Loughborough University

# Non-fungible token categorisation using machine learning classifiers

by: Manazir Najib
Student ID: B720039

Loughborough University

21COP328: Data Science Project

Supervisor: Dr Mohamad Saada

Submitted on 25th August 2022

# Abstract

Until the broad use of blockchain based advancements, verifying authenticity, ownership and securing of digital assets have been overwhelmingly challenging. They have been susceptible to tampering, scamming and other fraudulent acts which causes clients to fall under dept, financial downfall or complete abandonment of the digital marketplace. Years of explorations and deep research into blockchains, have led to the development and successful deployment of Non-Fungible tokens – NFTs. This are digital tokens, which are assets on the web that can be verified, have ownership rights and transaction history all indented into the asset, therefore it can never be mistaken who owns the digital asset. This frontline innovation proceeds to develop and attract the attention of the globe, while these NFTs are readily being used in regular everyday applications.

This research project highlights the history of the NFT, stating from the foundations of the problems that occur within a centralised system; where the mediator observes too much power and can be involved in fraudulent activities; the necessity of a decentralised one; where a third-party mediator is eliminated, and everything is public so is verified on a blockchain system via everyone and is immutable. The research is focused on thesis that covers the ideas relating to the Ethereum block chain, and similarities of it to a historical public ledger, but with digital imprints intended into each block that holds a history of previous owner and purchases and is updated every time there is an interaction with the digital asset.

The thesis gives detailed description of current successful uses of NFTs as art collection, which are a huge attraction for the public. Influencers are boosting the hype of NFTs which in return is causing prices of NFTs and the crypto domain to elevate higher than ever witnessed in history. However, there are demonstrations of drawbacks from NFTs, lac of security and intellectual property rights. Confusion and anxiety of NFTs that are stored on the decentralised system, but the blockchain is situated in a centralised locale, which can be deleted or destroyed costing consumer a lot of money.

The article also advances to analyse the CryptoPunks dataset which consists of 10000 avatars of 24x24 pixels that are divided into 5 types of beings (Male, Female, Zombie, Ape and Aliens), and give an overview of trends it may follow. It is highlighted that rare avatars such as the aliens which there are only 9 out of the 10000 in the CryptoPunks collection, is valued the highest. A CNN model is created to train the model to correctly classify the data to identify the class of the CryptoPunk avatar. This was very successful will 100% score on validation accuracy within 6 iterations and conclusions giving strong results with a slow learning rate.

This report still has more work to do, however sets a solid foundation for beginners into the crypto space and also give a good strongpoint for the continued research into this topic.

# Contents

# 1. Introduction

## 1.1 Motivations

*'Ideally, I was fascinated by the hype of NFTs and Crypto currencies. I have investments into the decentralised domain, and Ethereum coin. However, I have no knowledge of the meaning of these coins, I was pushed to get involved by friends. This project let me kill two birds in one stone. NFTs and digital assets like the Metaverse are definitely the future maybe not soon, but it is inevitable, therefore this project allows me to learn about the future digital assets, the functionalities and the uses of NFTS. Plus, it gives me an interesting topic for my master's project, and I believe if you are interested in a topic, you are more likely to invest more time and effort into that topic. '*

## 1.2 Aims and Objectives

The project aims to provide an outline of knowledge from the foundations of NFTs and the crypto domain. Feature the issues and advantages of NFTS and smart contracts, likewise, give case studies of real-world applications and problems risen by the use of NFTs. The project will aim to conduct analytical assessment of NFT dataset such as CryptoPunks NFT collection, to observe any patterns or trends that may follow with this collection. This can be further analysed with the use of machine learning algorithms that can identify different types of Punks or even assess the rarity of the punks if it possesses unique traits.

Objectives

The overall Part T project objectives are as follows:

- o Give insight of NFTs and Crypto domain
- o Collect dataset from CryptoPunk NFT collection using API
- o Analysis of NFT data
- o Develop machine learning based of collected data for classification
- o ~~Provide a rarity tool to highlight assets and traits of NFT collection~~

**Updated Objectives**

**Objectives needed to be updated as, it was not feasible to produce a rarity tool program, with the limited time that was left. Due to work/placement commitment in the US with limited access to Wi-Fi and resources it was planned to simplify the project and focus on research than programming.**

- o **Give detailed insight of History of Decentralised System**
- o **Give Overview on Ethereum Blockchain**
- o **Give insight of NFTs with case studies of applications and risks**
- o **Fundamental Analysis of NFT collection dataset**
- o **Model a CNN to identify CryptoPunk class type**

# 2. Literature Review

## 2.1 Introduction

A Non-Fungible Token (NFT) is a non-interchangeable unit of information stored on a decentralised public digital ledger, a blockchain most commonly the Ethereum Blockchain, which can be sold or exchanged. NFT data type varies is increasing as knowledge of NFT and crypto domain are maturing. Common NFT data units are comprised of digital files that include photos, videos and audio.

History gives evidence that a central mediator has been the focal point for the observation and completion of exchanges, purchases and transactions between parties. With the introduction of digital assets, cryptocurrency and crypto space, a new system has developed that comprehensively fulfils the transactions in a decentralised fashion. These done by smart contracts, which are lines of codes that are automatically executed, after the contractual conditions and terms are met, that sanctions the transaction between the parties.

Prior knowledge is required for a full understanding of NFTs and the uses of NFTs that will be discussed, analysed and evaluated in this report. The detailed research into the crypto domain is done below, that highlights centralised

exchange system, fundamentals of blockchain system and the uses of digital open records to generate smart contracts to allow for NFT trading.

## 2.2 Research Diagram

This **Figure 1** shows a research territory map of NFT. The map shows the links between each of the fundamental topics that were researched in detail. These topics are essential to develop a solid foundation of knowledge into NFTs and the crypto domain. The overlap shown in the diagram shows the connection between research area, and the size gives a portrayal of the refinement process of research being observed.
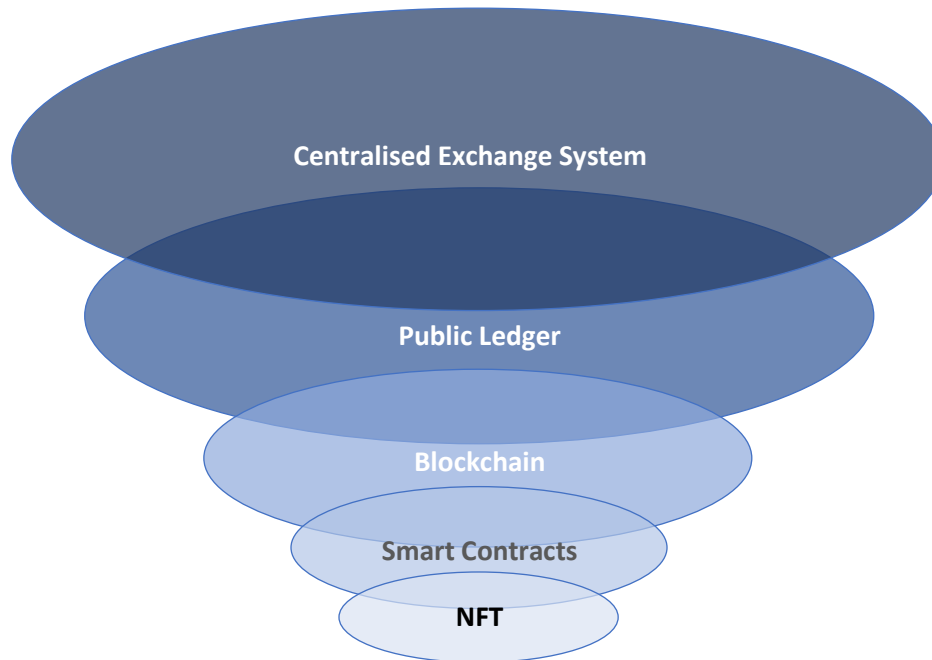


*Figure 1- NFT Territory Map*

## 2.3 Fundamentals of Research Diagram

### 2.3.1   History – Centralised Exchange

Crypto exchange is a commercial marketplace where digital currencies can be purchased, sold or exchanged. In USA, the biggest businesses permit the use of government issued money 'fiat money' such as the USD or GBP. This can be used to exchange for crypto currencies such as Bitcoin or Ethereum. The purchase, trade or exchange is done in a centralised manner.

A centralised exchange is a trade that utilises an outsider (third-party) to mediate the exchanges between dealers and consumers. The "third-party" ideas are the fundamental concept of centralised trades. The common principle employed in traditional banks, are that of a centralised system, thus trust of the 'third-party' is important to conduct trades or even to hold money. Similar to how a bank observes monetary exchanges between parties as a mediator, crypto trading platforms provide an interface which allows clients to purchase and exchange in crypto exchanges.

The bank or the online platform (third-party) has the history of transactions stored on a record. Often the 'third-party' charges money for the exchange to occur, which becomes very profitable for the mediators. For this centralised framework to function for monetary trading (crypto or fiat) the mediators are required to observe a series of regulations and guidelines set by the public authority. An example of this is anonymisation, where the all-customer contact information is hidden when stored on the ledger. (A ledger is a history of all trades a company or party has completed; in crypto exchange this is often a digital ledger).[1][2]

### 2.3.2   Public Ledger

In the past public ledgers were essential for communities, villages and society. This was a method for tracking trades between all habitants of the present town, such as blacksmiths, farmers and tradesmen. When the transactions history has been displayed publicly it works as a verification method by all parties.

In recent years with the increase in volume and interest in trades in the crypto domain, a similar mechanism was introduced that is being used to verify, authorise and record transactions. The public ledger is situated on the blockchain which automatically saves the history of transactions. When a trade offer has been processed the blockchain gains access to the public record for legitimacy of funds and assets between the clients, and the transaction will be approved and recorded. However, if funds are unavailable or have been pledged to another party, the sale is ultimately halted thus making it impossible for a double transaction.[3]

### 2.3.3   Blockchain

To overcome double spending, Bitcoin introduced the blockchain framework, a computerised public ledger employed for trades in crypto domain. The blockchain is a replicated structure of data that is distributed and partitioned among an organization's customers. Each block is unique and is timestamped between each update of the block within the blockchain. Each block is identifiable with its unique 'cryptographic hash,' the hash is a reference indentation of the previous update of the block, apart from the initial block – this is shown below in **Figure 2.**
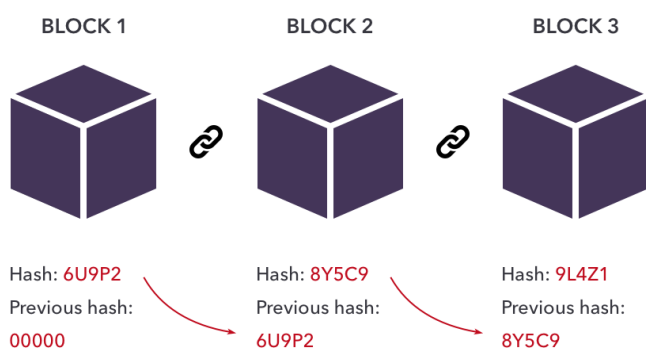


*Figure 2- Blockchain Network [6]*

The blockchain is thus created form the sequence of immutable block that holds the history of the block prior. The encrypted nature (outside alterations are impossible) of the blockchain, gives trust between clients without the need for a 'third-party' mediator.[2]

Blockchain structure has been described as a trust less network, as transactions can take effect without the intermediate being involved, so parties can come into trade without confidence between them. The lack of the mediator ensures quick transaction periods. The extensive use of cryptography is an essential characteristic in the blockchain frame network.

#### 2.3.3.1  Blockchain Network

Here is the detailed complete structure peer to peer interaction of blockchain network-

1.  Clients interact with the blockchain through a combination of private and public keys. The private key is used as a signature, which authorises their own transactions which is then identifiable on the public network via the public key. This distorted coding gives validation, reliability and uniqueness into the association. Each significant trade is broadcasted to other peers via the client's personal hub.
2.  Neighbouring associates relay the transmission after validating the incoming transaction. All invalid transactions are automatically voided and removed. Eventually the entire network has received the same transmission and trade history.
3.  All transactions and transmission get validated within the asked timespan, and the blocks are timestamped upon completion. This is mining, and the mining node advertises the block back to the network after the timestamping.

4. The nodes have to run two important checks before the exchange block is added to the blockchain. (i) the transactions are valid (ii) the reference hash is valid from the previous block. The added block applies the transactions and updates the global view on the network.[4]

### 2.3.3.2 Ethereum Blockchain

Ethereum is fundamentally a machinery software platform that permits construction and deployment of many decentralised applications. Ethereum follows Bitcoin blockchain innovation. The blockchain is compatible with Turing-complete programming language, which is normally used when creating smart contracts on the Ethereum blockchain. This blockchain technology gives versatility to the operator, as it consists of an intellectual tier, with the ability to characterise their rights, trade formats and state conversion procedures on command through the use of smart contracts – which are cryptographic rules collected together for a transaction or exchange and is only executed assuming explicit terms are fulfilled.

For the Ethereum blockchain, virtual currency Ether is used. This Ether is what drives the various applications that are scattered through the network in Ethereum blockchain. As Ether is a transferrable currency, that can be used to for tasks, Ether can utilise the decentralised applications, create smart contracts, create tokens and make traditions exchanges between clients. [5]

## 2.3.4 Smart Contracts

### History

In 1994 Nick Szabo, a first-generation Hungarian American computer scientist, introduced a breakthrough invention in the blockchain framework, with the first smart contract. ``A computerized transaction protocol that executes the terms of a contract'' this quote was stated by Nick. [2] The smart contract interprets the legally binding conditions (security, bonding, terms, etc.) into computer code. The lines of code guarantee that after the circumstances and terms of the contract have been met, the agreement is automatically self-enforced, any break of agreement will be archived digitally, and the guilty party will be punished immediately. The automatic deployment of the contract removes any dependencies from third party mediators, and the prevents perceiving malevolent or accidental trades.[6]

The Ethereum was the first blockchain to execute smart contracts, this prompted the advancement of numerous applications within marketing such as NFTs. Bitcoin is restricted by scripting and language that is permitted on the blockchain, whereas Ethereum accepts computationally universal programming language with high-level level scripting codes. These complex codes are assembled into simpler form of byte code that can be conveyed on the Ethereum network. Then the contract can observe terms, conditions and access the assets within the contract.

### Function

The following gives detail of the processing involved in a smart contract.

**Example scenario**

*"Tom and Jerry decide to get set up a private blockchain network for trade between them. The digital assets being traded are A and B. Tom writes and deploys a smart contract on the network. The contract has three main functions that have been defined.*

- *i) 'deposit' – this allows to deposit units of A into the account*
- *ii) 'trade' – defines trade terms e.g., 4 units of B for 1 unit of A*
- *iii) 'withdraw' - permits Tom to withdraw all resources in the contract*

*Tom deploys a transaction, addressing the smart contract and 'deposits' 5 units of A, this transaction archived onto the blockchain. Jerry, from his total resource of 20 units of B, initiates a 'trade' with 12 units of B, this gives him a return of 3 units of A. This exchange is recorded onto the blockchain and updates the digital ledger. After Tom uses the 'withdraw' function, the smart contract verifies the signatures to identify the correct establishment of resources, and distributes Tom with the updated properties of 2 units of A and 12 units of B. "[3]*

1. When the smart contract is created and established, it has its own state so all assets within the blockchain can be regulated by the contract. Within the blockchain the smart contract has an individual record, and the blockchain supports this feature. Therefore, it is seen as an independent entity, which can handle the entries of the transaction, e.g., from the scenario the contract can hold elements A and B.
2. The smart contract permits business proposals to be compiled into lines of code e.g., trading of "*4 units of B for 1 unit of A*".
3. For a strong, successful contract it is essential that all potential outcomes of the contract have been explored and described within the code, e.g., would be when coding the 'trade' function it is important to consider the outcome if offered a unit of B which is not a multiple of 4. To solve this problem the contract can either reject the trade or parse the resource with equivalent exchange.
4. All transactions are fundamentally data driven. The deployed smart contract ideally exchanges data structure indicating a transfer of value as a response.
5. The initiation of a smart contract upon receiving message/transaction to its address.
6. It is exceedingly difficult to bypass the blockchain network with faulty or non-deterministically written contracts. A non-deterministic smart contract would be spread across every node on the blockchain network; however, this would give invalid and randomised results. This problem is stopped by the smart contract being rejected from the global blockchain mainframe or restricting to a programming language that prevents these concepts. A well-defined smart contract is when the same input value should provide the same output value each time, this is its definitive characteristic.
7. All members of the network can observe and review the smart contract code as it resides on the blockchain
8. Everyone on the network possesses a digital unique key, which is used with every dealing with the smart contract. As the interaction with the digital signature takes place on the blockchain, it leaves a 'footprint' on the network. All other participants on the grid also receive this 'footprint,' this is the verifying step of the clients' interactions with the smart contract. [3]

*Assurances*

When the blockchain used supports the smart contract between parties, the executing parties can be assured of certain aspects when engaging into a smart contract deal–

1. Be able to examine the code written on the smart contract, and view results obtained by this code previously before confirming engagement with the contract
2. Have assurance of successful deployment as the code has already been conveyed into the global network which is operated independently
3. Being able to track and have complete verification of all procedures as each interface is digitally signed onto the blockchain

When all conceivable results are calculated and/or eliminated, there is no chance to dispute transactions and interactions that occur on the blockchain through smart contracts as the clients cannot dispute the outcome of the results embedded onto the crypto framework with this authenticating method.[3]

*Advantages*

Operating smart contracts have accumulated a number of benefits for clients investing money into businesses and transactions.

1. **Consistent Outcomes** – Smart contract offers an approach which can decrease or fully eliminate the 'third-party' mediator like courts or formal ensembles. This is achieved when the transactors agree to be involved into the self-executing contract, this autonomous process ties the client into terms and conditions which have been defined on the fundamental un-alterable code written on the contract. This prevents involvement from other parties and interactions that are beyond the contract boundaries.
2. **Low investment funds** – Mediators that charge high transaction fees are completely removed from the equation; this lowers the cost of initiating into a smart contract.
3. **Security** - There is prominent level of data encryption, which is commonly used in Bitcoin exchanges. This ensures smart contracts are highly protected.

4. **Transparency and trust** -   All trades are encoded and highly secured where all the records are easily accessible by all participants of the network, this gives reassurance to the clients willing to deal in smart contracts.
5. **Accuracy** – Smart contracts guarantee that all fundamental legally binding terms are documented in specific detail. This avoids any issues that can happen due to negligence of a contractor [6]

*Limitations*
1. **Making Contract Changes** – altering a smart contract is impossible or very difficult with inflated costs. Unlike a written text-based contract, which can have drafts written with new law changes. The change to blockchain would be impossible, and this is where the smart contract resides. The yield transaction cost would be too high to even try.
2. **Negotiations** – Businesses can profit from creating contracts that are ambiguous and revealing the knowledge later on which can bring profit to their cause. This is non-existent in smart contracts. For a smart contract to be issued, exact parameters and terms must be given.[7]

### 2.3.5   NFT

Non-fungible tokens, abbreviated to NFT are forms of digital data that are demonstrative of creative physical or digital work, intellectual property which include music, digital art., games and imagery. 'Nonfungible' – this implies that each token is not equivalent or exchangeable for another, this creates the unique attribute of every token which highlights a single specific object. The tokens that carry the digital data can be valued in cryptocurrency. Although majority of the NFTs are part of the Ethereum blockchain, but different from the Ethereum coin which is fungible- 1 coin can be exchanged for another coin- or for other similar assets.

Swift high-tech advancement and security risks develop proportionally to each other. The individualism and non-fungibility of NFTs limit, even completely eliminates any problems with verification and fraudulent work. This is achieved through the incorporation of digital signatures into each individual token, such that the owner can locate the asset effortlessly. Likewise, these benefits consumers purchasing items such as tickets or graphic arts as they can easily track the original owner of the item, thereby solving the issue of deceptive and counterfeit items and giving confidence to the consumer for a legitimate purchase. NFTs have paved a new path for art and creative businesses, which struggled in online marketing previously due to lack of 'exclusive ownership' on the web. [4]

*NFT History*
NFTs were first generated in 2014, which mixed art and technology. This was initially created to allow artists to monitor and safeguard their work by establishing ownership and legitimacy on a blockchain. The NFT scene had a significant milestone with the emergence of two NFT projects, CryptoPunks and CryptoKitties in 2017. The two projects are relayed on the Ethereum blockchain to carefully store, safeguard, and validate ownership of digital assets.

NFTs audience grew in October 2017 with the CryptoPunks collection, but a single sale form Mike Winkelmann drew even more attention to the NFT domain in the digital world. His digital art collection sold for about $70 million; and following this sale the trend and interest for NFTs has been increasing exponentially from other designers and art fanatics. Beforehand, NFTs were neglected and observed as a restricted asset on the blockchain network, yet have created unique market for themselves, and as of July 2021 has trading volume of $1.2+ billion.

*Deployment*
Digital art is embedded onto the Ethereum blockchain through a process called minting. 'Minting' is when NFT is created and put into distribution on the blockchain network much like new coins when they are distributed into circulation. NFT represents the digital art, which can be exchanged on the market, with all interactions being tracked on the web.

It was recorded in 2021, that within a 24-hour normal trading day, around $4 billion is being traded, and the trade volume is always increasing massively every day. The NFTs are traded on platforms on the web that are easily accessible to both creators and consumers, using their crypto wallets.

With the high demand for NFTs recently various marketplaces have been developed for trading digital assets like NFTs. Figure 3 is show below, this highlights some of the biggest marketplaces for NFTs on the web today and the

volume of trading occurring in the websites. Different marketplaces exchange in different NFTs and maybe even on different blockchains.

| Market | Volume |
|---|---|
| OpenSea | $6.5bn |
| Axie Marketplace | $2.1bn |
| CryptoPunks | $1.3bn |
| Rarible | $210m |
| SolSea | $79m |

*Figure 3 - Top Marketplaces [6]*

Opensea is on the Ethereum blockchain. This marketplace trades in-game items, collectibles artwork, music and GIFS, these are all forms of NFTs. Currently they are the industry leaders in global NFT trading.

Axie Marketplace is the official trading centre for the NFT game Axie Infinity. This market is coded for purchasing Axies, Territory and more in game items. Axie are digital fantasy creatures, raised as pets, trained and used in battling other players in Axie Infinity NFT game.

SolSea is used for exchanging similar assets, however these NFTs are stored, sold and bought on the Solana blockchain.

## Applications
### Digital Art
Digital art consists of music, videos, paintings, images and more that is on the virtual or digital medium. This digital content much like physical artwork can be sold or created, there are artist and art collectors. Physical art however is defenceless, it can be stolen, or counterfeit art could be made and sold instead. NFTs give artists and consumers some reassurance when trading. NFTs embed a hash into each piece of digital art or etc. that has been submitted into the blockchain, this makes each individual piece unique. It is possible for the developer of the NFT to mark the digital asset with their signature, which builds credibility for the artist and the produced token. NFTs give assurance that all copies made of the original token, is owned by consumer, where one token is not interchangeable to another, this is a huge fascination to art collectors. [10]

Artists and creators benefit generously from utilizing NFTs and blockchains to trade art. The creator's profit from the sale of each token, likewise the profit is boosted further with royalty bonuses given to the designer when the original piece is transferred to new owners every time. The idea of royalty for the designer used to be considered illogically, as after several transfers of the asset ownership of physical artwork is hard to keep track, whereas NFT now establishes a solution to this problem where the artist is well reimbursed for recurrent sales. [3]

### Collectibles
Collectibles are a major player in the NFT domain, with the introduction of CryptoPunks and CryptoKitties. Later collectibles followed these trendsetters that were opened to the public marketplace. In 2017 CryptoPunks was released into the Ethereum network. CryptoPunks each have unique characteristics and appearance that come in various combinations. Each CryptoPunk are produced with its own algorithm and unique traits, this along with its rarity defines the value of the token.

### Fashion
Extravagant fashion brands are opting to include NFTs for their assets, where they can utilize the NFT's properties to individual possession, perpetual quality and 'royalty acquisition.' [5]. Large ecommerce fashion brands utilise online presence and marketplace to grow their infrastructure and attract more customers. However, the problems they face are other third-party marketplaces on the web that sell counterfeit products with the same logos and brands, which are making the original owners of the fashion brand lose business. These problems caused by the fraudulent designers can be stalled or completely eliminated with the use of NFTs.[10]

During the global coronavirus pandemic fashion industries have excelled their attempt to embed NFTs into fashion tech, especially due to the closure of physical stores. Organizations have currently started installing digital NFTs on to their resources to establish ownership and hold distinctiveness of their brand.

It is still very unlikely that NFTs will totally replace fashion wear, as fashion industries highly rely on sale of physical products. However, NFT's can still prove to be an extravagant design expansion for fashion businesses.

## Licenses

Companies expend considerable effort and time on authenticating critical documents. These issues can be solved by assigning NFTs to each individual documents for each customer, this will improve administrative processes. Similarly, the issuer of the endorsements and licenses can dispose of their responsibility for tracking and recording reports for the documents, as the marked NFT can be used for validation and tracking on the network. The distributing of the certificates and licences on the blockchain make them impervious to any alteration and fraudulent activities, this ultimately lessens the probability of experiencing counterfeit records.[4]

## Gaming Industry

'Play-to-earn games', this is now the common name for NFT games. These games award players with NFTs or other tokens after playing a certain amount of time. Players may even purchase game items. However, if the acquired item is an NFT, the player may be permitted to trade again on the market, some players may even make a profit as price fluctuate over time.

The mix of NFTs into blockchain games, has boosted game economies profoundly and will continue to do so within the Metaverse. Further development of NFTs where playtime is rewarded with Currency, and NFTs can boost the rewards, or influence better item drops. Currently several games, DeFi Kingdoms and Axie Infinity accommodate these features but will be developing even more new elements and processes to best utilise NFTs. Axie Infinity is a popular play-to-earn game, where the game prize is a 'Smooth Love Potion'- SLP NFT. These tokens are widely listed on the crypto market such as Binance where players can trade these on the crypto domain for ford money (cash).[20]

## Domain Names

A more unobtrusive application of NFTs is the creation and ownership of blockchain-based domain names, many use Ethereum Name Service (ENS) or 'unstoppable domains' to gather attention for the domains. Clients can simply long, undesirable web addresses to much neater attractive address lines on the web, this is more inviting and user friendly. There is very high demand for personal domain names, which is tough on ENS and 'unstoppable domains', however thus far they have been quite successful in completing and generating all requests.[3]

## Virtual Reality

Virtual Reality worlds are constantly developing, and the technology is advancing rapidly. These worlds can have digital assets in forms of NFTs which can influence gameplays and also the worlds can be controlled via smart contracts. Assets such as land, buildings and other resources can be owned, and the service provides can create a smart contract for the sales of these assets.

Assets in the real world need to be valued by economic institutions using a variety method, that may be questionable or unreliable. This causes an issue of trust for the clients in the real world to accept the determined price of assets. Likewise hiring an outside auditor to evaluate the authenticity of prices is also a relatively expensive procedure. Consequently, assets and items remain undervalued and creates disruptions between exchanges and offers which can make asset owners lose confidence in the exchange.

On the other hand, in the virtual domain any form of digital interaction or activity is based upon the blockchain technology. As all trades are recorded onto the Ethereum smart contract, clients do not have to worry about exchanging assets in the virtual world in the crypto domain, this also give them more confidence in the nature of the asset.

## *Challenges*

## Centralisation of a Decentralised System

Although NFTs and smart contracts revolve around a blockchain network which is a decentralised system. The marketplace where assets and NFTs are being sold is a centralised system. This can be a concern if big marketplaces like Opensea where a lot of NFTs are exchanged every day, and it is stored on the web address, the owners of this wed address can block purchases, and delete the contents on the market network.

Furthermore, another concern is where the NFTs like digital artwork are stored. The items being sold and placed on the blockchain is actually hosted from a centralised location, where the servers can be deleted meaning all contents could be lost. Web addresses that have been deleted can affect the NFT, as the address of the NFT will only give error messages, even though it has been purchased.[12]

## Evaluation

Pricing NFT is a difficult problem to solve and is one of the primary reasons for clients avoiding business concerning NFTs and cryptocurrencies. The prices of NFTs are forever fluctuating and is rarely stable for long periods of time, this is due to various reasons. The current value of the crypto market has a big role to play, but more relative terms would be the creativity, individuality and scarcity of the NFT. This cannot be measured and have abstract views of valuation, so it makes pricing NFT difficult. Another issue with evaluating the quality of NFTs is the power the social media and celebrities can have to influence the hype and trend of NFTs and stock market in general. Trade volumes of certain NFTs can be boosted if a celebrity has purchased or mentioned the collection example there was a huge boost to the trading volume of BoredApes NFT collection as Neymar Jr. purchased one and tweeted about it in early 2021. [8]

## Legal

Classifying NFTs and considering legal terms for NFTs have been a major issue, it is not the same across the world. Countries such as UK, Japan and the US have different approaches of defining NFTs and classifying them. Countries like France and Malta have implemented suitable principles that regulate interaction of the digital assets, whereas other regions are already existing laws. However, this can sometimes be confusing to the consumers as they are required to follow these complex rules which may sometimes even be contradictory. NFTs have been attracting consumers and designers progressively, therefore it is necessary to establish a worldwide regulation that can be used to identify and trade NFTs proficiently.

Another problem revolving around the legal field is intellectual property that defines the individual's possession and tights to specific NFTs. Prior to purchasing an NFT, it is essential to determine authenticity of the merchant trading the NFT. Issues of fraudulent parties have increased where malicious parties selling and replicating multiple copies of an NFT. Therefore, the buyer can use the bought NFT, however they do not get ownership of the NFT, which means if it is deleted by the owner the buyer would lose the data of the NFT.[3]

## Security and Privacy

Anonymity and privacy of NFTs still require further advancement. Data stored on the platform can be lost or misused by malicious parties. Currently NFT transactions depend on pseudo-anonymity which is provided by transactions occurring across the Ethereum platform instead of complete secrecy. Complete anonymity has not been achieved yet, because the user can be discovered if the connections can be made from the corresponding user network address and true identity. Then clients' activities will be fully exposed to the public.

Security concerns over smart contracts have been a major issue with the continued development of NFTs, A recent attack on DeFi network has hindered the safety of smart contracts and NFTs. The cyber hacking concluded with the company losing $600 million. Although smart contracts have security and all transactions and lines of codes are immutable, however there is no definite template that can provide functioning terms for each different smart contract and transactions. Likewise, another problem is the designers of the contracts lack proper instruments to create, verify and review the smart contracts.

## Environmental

Evidence has risen to show the detrimental effect to the environment from the huge growth in NFT market and overuses of blockchain environment through various crypto transactions. Data shows that Ethereum network consumes 45 terawatts-hours of power, countries like Qatar and Hungary have similar power consumption. [8]

Minting NFTs (large scale) and uploading them onto the crypto domain and blockchain marketplace requires a lot of energy. This a result demands a lot of gas fuel consumption which has been increasing in price exponentially. Smart contracts require codes to run which require computational power (complicated contracts require more power), and also need a lot of resources for storing data and processing, this power is consumed after every transaction.

# 3. Software/Tools

For this thesis, a few key tools and software have been used to appropriately gather data, filter out unwanted data and run analysis on the data set. Likewise, to develop machine learning algorithms for classification we need to identify which methods would be best suitable for the CryptoPunks NFT collection dataset. Full explanation of software and tools are shown below.[14]

## 3.1 API

'Application Programming Interface', API is innovative piece of software which permits two applications to communicate and exchange data with one another. The API is a courier that conveys one parties request to the provider which in return conveys the response back from the provider.

Independent functionalities with induvial operations are defined and characterised by the API, this enables the implementations and definitions to function without any overlap or compromise. Therefore, APIs are an excellent foundation for program development. This gives ease to designers as they do not have to create code from scratch, due to the high reusability of APIs. The reusing ability of APIs permits developers to efficiently process complex methods and significantly increase speed of development. [16][18]

### 3.1.1 Moralis's NFT API

With NFT APIs, the function described above is used to exchange and acquire data from an NFT marketplace. For data acquisition in this thesis, Moralis's API is used to collect data of CryptoPunks NFT collection from Opensea marketplace. From research this API seemed to excel at performance more than others and is easy manageable. The execution can be done upon one request to the marketplace to gather data from a time range, whereas other APIs would require multiple contacts with Opensea.

One line of code can be used to gather all intel from the NFT collections, intel such as:

- NFT ownership
- Transaction details
- Contract type
- Minting time
- Token Address
- Metadata on the NFT

Upon receiving the required data through the use of the API, the data was sorted chronologically, serialised and stored on the Google drive. This is ultimately the dataset that would be used for data analysis and machine learning procedures, this prevents the need for requesting frequent requests to the marketplace to get data, this saves time and prevents traffic load on the network.

## 3.2 Python/Colab

The programming and data analysis is developed in Python environment, more appropriately Google Colab. The machine learning will require high memory capacity and fast analysis, Colab uses python interface but over online servers with high GPU capacity for peak performance. Python/Colab is the most appropriate programming language used for this project, a preference over R or Tableau, as complex investigations will be performed with development of machine learning algorithms, such as neural networking. Benefits to using Python is that it provided an easily to manage user interface, with a simple syntax. Contains a huge standard of libraries that helps with uploading, reading and cleaning data. This software open source and can be executed on any platform (Mac OS, Windows or Linux), also the notebook allows for easy readability and progression of the project with clear apparent analysis and visualisations to support the evidence.

## 3.3 Machine Learning

Deep learning

Convolution Neural Network

TensorFlow has been used to optimise the building of the deep learning models- this library is proficient for leaning and training iteration which are executed on complex structures and large datasets. This library is commonly used in

Python and many useful tools embedded within, an important one is the tools that restrict overfitting of the machine learning models.

Keras package is also used within this project, another library index compatible with Python. This works as a development API, that is efficient for constructing model with 'experimental neural network implementations. Keras is highly adaptive package so is very versatile with all data sets. [20]

# 4. Data Handling and data summary

## 4.1 Data Choice

Data for NFTs are endless, and there is a broad choice of data to select from. For this research it has been allocated that a collection NFT is used. CryptoPunks is a collection of NFTs containing 10000 unique 24x24 pixel images that have been tokenised on the Ethereum blockchain since 2017. Each CryptoPunk avatar has a unique number of traits and attributes. The designs vary from plain individuals, zombies, apes and aliens to these selections with variety of clothing and accessories. Figure 4 shows an example of some CryptoPunk NFTs



CryptoPunks ⬤          Last ◆ 75.88      CryptoPunks ⬤          Last ◆ 69.93      CryptoPunks ⬤          Last ◆ 60
CryptoPunk #9763                          CryptoPunk #1385                         CryptoPunk #1

*Figure 4- Example CryptoPunks*

CryptoPunks have a lot of data, therefore research into this is very plausible with a lot of information to make predictions and training models. Main hype for CryptoPunks is that it is one of the oldest NFT collections that was introduced into the Ethereum blockchain, likewise there is a lot of history and sufficient transaction patterns which can be analysed and used for future referencing, purchasing and even advising. [21][22]

## 4.2 Dataset

The data set was gathered as described above using the Moralis's NFT API. A range of dates were selected from July 2017 to October 2021, this should gather enough data for good modelling and training parameters. the data has 10000 unique avatars – 6039 male, 3840 female, 88 zombie, 24 ape and 9 aliens. All of them have a unique randomised appearance with assets like beanies, beards, glasses etc, or some have no attributes at all.

For the given timespan there is a total of 167492 data entries, each with the columns highlighted in Figure 5 below.

```
[ ]  df.columns


    Index(['txn_type', 'from', 'to', 'date', 'timestamp', 'source', 'eth',
           'punk_id', 'from_wallet_address', 'to_wallet_address', 'type',
           'accessories'],
          dtype='object')
```

*Figure 5- Columns*

### 4.3 Missing Data

The dataset has been very well kept, and as it has been embedded into the Ethereum blockchain, all fields of the data are completed. This shows the befits of using blockchain networking.

### 4.4 Cleaning Data

However, there have been cases of duplicate datasets, which have been removed with simple lines of coding that can be seen in the programming language. Likewise, not all fields are required for the data analysis and machine learning, so the dimensions were reduced. For the data analysis the columns were narrowed down to the following-

**txn_type** – this shows the interactions with CryptoPunks (Bids/ Withdrawals)

**date** – date of transaction – had to be parsed into date type from numeric

**eth** – price for the trade

**punk_id** – identification of the CryptoPunk

**type** – highlights the attribute

**accessories** – gives an overview of the appearance of the avatar

Figure 6 shows a typical row with all usable data.



| | txn_type | date | eth | punk_id | type | accessories |
|---|---|---|---|---|---|---|
| 0 | Bid Withdrawn | 2021-09-04 | 321.0 | 0 | Female | [Green Eye Shadow, Earring, Blonde Bob] |

*Figure 6- Sample Entry*

### 4.5 Understanding Data

While observing further into the data gathered it was evident that not all 160000+ entries were actually required, as some were duplicates whilst some were irrelevant bids that were rejected immediately. Therefore, a method used was to highlight the media sale value for each individual punk type and eliminate any bids that are lower than 50% the median. This approximation took a while to run on Colab but reduced the total data entries to 123848. However now all trends and prediction made will be more reliable and will have minimal anomalies due to the omission of 'un-serious' bids.

Punk accessories are given as a list varying from 0-7, with total of 87 unique accessories, this will need to be split up and analysed separately to prevent confusion. The type of column is split into 5 criteria described previously, therefore if required this can be encoded for easier training models.

The punk_id can be cross referenced with where the data is stored, or where the images are stored for the avatars to display the avatars, this is useful for CNN.

# 5. Analysis

*Link to code provided in Appendix as well as submitted separately.*
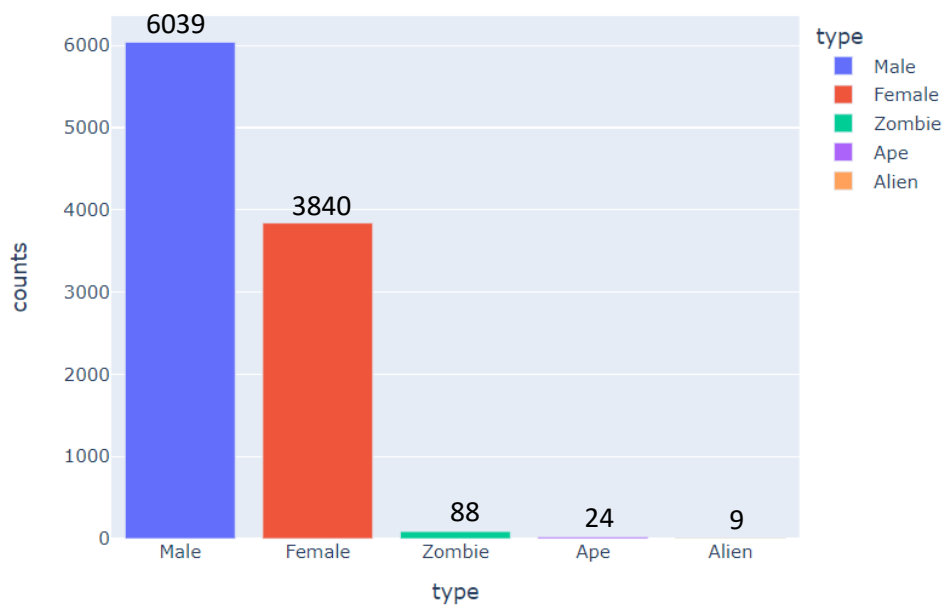
## 5.1 Fundamentals



*Figure 7- CryptoPunk Type Count*

**Figure 7** gives the primary analysis of the given data; this confirms the 5 different types of CryptoPunks listed above. It is clear that majority of the punks are from the types 'Male' followed by 'Female'. The other variants seem very rare with the lowest count shown by 'Alien' group of only 9 out the entire 10000. The types are colour coded for visualisation, but the proportion of 'Ape' and 'Alien' are so low relatively that it is difficult to view in the bar chart above.
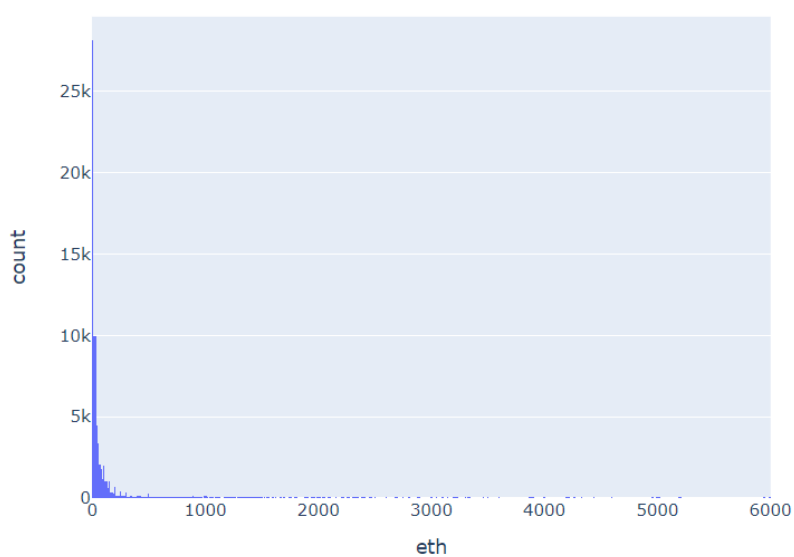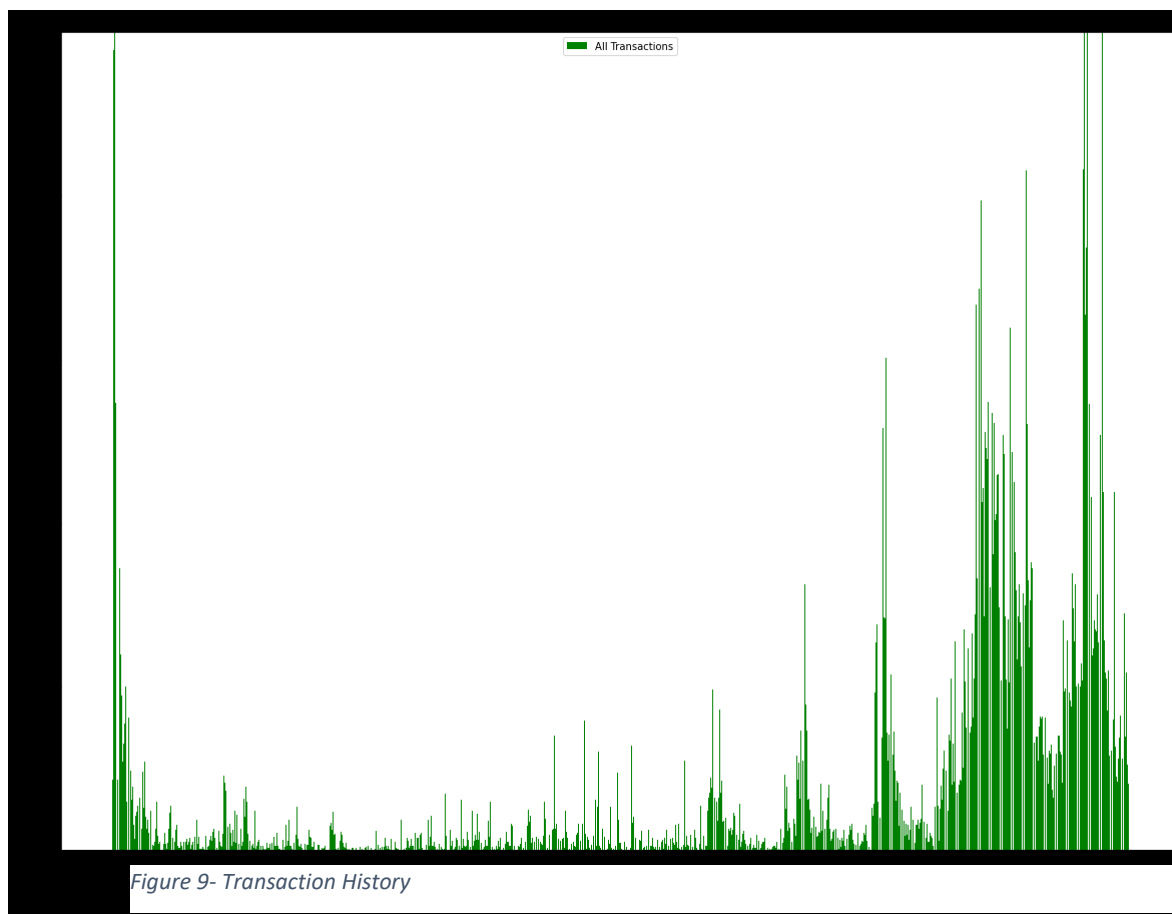


*Figure 8- All Transactions*

**Figure 8** displays the frequency of CryptoPunk transactions in a histogram. Most of the transactions are relatively low cost at 2.5 Ethereum with 28000+ transactions. Although the scaling is rather larger the highest transaction actually reaches 4200 Eth, which equates to about $15 million during that period.



*Figure 9- Transaction History*

**Figure 9** gives an evaluation of the number of transactions that has occurred every day since July 2017- October 2021. Recent transactions reach above the threshold and surpass the 1000+ limit.
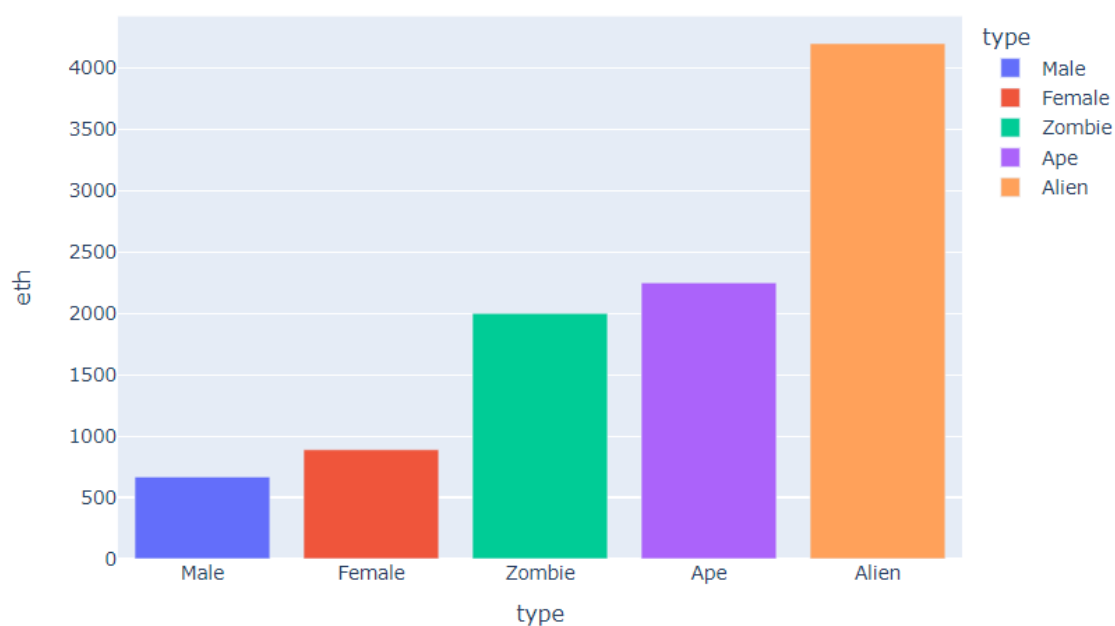


*Figure 10- Max Sold by each Type*

**Figure 10** encases the previous graphs (Figure 8) and shows what class type of CryptoPunk has sold for the highest price. This shows the 4200 eth (or $15 mn) was from the Alien type avatar. Whereas the lowest priced Avatar was from Male group. This data is however only representative until October 2021, prices and evaluations may have changed since.



*Figure 11- Price Trend*
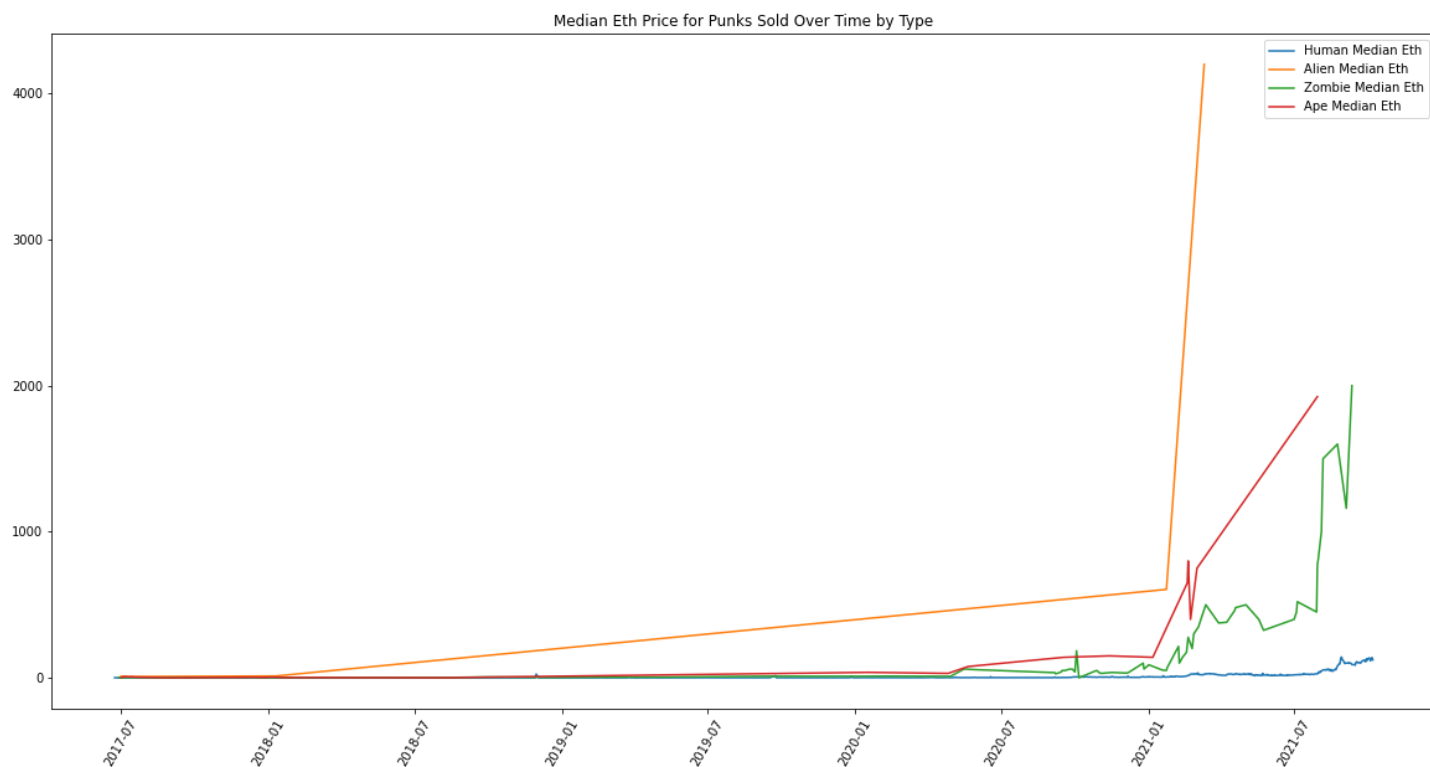
**Figure 11** shows the sale history of the different types of Punks. The median value was calculated and then plotted (colour scheme decided by type), this was again plotted from July 2017 – October 2021. Alien again is off the charts with a sudden rise in pricing at the beginning of 2021 and following the same trend Human avatars are the lowest priced.
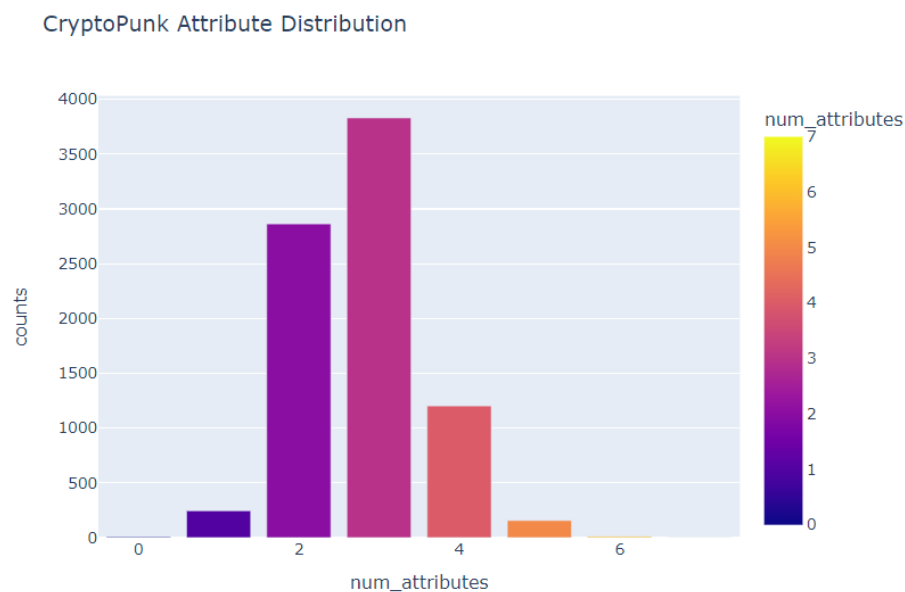


*Figure 12- Attribute Distribution*

**Figure 12** portrays the distribution of attributes given to the full randomised set of 10000 CryptoPunks. The Median (most occurring) occurring value of attributes is 3 with 3830 avatars. And the least occurring with only one entry is 7. 0 attributes and 6 attributes are registered very low as well with 8 and 11 respectively.



Avg. Price of Human CryptoPunks compared with Attributes
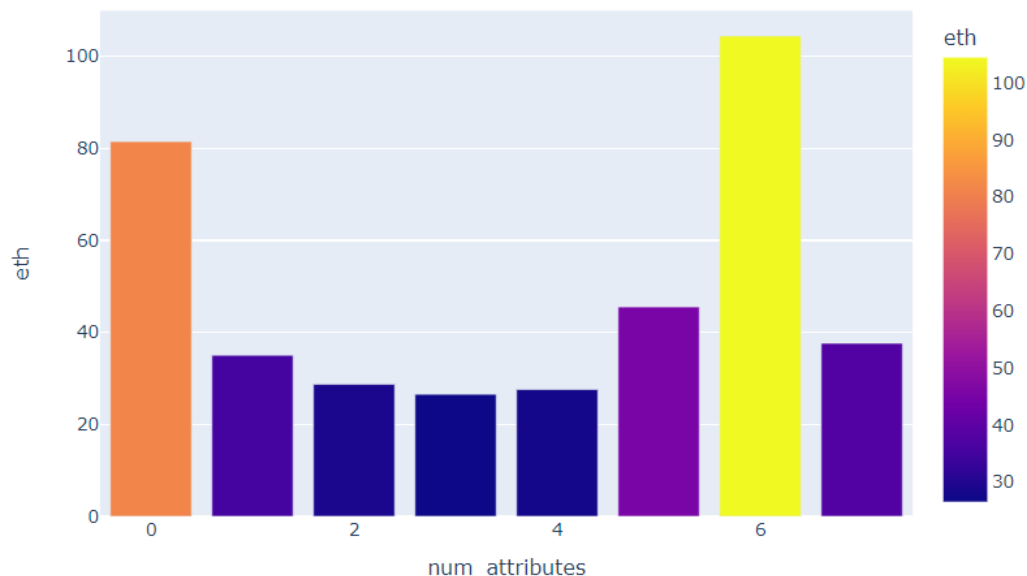
*Figure 13 - Human Attribute Pricing*



Avg. Price of Alien CryptoPunks compared with Attributes
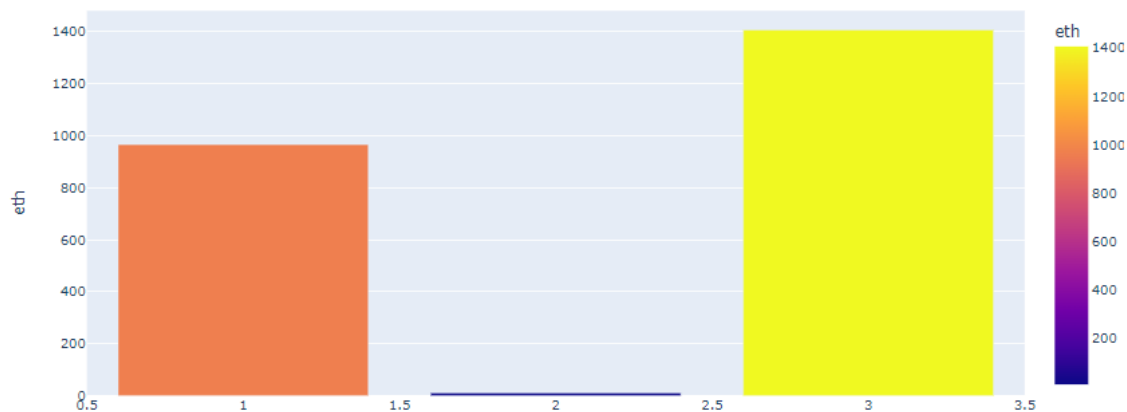
*Figure 9- Alien Attributes Pricing*

**Figures 13** and **14** are plotted for a comparison between the rarest type and the most common type of avatars in the CryptoPunk collection. Humans with 6 or 0 attributes give a higher value than normal ones, but only peak at 104 Ethereum, whereas 3 attribute Alien is evaluated at 1406 Ethereum average price ($4.9 mn). Aliens do not have more than 3 attributes, but humans cover all 0-7 (8 possible).

## 5.2 CryptoPunk Type Classification

### 5.2.1 Data

*Dataset*

The data set now consist of 10000 pictures of the CryptoPunk NFT collection with punk_id and attribute list. Therefore, there are 5 types which the avatars are split into those are Male, Female, Zombie, Ape and Alien. The CryptoPunk generation is random therefore it is unnecessary to randomise the data before splitting into training and validation and testing sets. The initial split is 70-30% in to training and testing with a further 10% from the testing is made into validation set.

*Preliminary Analysis*

It was important to glimpse at the images and understand what the raw data is and how it can be optimally used for training.
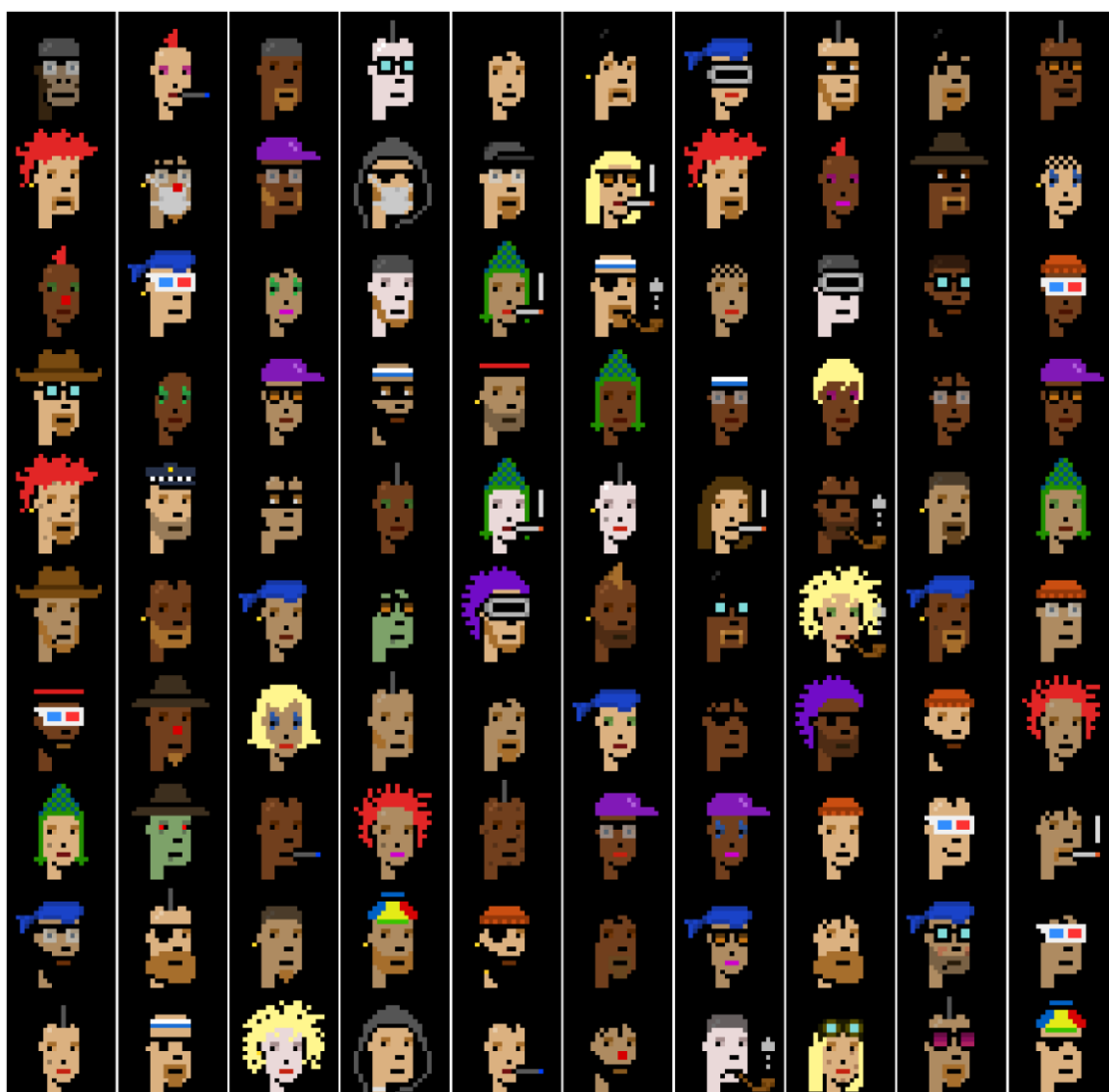


*Figure 1510- Random Sampling of CryptoPunks*

**Figure 15** displays a random sampling of 100 CryptoPunk images, some punk may look identical but, there are very slight difference. None of them are the exact same. This initial overview can give an insight of how features of the avatars may be Identified during CNN, for example spiky red hair, seems to be an easy indicator for Human types, likewise with the green skin its highly likely to be zombies. CNN can also distinguish attributes such as beards, glasses, eye patches and bandanas. These features will really help develop the model is an accelerating manner.

*Pre-Processing*

The images were processed into Colab using TensorFlow functionalities with bath size of 32 and image size was increased to 160, no randomisation was required as CryptoPunks are generated randomly anyway. The underlying image data was encoded into 0-255 numerical values which are representative of the RGB pixel values. After rescaling the samples, they were allowed to pass through the CNN layers. This was done by normalising the data. This step is essential for the pre-processing, this ideally rescales the previous values to lie within 0-1, restricting the range of the data.  A suitable batch size is important, as the batch size determines the error gradient of the network.[25]

### 5.2.2   CNN Model

```
# initialize model
inputs = keras.Input((42, 42, 3))

x = layers.Conv2D(filters=wandb.config.filters, kernel_size=(3, 3), activation="relu")(inputs)
x = layers.MaxPooling2D((2, 2))(x)
x = layers.BatchNormalization()(x)
x = layers.Dropout(wandb.config.dropout)(x)
x = layers.Conv2D(filters=wandb.config.filters, kernel_size=(3, 3), activation="relu")(x)
x = layers.MaxPooling2D((2, 2))(x)
x = layers.BatchNormalization()(x)
x = layers.Dropout(wandb.config.dropout)(x)
x = layers.Conv2D(filters=wandb.config.filters, kernel_size=(3, 3), activation="relu")(x)
x = layers.MaxPooling2D((2, 2))(x)
x = layers.BatchNormalization()(x)

x = layers.GlobalAveragePooling2D()(x)

outputs = layers.Dense(units=5, activation="sigmoid")(x)
model = keras.Model(inputs=inputs, outputs=outputs)
```

*Figure 16- CNN Model*

```
Layer (type)                    Output Shape            Param #
=================================================================
rescaling_3 (Rescaling)         (None, 160, 160, 3)     0

conv2d_9 (Conv2D)               (None, 158, 158, 32)    896

max_pooling2d_7 (MaxPooling     (None, 79, 79, 32)      0
2D)

conv2d_10 (Conv2D)              (None, 77, 77, 64)      18496

max_pooling2d_8 (MaxPooling     (None, 38, 38, 64)      0
2D)

conv2d_11 (Conv2D)              (None, 36, 36, 64)      36928

max_pooling2d_9 (MaxPooling     (None, 18, 18, 64)      0
2D)

flatten_3 (Flatten)             (None, 20736)           0

dense_7 (Dense)                 (None, 64)              1327168

dense_8 (Dense)                 (None, 10)              650

=================================================================
```

*Figure 17- CNN Architecture*

This model was made from a pre-existing template which utilises TensorFlow and Keras bundle in Python to construct a functioning convolutional neural network used for the classification.  **Figure 16** shows the codes to create the CNN modelling layers and **Figure 17** gives the architecture model for the CNN, this is discussed in detail as follows:

*CNN (conv2d)*

A straightforward premade CNN template was followed and changed appropriately to observe the process for this NFT dataset. Initially the kernel count is a major factor for the model, which is used to extract the features given from the images. Customary methods were followed where padding for the input and output are recorded to be the 'same'. A globally acceptable activation function was utilised. 'Relu' is often praised for its speed and efficiency, and simplicity. Experimental data has proven in the matter of deep learning events and training a model, data is processed faster and with more precision when enabling the 'relu' function instead of the 'sigmoid' ability.

**Figure 17** highlights the uses of 3 convolutional layers all using the 'relu' activation function, but each layer is decreasing in dimension as it passes through each layer down. [23]

*Pooling Layer(max_pooling2d)*

These layers are mainly for decreasing the size of the elements present after each convolutional layer, as a result less computational power is required, and fewer parameters need to be observed for the model to learn from. After the images pass through the convolutional layer above, this layer summarises them, and progresses forward to the new layer with the summary of the data instead of individual elements. The most common strategy used is MaxPooling, which is best for extracting the most prominent elements from the layer above. 3 MaxPooling layers were placed as shown in **Figure 16/17** which halved the dimensions after each passing. [24]

*Flattening layer(flatten)*

This is the penultimate layer which flatten the 3D construction that has developed from all the layers above compiling together and manipulates them into vectors from where the classifications can be determined. Figure 15 shows the use of the flattening layer after the successive use of 3 CNN layers. [24]

*Dense Layers (dense)*

All the neurons come together into the dense layer; this establishes the deep connection this layer has with the overall neural network. Commonly the layer is the final layer in a neural network model. This layer computes multiple matrix-vector calculations, which provided values and parameter used for the training and updating (backpropagation).[24]

It is preferable for the dimensions of the dense layer equal the number of classification elements; in this case it is 5. **Figure 16** highlights this at the bottom, but this uses the 'sigmoid' function

This is an additional tool used to simplify the readings from the CNN model. Weights and Biases is commonly shortened to W and B, are instruments for learning in neural networks. As described above neurons are what connects each of the layers described in figures 14/15. This tool relays the information between neurons with the weights and bias dependant on the number of data entries for each class available. Wights influence the effect the input will place on the outputs whereas the Bias is a constant which is always added to the layers, this ensures that there are always neurons being activated even though weightings might equate to 0.
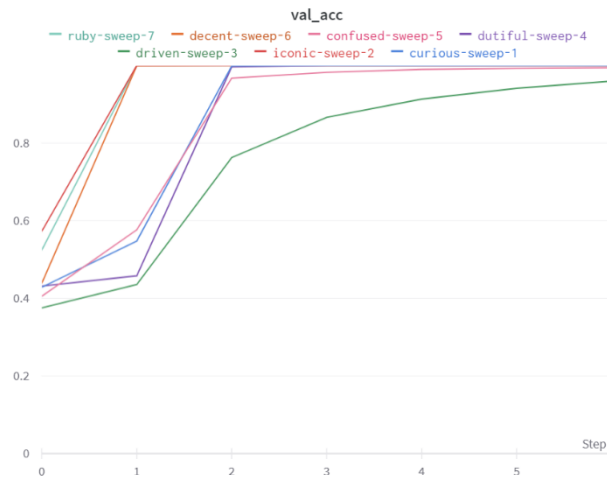


*Figure 11- Validation Accuracy - AUC*

**Figure 16** shows the AUC metric for each sweep. All apart from the one tends towards 100% after just 6-7 Epochs. Certain well performing sweeps reach full score after 1 Epoch.

Epoch is the number of cycles the algorithm will undergo to refine the accuracy and build a better model. The iterations use a lot of memory space and power consumption, but unfortunately the result does not require further analysis with extensive iterations.

**Figure 17** shows how each sweep performed with varying values for each criterion. Main criteria for concern are learning rate and filters as well as accuracy and validation accuracy (discussed previously).  Accuracy for all sweeps is very good the lowest performance giving 94.7% score Learning rate of 0.00005 has provided the highest validation scores, this is also true with high number of filters. [22]



*Figure 12- Weights and Biases Scores*

# 6. Discussion

## 6.1 Fundamentals

The fundamental analysis was required for an initial understanding of the data and to distinguish patterns and diversity within the data set. From the initial ba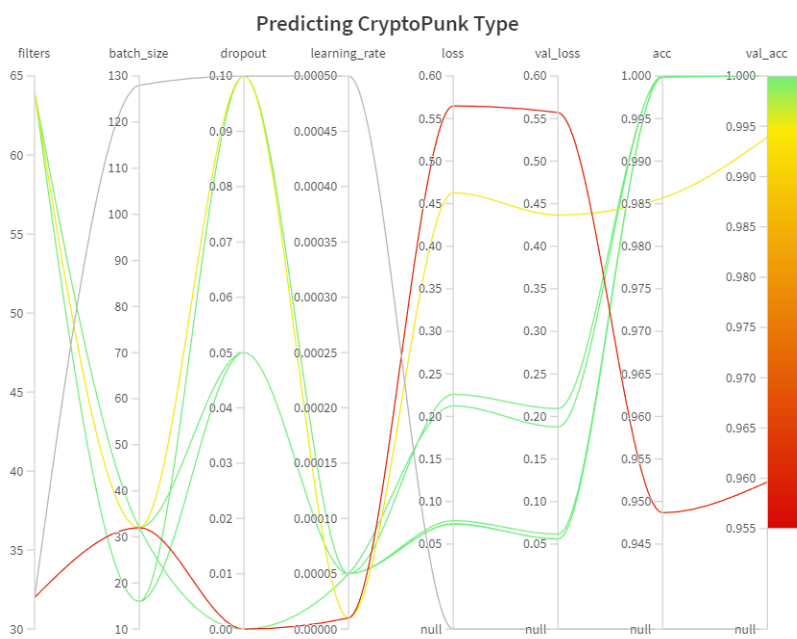r chart, **Figure 14**, representing the counts of the different types of avatars present in the CryptoPunks collection, the human avatars make up 98% of the total data. This mean not enough evidence will be found for the other types, which are reliable. However, for the human avatars very good predictions can be constructed.

**Figure 8** give evidence that a lot of the NFT collection has been exchanged several times, for relatively low prices. **Figure 9/10** confirm that these low cost NFTs are from the Human type, most likely Male. Figure 9 also gives evident of the research that has been conducted, which states that there has been a superior hype in recent years for NFTs and generally within the crypto marketplace. The number of transactions increase exponentially from the start of 2021 and reach beyond the scope record by July of 2021.

Reasoning behind **Figure 9** displaying the overly high cost of Alien type Punks, can be determined by the rarity of these avatars. It has been recorded that there is only 9 presents out of 10000, this scarcity heavily influences the market and demand for this specific avatar.

**Figure 11** gives similar results shows the explosive jump for the demand of Alien type CryptoPunk, again due to the rarity of this punk. The spike occurs at the beginning of 2021, when crypto domain started treading heavily. This class type is followed by the Ape and Zombies, which are also fairly rare in comparison to the 98% made up just by the Human avatars.

**Figures 12/13/14** refer to the distribution of attributes for the CryptoPunks. **Figure 12** shows that most avatars consist of 3 attributes and very few have the polar extremes 0 or 7. **Figures 13** and **14** are comparing the Human avatars with the alien avatars. Humans cover all 8 units of accessories with the most valuable asset having 6 attributes, this has sold for 104 Ethereum, followed closely by an avatar with 0 traits which was sold for 81 Ethereum. Contrastingly Alien avatars although threes only 9 total, they only cover 3 attribute tiers, 1 -3. However, the Alien avatars overpower the human is cost the 1 attribute and 3 attribute CryptoPunks sell for 966 Ethereum and 1406 Ethereum respectively. There is no definitely result or predictions from these attributes, nothing decisive can be concluded, as there is no pattern that 0is followed when evaluating the price of the NFT that also agrees with the number of features the NFT has.

## 6.2 CNN

Before building the CNN modelling it was important to get the understanding of the basics of the data, this model was designed to try and distinguish between the 5 types within 10000 CryptoPunks, Male 6039, Females 3840, Zombies – 66, Apes 24 and Aliens – 9. From this a basic performance model can be created which his that Male avatar population coves 60.39% of the types, therefore that is the threshold that will determine the success or failure of the CNN model.

The CNN modelling gave very promising results. The AUC validation score reached 100% after 6-7 Epochs. Some of the sweeps reached full score after 1 iteration, these had the optimum parameters for the models. According to **Figure 17** which shows various sweeps with different variations of the criteria. The ones that gave best results had learning rate of 0.00005 and 64 filters. This shows that the models perform well when learning at a slower rate, however this can be a downside as this would consume a lot of energy and require additional computation power. This is the same for the amount of filter, high filters worked better but processing the information through more filters require large number of calculations that mostly require a GPU due to the high memory requirement.

Another reason for the high scores achieved in this CNN compared to CNN done on other datasets, was due to CryptoPunks being deigned in a 24x24 pixel foundation. Therefore, during encoding it can be easy to identify as the encoding and normalisation would benefit the pixel art, unlike other data sets which need to be pixelized and rescaled vastly.

### 6.3 CNN – Classification with Attributes

Another model was scripted for development for this project, however due to personal commitment and lack of resources it could not be completed for this report. However, a brief, outline of the model is discussed below.

- o A prediction is made with the Male avatar with 0 accessories as the baseline
- o The CNN is required to achieve a higher score that the prediction for the model to be successful, but from overview the accuracy would have to reach 92+%., this would be difficult to achieve.
- o Similar processes would need to be undertaken using the AUC validation curve and sweeps varying which will provide the Wights and Biases visualisation
- o The results will highlight which features are predicted well and which need work compared to the baseline and can highlights problems occurring due to lack of data such as from the Ape or Aliens classes. These would be hard to identify as there is not enough data in the system or would require an excessively high number of iterations.

# 7. Conclusion

## 7.1 Report

This report has varied in success and failures but has definitely provided with a solid foundation for research into Crypto and NFTs. Initially the report highlights key definitions revolving around the crypto domain. This is given in detail and in a hierarchy order, so that links may be followed and understood. So that all bases are covered, and it is easy to understand the terminologies being used in the report the research diagram was used.

NFT was the primary task for this report, that has been covered extensively with few case studies. This report highlights that NFTs are non-fungible tokens, which are digital assets of data, that can be owed by individual on a decentralised block chain. This covers the benefits of NFTs and various opportunities and current applications of NFTs which have drawn huge crowds and businesses to add NFTs in promotions and security and contract making. However, all things have a flaw, this is also g=highlighted. It is shown that NFTs are prone to security risks, and no true ownership can be distinguished, and the intellectual property rights may cause a lot of confusion. Likewise, clients may be afraid in investing into NFTs as although these assets are on the Ethereum block (decentralised) the marketplace and the storage for these assets are in a centralised location or on centralised servers, these can be destroyed or deleted which would cost consumers a lot of money (millions of dollars even). Similarly, NFTs are causing problems to the environment due to huge power consumption and fuel shortages due to high demand for gas and high fuel prices.

After the report focuses on analytic work. This is achieved by fundament analysis of CryptoPunks dataset, which is one the first NFT art collection that has a random assortment of 10000 avatars in 5 different class types. The data was split into training- testing- and a further validation data set, which was then used to pass through the CNN model built for classifying the Punks. A CNN template was used to build the model, which was refined using hyperparameter optimisations the model, and the results were weighted and biased and with the help of a new tool the results were observed.

The results were very promising, with validation AUC achieved 100% score after only 6 iterations, whilst the graphs shows that slow learning rates and high use of filters will produce the best results.

Further research and time are required to develop a better understanding of the Convolutional Neural network. Research should be focused on including attributes to the classifications, which may be a complicated parameter to use. Likewise, researching into creating a more efficient model using different activation functions may prove to be beneficial in terms of less computational work and power consumptions

Furthermore, ideas of improving the research may lead toward creation of own unique CryptoPunks. There has been may research done on creation of NFTs, so those can provide a foundation of innovation.

## 7.2 Issues

No project is ever completed to what was initially predicted or expected. All decisions lead to a mixture of success and failures. This was the case wit this project, there are many places where success has been achieved as an overall consensus. However, the main drawback from this project was time management, time allocation and personal

commitments. The project was rushed and plans for the project had to be changed and adapted, this the reason for the project being more theory based than analytical and program orientated. Personal commitment to placements abroad without internet has vastly affected the research project, however some progress was still made to progress upon.

## 7.3 Future of NFTs

NFT is revolutionary, and the development is still in its youth. The versatility of NFTs retaining history and self-verification of purchases of digital data in a token. This enables clients to trade freely in an open marketplace. This functionality is giving advances and more opportunities into the crypto domain.

### 7.3.1 Online Events and Communities

First use of NFT was using the token as a 'membership pass' for a digital event. Popular trending NFTs are influenced by media and general population trends. Currently big NFT collections like CryptoPunk or Oni Ronin have maximised this concept for maximum profit. Clients minting or trading into such collections would be able to gain access into private seminars and events, free commodities and sponsored assets, and be entered into raffle draws.

In addition to providing access to online events and networking, NFTs are more favourably being used to provide exclusive entry to real-life events. With issues with fraudulent tickets and digital theft and scamming, the development of NFTs is thought to be the solution to these problems, as they are immutable, and ownership can be directly traced on the blockchain.[10]

### 7.3.2 Exchangeable Game Assets

Gaming has always carved a path to future ideas and development of modern technology. NFTs are now a huge boost to gaming industry, where clients gain ownership of in-game purchased assets. Metaverse has projects where they have 'NFT avatars,' where players can choose to trade the avatar or even rent it on the game market. The development of the games, the NFTs purchased can be used to obtain rare quests which earns further NFTs, game items or crypto currencies, which can be exchanged for crafting or upgrading the avatar.

Many game developers understand and accept that NFT games are the ultimate fate of gaming. There are various advantages that is provided from NFT games which customary games cannot provided. Some are listed as follows –

**Income**

NFTs can be seen as a symbol of hope for third-world countries that are underdeveloped, with weak economies and high unemployment rates. Play-to-earn games give a wonderful opportunity to the residents of these areas, where they can simply earn money by playing games. Axie Infinity is very popular in nations like the Philippines, where it has been recoded of average earnings for players to be around $300-$1200 monthly. [11]

**Ownership**

NFTs involved in games introduces bartering for any NFT or in-game items that happen to be an NFT. As the player retains ownership of the item, with their unique signature it is easy traceable to who owns the item and where the money should be transferred to.

Each NFT is extraordinary and can't be duplicated since they are situated on the blockchain. This gives games assurance that their hard-earned digital assets are always safe and cannot be stolen or copied from their accounts via the attack of malicious and fraudulent users. [10]

**Exchange Commission/Business**

Not only players, but game developers heavily benefit form NFT gaming. When an NFT item changes possession, the game designer may charge a fix rate where a certain percentage is always distributed to the maker (commission). This will profit the designer with every purchase of the NFT, especially beneficial when the NFT is sold for high values. This overall creates a business framework that is highly independent where both parties (consumer and developer) profit on the NFT market.

Game developers can attract sponsors and funding for the games through NFTs and using them as leverage. Great financial investors, with knowledge of crypto domain and NFTs, are eager to fund promising projects like Axie Infinity and other games on the metaverse. [12]

### 7.3.3 Digital Identities and Resources

Digital identities are important and vital for certain users. In the metaverse, through NFTs it is possible to obtain ownership of custom usernames. It has been reported that 600000+ usernames have been registered on a Ethereum Name Service with the Ethereum wallet. The NFT on the blockchain utilises this username and integrates it into a decentralised applications scattered on the network. 3Other NFT projects can give clients ownership of private domains as well.

Sandbox is a well renowned game developer; they have opened real estate in the metaverse. NFTs is used to own addresses of real estate. Famous celebrities have already taken interest and purchased various digital land.

The future is already in motion, with the shift of web 2.0 to 3.0, and NFTs play a huge role on this. They are constructing the establishment of upcoming digital networks, economies and resources. [18]

# References

1. Townsened, C., 2022. *Blockchain, Public Ledger, and Peer to Peer Sharing*. [online] United States Cybersecurity Magazine. Available at: <https://www.uscybersecurity.net/public-ledger-small-businesses/> [Accessed 26 August 2022]

2. Investopedia. 2022. *What Is Ethereum?*. [online] Available at: <https://www.investopedia.com/terms/e/ethereum.asp> [Accessed 26 August 2022].

3. Ieeexplore.ieee.org. 2022. *NFTs: Applications and Challenges*. [online] Available at: <https://ieeexplore.ieee.org/document/9677260/> [Accessed 26 August 2022].

4. MuleSoft Blog. 2022. Learn APIs | MuleSoft Blog. [online] Available at: <https://blogs.mulesoft.com/learn-apis/> [Accessed 26 August 2022].

5. Medium. 2022. Deep Learning w/ CryptoPunk. [online] Available at: <https://medium.com/analytics-vidhya/deep-learning-w-cryptopunks-6b27dc767fab> [Accessed 12 May 2022].

6. Investopedia. 2022. What Are Centralized Cryptocurrency Exchanges? [online] Available at: https://www.investopedia.com/tech/what-are-centralized-cryptocurrency-exchanges/ [Accessed 7 April 2022].

7. Investopedia. 2022. What is a Cryptocurrency Public Ledger? [online] Available at: <https://www.investopedia.com/tech/what-cryptocurrency-public-ledger/> [Accessed 7 April 2022].

8. Christidis, K. and Devetsikiotis, M., 2016. Blockchains and Smart Contracts for the Internet of Things. IEEE Access, 4, pp.2292-2303.

9. Zheng, Z., Xie, S., Dai, H., Chen, W., Chen, X., Weng, J. and Imran, M., 2020. An overview on smart contracts: Challenges, advances and platforms. Future Generation Computer Systems, 105, pp.475-491.

10. Mazur, M., 2021. Non-Fungible Tokens (NFT). The Analysis of Risk and Return. SSRN Electronic Journal,

11. MuleSoft Blog. 2022. What are APIs and how do APIs work? | MuleSoft Blog. [online] Available at: <https://blogs.mulesoft.com/learn-apis/api-led-connectivity/what-are-apishow-do-apiswork/#:~:text=API%20stands%20for%20%E2%80%9CApplication%20Programming,the%20response%20back%20to%20you.> [Accessed 12 May 2022].

17. Zzun.app. 2022. NFT-Price-Prediction-CNN - Using visual feature extraction, prices of NFTs are predicted via CNN (Alexnet and Resnet) architectures. | PythonRepo. [online] Available at: <https://zzun.app/repo/furkanoruc-NFT-Price-Prediction-CNN> [Accessed 12 May 2022].

18. Shahriar, S. and Hayawi, K., 2022. NFTGAN: Non-Fungible Token Art Generation Using

19. Generative Adversarial Networks. [online] arXiv.org. Available at:

20. Medium. 2022. Generate NFT CryptoPunks with Deep Convolutional Generative Adversarial Network (DCGAN). [online] Available at: <https://medium.com/mlearning-ai/generate-nft-cryptopunks-with-deep-convolutional-generative-adversarial-network-dcgan-db35f0a1adb4> [Accessed 26 August 2022].

21. Medium. 2022. Data Preprocessing and Network Building in CNN. [online] Available at: <https://towardsdatascience.com/data-preprocessing-and-network-building-in-cnn-15624ef3a28b> [Accessed 18 March 2022].

22. Analytics Vidhya. 2022. CNN Image Classification | Image Classification Using CNN. [online] Available at: <https://www.analyticsvidhya.com/blog/2020/02/learn-image-classification-cnn-convolutional-neural-networks-3-datasets/> [Accessed 18 March 2022].

23. TensorFlow. 2022. Transfer learning and fine-tuning | TensorFlow Core. [online] Available at: <https://www.tensorflow.org/tutorials/images/transfer_learning> [Accessed 18 March 2022].

24. GeeksforGeeks. 2022. CNN | Introduction to Pooling Layer - GeeksforGeeks. [online] Available at: <https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/> [Accessed 18 March 2022].

25. MLK - Machine Learning Knowledge. 2022. Keras Dense Layer Explained for Beginners - MLK - Machine Learning Knowledge. [online] Available at: <https://machinelearningknowledge.ai/keras-dense-layer-explained-for-beginners/> [Accessed 18 March 2022].
26. Fang, H., 2022. Applied Machine Learning. [online] 21COP528 - Applied Machine Learning. Available at: <https://learn.lboro.ac.uk/course/view.php?id=21420> [Accessed 16 March 2022].

# Appendix

| Manazir | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Data Science (MSc) | | | **NFT Timeplan: Non-fungible token categorisation using machine learning classifiers** | | | | | | | | | | | | | | | | | | | | | | |
| Loughborough University | | | Mar-22 | | | Apr-22 | | | | May-22 | | | | | Jun-22 | | | | Jul-22 | | | | Aug-22 | | |
| Week commencing: | 14th | 21th | 28th | 4th | 11th | 18th | 25th | 2nd | 9th | 16th | 23rd | 30th | 6th | 13th | 20th | 27th | 4th | 11th | 18th | 25th | 1st | 8th | 15th | 22nd | 29th |
| **Task** | Wk1 | Wk2 | Wk3 | Easter | | | | Wk4 | Wk5 | Wk6 | Wk7 | Wk8 | Wk9 | Wk10 | Wk11 | Wk12 | Wk13 | Wk14 | Wk15 | Wk16 | Wk17 | Wk18 | Wk19 | Wk20 | Wk21 |
| *Progression* | | | | | | | | | | | | | | | | | | | | | | | | | |
| Project Allocation | ☑ | | | | | | | | | | | | | | | | | | | | | | | | |
| *1.0 - Literature Review* | | | | | | | | | | | | | | | | | | | | | | | | | |
| General research | | | | | | | | | | | | | | | | | | | | | | | | | |
| Initial Research - History Centralised System | | | | | | | | | | | | | | | | | | | | | | | | | |
| Initial Research- Blockchain | | | | | | | | | | | | | | | | | | | | | | | | | |
| Initial Research- Smart Contract | | | | | | | | | | | | | | | | | | | | | | | | | |
| Initial Research- Etheruem | | | | | | | | | | | | | | | | | | | | | | | | | |
| Initial Research- Crypto Punk NFT | | | | | | | | | | | | | | | | | | | | | | | | | |
| Project Research Proposal | | | | | | | | ☑ | | | | | | | | | | | | | | | | | |
| *2.0 - Pre-Processing* | | | | | | | | | | | | | | | | | | | | | | | | | |
| Collecting Data using API | | | | | | | | | | | | | | | | | | | | | | | | | |
| Cleaning Up Data | | | | | | | | | | | | | | | | | | | | | | | | | |
| Appropriate Labeling | | | | | | | | | | | | | | | | | | | | | | | | | |
| Storing on SQL Database | | | | | | | | | | | | | | | | | | | | | | | | | |
| *3.0 - Data Analysis* | | | | | | | | | | | | | | | | | | | | | | | | | |
| Make Hypothesis | | | | | | | | | | | | | | | | | | | | | | | | | |
| Basic Analysis | | | | | | | | | | | | | | | | | | | | | | | | | |
| Expereiment with different ML Algorithmns | | | | | | | | | | | | | | | | | | | | | | | | | |
| Build Classifying Algorithm | | | | | | | | | | | | | | | | | | | | | | | | | |
| Build CNN | | | | | | | | | | | | | | | | | | | | | | | | | |
| Testing on Unseen data | | | | | | | | | | | | | | | | | | | | | | | | | |
| Optimisation | | | | | | | | | | | | | | | | | | | | | | | | | |
| Final Testing/ Prediction | | | | | | | | | | | | | | | | | | | | | | | | | |
| *4.0 - Further Work* | | | | | | | | | | | | | | | | | | | | | | | | | |
| Compare Etheureum with Bitcoin | | | | | | | | | | | | | | | | | | | | | | | | | |
| Predict Price of Etheruem | | | | | | | | | | | | | | | | | | | | | | | | | |
| Analsys Volume purchase of NFT | | | | | | | | | | | | | | | | | | | | | | | | | |
| *Disertation Write-up* | | | | | | | | | | | | | | | | | | | | | | | | ☑ | |

| | Key |
|---|---|
| ☑ | Hand-in |
| | Contingency |

**https://drive.google.com/drive/folders/1hu1OEChyx0H9LOF-h8qybwtTcJKzpz86?usp=sharing**