Springboard Capstone Project Proposal

Intermediate Data Science Course
Brian Leip
2/4/2018

Project:

CryptoCurrency Price Analysis Using Machine Learning and Neural Networks to generate buy/sell and irregular activity alerts, and potentially forecast future prices

Problem Being Solved:

Providing buy/sell alerts and irregular market activity alerts to human users that are limited in the number of cryptocurrencies that they can watch and analyze in a 24 / 7 / 365 market. Also there is a potential for future price forecasting which is extremely useful but equally difficult in any financial market.

Potential Client and Use Cases:

Anyone interested in investing in cryptocurrencies, or already invested and interested in alerts to take action (buy/sell) on their current investments.

Current Approach to Solving Problem (subject to change):

- Gather cryptocurrency time series price and volume data from various exchanges, apis, data providers
- Clean, wrangle and consolidate the data
- Perform price analyses and forecasting on the data
- Potential Tools:
    - Machine Learning, Deep Learning, Neural network analyses
    - Keras deep learning tools
    - LSTM recurrent network
    - More to come as I progress through the course and learn more
    - Data Viz Tools:
        - Matplotlib / Pyplot
        - Seaborn
        - More to come as I progress through the course and learn more
- Useful Resources:
    - blog about analyzing cryptocurrency prices in Python - https://blog.patricktriest.com/analyzing-cryptocurrencies-python/
    - youtube video from Siraj about predicting prices using Keras deep learning bidirectional LSTM recurrent network - https://www.youtube.com/watch?v=G5Mx7yYdEhE&t=22s
    - scholarly paper - PERFORMANCE FUNCTIONS AND REINFORCEMENT LEARNING FOR TRADING SYSTEMS AND PORTFOLIOS - https://pdfs.semanticscholar.org/170c/c2f8373322ab91036bbb66fc52b5c5c37e83.pdf?_ga=2.107245569.2093925112.1504293489-29180410.1504293489

- Cryptocompare API - https://www.cryptocompare.com/api/
- Bitcoin blockchain - https://blockchain.info/
- Ethereum blockchain - https://etherscan.io/

Deliverables:

1. Aggregated time-series price and volume data for multiple cryptocurrencies
2. Buy / Sell Alerts on cryptocurrencies as result of machine learning analysis
3. Irregular Activity Alerts on cryptocurrencies as result of machine learning analysis
4. Optional: Cryptocurrency price forecasting as result of neural network analysis

REFERENCE A - Capstone Proposal Instructions from Springboard

Project: Submit Your Capstone Project Proposal

🕐 1 Hour

Finalize one Capstone idea based on the feedback you got from your mentor and peers on your Section 1 submission, and on your newly acquired understanding of the tools and data wrangling. Submit a project proposal - a short (1-2 page) document that answers the following questions:

1. What is the problem you want to solve?
2. Who is your client and why do they care about this problem? In other words, what will your client DO or DECIDE based on your analysis that they wouldn't have otherwise?
3. What data are you going to use for this? How will you acquire this data?
4. In brief, outline your approach to solving this problem (knowing that this might change later).
5. What are your deliverables? Typically, this would include code, along with a paper and/or a slide deck.

The proposal will be part of a github repository for your project. All code and further documentation you write will be added to this repository.

Once your mentor has approved your proposal, please share the github repository URL on the community and ask for feedback.

REFERENCE B - My Original 3 Ideas for Capstone Project

Github link:

https://github.com/BrianLeip/Data-Science-Intensive-Springboard/blob/master/Capstone%20Project/Ideas%20for%20Capstone%20Project.md

CAPSTONE PROJECT IDEAS

Springboard - Data Science Intensive Course

Brian Leip

IDEA 1: Alt-Coin Opportunity Identification

CryptoCurrencies, originated by Bitcoin in 2009 have recently experienced one of the most explosive bull runs in financial history over the last year. A large number of alt-coins (the name given to Bitcoin competitors) have appeared on the scene, designed to fix problems of previous CryptoCurrencies, or solve a problem in the global financial markets or online payments system. Of the 1,000+ alt-coins, there are some with excellent potential, and others that are scams or copycats. Buyers need to differentiate the wheat from the chaff. Coins will often initially launch to the market on an ICO at relatively low cost, and based on supply/demand of the marketplace some coins will rise in value, often extremely quickly. XRP (Ripple) for example, rose 36,018% in 2017 to become the #2 coin by market cap after Bitcoin. $1,000 invested in XRP would have returned $360,018. NEM rose 29,842%, Ardor 16,809%, Stellar 14,441%, Dash 9,265%, Ethereum 9,162%, Litecoin 5,046%, and so on. Bitcoin itself rose 1,318% which is a phenomenal annual return compared to typical financial markets (equities, forex, fixed income, real estate, etc) but not enough to match the 13 alt-coins that surpassed it on a returns basis.

Being able to spot and identify alt-coins with good potential that are rising in favor could be an extremely valuable skill. Using publicly available data from CryptoCurrency exchanges (price time series data for multiple CryptoCurrencies across multiple exchanges, volume, trends, etc), the data can be analyzed to spot patterns and identify

the alt-coins with the most potential early on, thereby allowing the buyer to research the coin and capture the exponential returns as the rest of the market catches on.

Data Sources: [CryptoCompare API](#), [Bitcoin Blockchain](#), [Ethereum Blockchain](#), possibly others

---

IDEA 2: Bitcoin or Ethereum Blockchain Analysis

Blockchains have been compared to a decentralized database, and all transactions and history are publicly available for all to see and verify and analyze. There is a gold mine of data that could be analyzed to uncover useful information.

For example:

- With Bitcoin transaction fees rising to $25-$30, that high fee has rendered many accounts incapacitated. How many accounts are affected by this and is it material compared to the total supply?
- Is it possible to determine what is the active supply of Bitcoin / Ethereum vs amounts that have been lost or are being held long term.
- It is said that a large amount of coins are held by a relatively small number of people. What does that distribution look like?
- Mining has become more and more centralized in China. How has that concentration affected the functioning of the network?
- There are more transactions waiting in the Mempool than either Bitcoin or Ethereum network can process. What does that data flow look like and are there any clues in solving the capacity problem?

Data Sources: [CryptoCompare API](#), [Bitcoin Blockchain](#), [Ethereum Blockchain](#), possibly others

---

IDEA 3: Use Machine Learning to "play" retro video games (Atari, NES, etc) and get high scores or finish levels and possibly surpass human ability

Artificial Intelligence, Machine Learning, and Neural Networks are evolving and progressing rapidly, with many experts predicting that they will replace up to 50% of human jobs within our lifetimes. Understanding these powerful tools and having the ability to design and improve them are keys to relevance in a future driven by AI.

Video games are a perfect sandbox for Artificial Intelligence since there are clearly defined rules, it is easy to quantify and measure progress, and goals/wins/losses are often black and white.

In this project, I would use AI, Machine Learning and any tools under that umbrella such as Neural Networks in order to learn how to play video games, improve over time, and eventually surpass human ability.

Data Sources: [Open AI Gym - Atari](), [Starcraft II: API, Dataset, PySC2 toolset](), possibly others