preference of investors are collected from an app, which calculates the overall state of the 'market'. Then, the app recommends to the user which crypto assets proposed by the user would be a sensible investment. Information about the adoption choice of other investors is considered when making this recommendation.

The number of assets, investors and assets' features and investor preferences were fixed within the period of analysis. In a normal use-case scenario, it's highly likely that all these would fluctuate and evolve with the asset's adoption probabilities and expected returns. This revelation clarifies the fact that crypto related assets have a tendency to change with time, social acceptance and trends. Therefore, it is important to consider these factors when building a crypto-related Recommendations System.

When considering ownership desire of NFTs, it is understood from domain research that the increase in price of an NFT has the possibility of being a factor to be considered when making a purchase. As mentioned under the Problem Domain section of this literature review, it is understood that NFTs have very little spill-over with other Crypto assets. However, knowing Crypto price prediction models is important since Wavelet coherence analysis indicates a co-movement between these two markets (Dowling, 2021b). These models can be used separately on each NFT asset to anticipate the pricing with related to time, sales & bids. The author finds this research to be related to address the research gap in this thesis since an appropriate price prediction could be used to enhance NFT recommendations to users.

Past research suggests a model which employs time series techniques, can predict the price for the next few days by splitting the data into train and test runs (Ferdiansyah et al., 2019).

In terms of RMSE (Root Mean Squared Error), the result is insufficient. The authors of this research have shown interest in testing out this method with modified Long short-term memory (LSTM) layers by adding dropout and modifying the number of epochs. Using different instability data-sets can also be tried out to test how good the prediction results could get. Furthermore, sentiment analysis is also proposed as future work to be combined with the LSTM method. This could be used to identify how public sentiment 'causes the value of crypto to adjust, with related to past price-fluctuations.

2.4.1 Benchmarking

2.5 Review of Different Problem-solving Approaches

There are several baseline techniques of Recommendations Systems that have been used by the biggest data-driven companies around the world.

(Larry, 2019) There have been many attempts to expand the capabilities of Recommendations by making use of public opinion. Collaborative Filtering was one approach to achieve that. Another identified approach was to make use of user-data on social media. This has been integrated into Machine Learning-based Hybrid Recommendation Architectures in many ways. In recent research done by Amazon (ibid.) it is understood that when a timeline is considered for recommendations, an *Autoencoder* Deep Learning model is capable of Recommending the best possible combination of movies to users.

2.6 Review of Evaluation Approaches

2.7 Tools

2.8 Chapter Summary

REFERENCES

Ayushi, Smriti and Badri Prasad (Nov. 8, 2018). "Cross-Domain Recommendation Model based on Hybrid Approach". In: *International Journal of Modern Education and Computer Science* 10.11, pp. 36–42. ISSN: 20750161, 2075017X. DOI: 10.5815/ijmecs.2018.11.05. URL: http://www.mecs-press.org/ijmecs/ijmecs-v10-n11/v10n11-5.html (visited on 07/12/2021).

Bartolucci, Silvia and Andrei Kirilenko (Aug. 12, 2020). "A model of the optimal selection of crypto assets". In: *Royal Society Open Science* 7.8. Publisher: Royal Society, p. 191863. DOI: 10.1098/rsos.191863. URL: https://royalsocietypublishing.org/doi/full/10.1098/rsos.191863 (visited on 07/07/2021).

Chen, Rung-Ching and Hendry (June 2019). "User Rating Classification via Deep Belief Network Learning and Sentiment Analysis". In: *IEEE Transactions on Computational Social Systems* 6.3. Conference Name: IEEE Transactions on Computational Social Systems, pp. 535–546. ISSN: 2329-924X. DOI: 10.1109/TCSS.2019.2915543.

Cheng, Li Chen and Ming-Chan Lin (Oct. 2020). "A hybrid recommender system for the mining of consumer preferences from their reviews". In: *Journal of Information Science* 46.5, pp. 664–682. ISSN: 0165-5515, 1741-6485. DOI: 10.1177/0165551519849510. URL: http://journals.sagepub.com/doi/10.1177/0165551519849510 (visited on 07/16/2021).

Chevet, Sylve (2018). "Blockchain Technology and Non-Fungible Tokens: Reshaping Value Chains in Creative Industries". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3212662. URL: https://www.ssrn.com/abstract=3212662 (visited on 04/18/2021).

Dowling, Michael (Apr. 29, 2021a). "Fertile LAND: Pricing non-fungible tokens". In: *Finance Research Letters*, p. 102096. ISSN: 1544-6123. DOI: 10.1016/j.frl.2021.102096. URL: https://www.sciencedirect.com/science/article/pii/S154461232100177X (visited on 07/17/2021).

— (Apr. 29, 2021b). *Is non-fungible token pricing driven by cryptocurrencies? | Elsevier Enhanced Reader*. DOI: 10.1016/j.frl.2021.102097. URL: https://www.sciencedirect.com/science/article/pii/S1544612321001781?via%3Dihub (visited on 06/23/2021).

ERC-721 Non-Fungible Token Standard (2021). ethereum.org. URL: https://ethereum.org (visited on 07/19/2021).

Ferdiansyah, Ferdiansyah et al. (Oct. 2019). "A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market". In: *2019 International Conference on Electrical Engineering and Computer Science (ICECOS)*. 2019 International Conference on Electrical Engineering and Computer Science (ICECOS), pp. 206–210. DOI: 10.1109/ICECOS47637.2019.8984499.

Howcroft, Elizabeth (July 6, 2021). "NFT sales volume surges to $2.5 bln in 2021 first half". In: *Reuters*. URL: https://www.reuters.com/technology/nft-sales-volume-surges-25-bln-2021-first-half-2021-07-05/ (visited on 08/31/2021).

Larry, Hardesty (Nov. 22, 2019). *The history of Amazon's recommendation algorithm*. Amazon Science. Section: Latest news. URL: https://www.amazon.science/the-history-of-amazons-recommendation-algorithm (visited on 05/25/2021).

What are you missing? Using basic machine learning to predict and recommend NFTs with OpenSea data - OpenSea Blog (Jan. 30, 2020). URL: https://opensea.io/blog/analysis/predict-and-recommend-nfts/ (visited on 08/27/2021).