

UMD DATA605 - Big Data Systems

Sorrentum Tutorial

Dr. GP Saggese

gsaggese@umd.edu

Overview

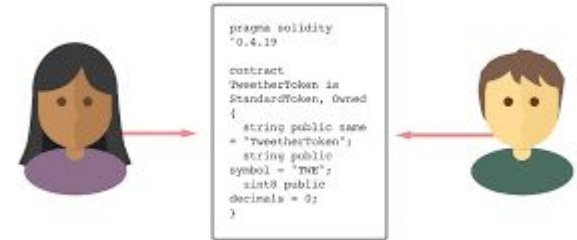
- www.sorrentum.org
 - DeFi protocol to build advanced financial applications
 - Open-source project backed by a non-profit foundation
- [GitHub](#) repo
 - Please watch, star, and fork the repo!
- **Two products / companies built on top of Sorrentum**
 - Kaizen (crypto high-frequency trading firm)
 - DaoSwap
- **Research project to foster entrepreneurship in DeFi (jointly with UMD)**
 - We have built a toolkit to quickly build Web3 DeFi apps
 - We give students a product idea
 - Students implement a prototype with our help
 - We fundraise for the prototype
 - Create a company out of it
 - Governance token as a start-up commune
- **How to contribute**
 - DATA605 class projects are part of Sorrentum
 - You can keep contributing to Sorrentum
 - Office hours for 5 research team every week (Arbitrage, Growth hacking, ML Model, NLP, Web3)
 - 62 students interested, 20 already contributing
 - Sign up [here](#) if interested



Sorrentum Primitives

- Smart contracts

- Staking, escrow, vesting
- P2P transactions
- Off-chain secure computation
- DaoSwap, DaoETF



- Data pipeline

- Standardized flow to on-board, clean, normalize, serve market data and alternative data

- Data flow

- Machine learning framework for describing, simulating, and deploying financial models

- Risk models

- Statistical / fundamental models
- Portfolio optimization

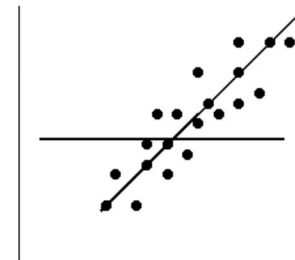
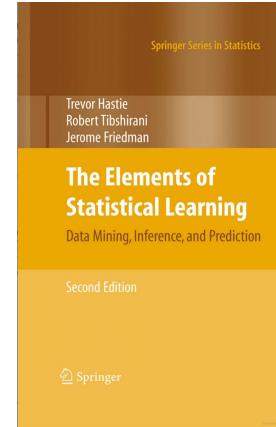
- Market execution

- Connect to centralized and decentralized exchanges
- Manage orders and portfolios using algorithmic trading

Sorrentum: Financial Machine Learning

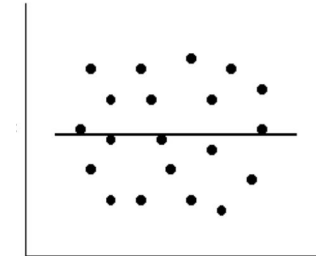
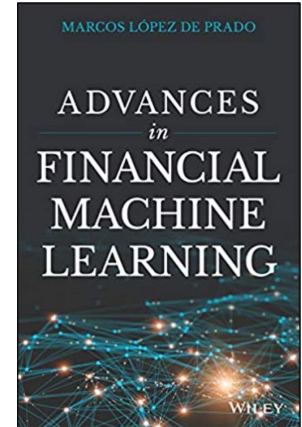
- Financial ML is very different than “traditional” ML
- Problems
 - Non stationarity
 - Fat-tails (vs Gaussianity)
 - Small data sets
 - Enormous amount of noise
 - No way to run experiments (vs A/B testing)
 - One single realization of the universe
 - Decisions are made collectively by millions of individuals
 - Non-completely rationally
 - Behavioral biases and fallacies
 - Markets are influenced by emotions, fear, and greed
- Solutions
 - Use scientific method based on hypothesis testing
 - Rigorous systematic research flow
 - Use “simple” models to deal with noise
 - E.g., linear model with regularization
 - Make a lot of bets + compounding

Traditional ML



Hit rate: 80%
 R^2 : 70%

Financial ML



Hit rate: 51%
 R^2 : 0.1% (10bps)

“Imagine how much harder physics would be if electrons had feelings!”,
Prof. Andrew Lo (MIT)

Blockchain

- **Replicated, shared, synchronized data based on consensus**

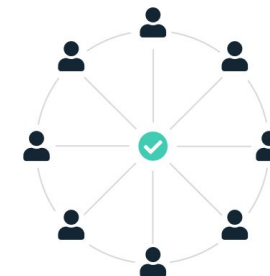
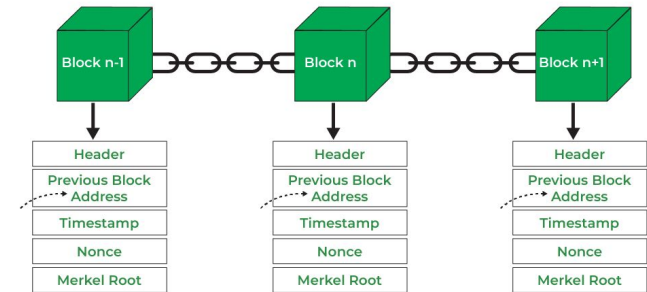
- Multiple parties can operate on data without trusting each other or a central authority
- Permission-less (public) vs permissioned (private)
- Intuition:
 - a distributed database
 - a distributed append-only immutable ledger

- **Consensus algorithm**

- Multiple peer nodes replicate and update a copy of the data independently
- Rules to determine what blocks become part of the blockchain
 - Proof-of-work vs proof-of-stake
- Lack of central authority
 - No single point of failure
- Designed to be resistant to malicious tampering
 - Up to a certain bound (e.g., 51% attack)

- **Example of blockchain + consensus algorithms**

- Bitcoin, Ethereum, Solana, Cardano



DECENTRALIZED CONSENSUS



CENTRALIZED CONSENSUS

Smart Contract

- Smart Contract / Protocol

- A computer program that automatically executes and controls actions according to the terms of a contract

- What to trust? What is trust?

- Based on law (e.g., US currency is legal tender, fiat currency)
- Based on social costumes (e.g., when you are in line at post office)
- Based on math / code
 - Trust is in the infrastructure
 - Code is open source
 - Code and data is immutable (i.e., blockchain)
 - Trust that it's unlikely to hijack 51% of the nodes
- Example:
 - A vending machine (can you trust it?)
 - An escrow company (can you trust it?)
 - Bitcoin protocol, Ethereum smart contracts (you can trust it)

- Nodes maintain / update state on a blockchain

- Reduce need for trusted intermediaries
- Reduce malicious and accidental exceptions
- Bitcoin a special-purpose blockchain / smart contract
 - First application was money
- Ethereum a blockchain for general-purpose smart contracts
 - Turing-complete
 - EVM

DApps and DAO

- Decentralized Applications

- Aka DApp, dApp, Dapp, dapp
- Operate autonomously (without human intervention) using smart contracts running on a blockchain
- Are not owned by any entity
- Use tokens to represent ownership and incentivize behaviors
 - Aka tokenomics or cryptonomics
- Typically open-source
 - You can do an (hostile) fork (in the old meaning of create a competing project)

- Decentralized Autonomous Organization

- DAO
- Company with rules encoded in a contract
- Transparent
- Controlled by the organization's members through governance token
- No centralized leadership
- Not influenced by central government
- E.g., Uniswap

Current Financial System is Failing

- **Our financial system has not substantially changed over the past century**
 - Banks, brokers, exchanges, insurance companies, central monetary authorities
- **One revolution was digitization**
 - On-line banking
- In 2023 (55 years after the man was on the moon)
 - 3% for a credit card swipe to an oligopoly of companies
 - Saving rates are 0 or negative
 - A wire transfer takes 2-3 business days
 - Banks are closed on weekends
 - Stocks are settled T+2 (transfer of ownership)
 - The government decides what you can or cannot do with your own money (e.g., accredited investors)
 - The government can fund wars with your tax money
 - Inflation is rampant around the world
 - In 2008 Bailouts of the very institutions that caused the crisis
 - There are 1.7b people in the world that have no access to banking system



Fight Club, 1999

Tyler Durden's Project
Mayhem: erase debt by blowing
up buildings where credit
information is stored

Decentralized Finance

- **Idea: instead of renovating a crumbling financial system, rebuild it from the ground up**
 - Offer financial instruments without relying on intermediaries (e.g., brokerage, exchanges, banks)
 - Replace old crap with dApps and smart contracts running on blockchain
- **The bright side**
 - No layers of bureaucracy
 - No middle-man taking commissions
 - No delay
 - Interoperability between services
 - Transparency
 - Equality (no special treatments or bribes)
- **The dark side**
 - Scams, greed, stupidity, rabbid speculation -> bubbles + bankruptcies
 - Do not learn from the past and make the same mistakes
 - As humans it seems that we can't help ourselves
- **My take: the impact of Web3 / Defi will be as big as the Internet**
 - From free information (Internet) to free society (Web3)
 - Put society and democracy on the blockchain
 - Get rid of rent-seeking tech companies
 - You can't defeat economy as you can't defeat gravity

Sorrentum: Projects

- Arbitrage

- = take advantage of discrepancy of prices in time or space to make a profit
- Key to a well-functioning markets
 - Capitalism at its best
- Mechanical (or risk-less) vs statistical arbitrage
- Cross-exchange
- Single-exchange

- Natural Language Processing (NLP) and markets

- Understand effect of news and social media on markets

- Model research

- Predict noisy financial quantities with ML
- E.g., price, volatility, volume, bid-ask spread

Sorrentum: Projects

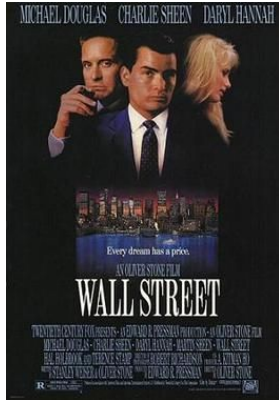
- DeFi / Web3 Research

- DApps
- DaoSwap
 - Lit liquidity pool
 - Crossing network
- DaoETF
 - Build investment tokens
 - Passive and active strategies
- Tokens
 - SORRE - Utility token
 - NTUM - Governance token

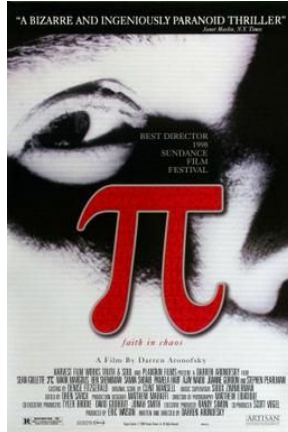
- Growth hacking

- How to build a community around these research project
- How to fund raise for DeFi ideas

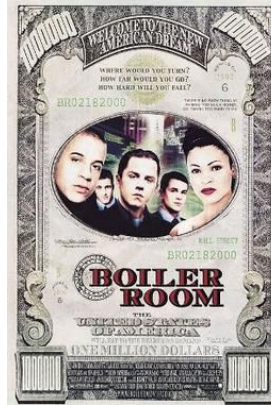
Additional Resources



[Wall Street, 1987](#)



[Pi, 1998](#)



[Boiler room, 2000](#)



[Margin call, 2011](#)



[The Wolf of Wall Street, 2013](#)



[The Big Short, 2015](#)

Sorrentum Sandbox

- [Tutorial](#)
- Two data node [examples](#)
 - Binance
 - Reddit