Startup.ML

QuantCon 2016

Machine Learning Based Cryptocurrency Trading

Outline

Exchange tradable cryptocurrency instruments

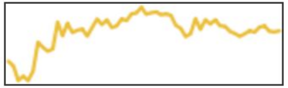
Exchange mechanics and fees

Architecture of an ML-based systematic strategy

Machine learning choices

Fellowship process to develop systematic strategies

Cryptocurrency Flavors and MCaps

#	Name	Market Cap	Price	Available Supply	Volume (24h)	% Change (24h)	Price Graph (7d)
1	Bitcoin	$ 6,477,017,277	$ 420.91	15,388,275 BTC	$ 32,935,500	-0.00 %	
2	Ethereum	$ 914,022,103	$ 11.61	78,716,970 ETH	$ 6,271,040	1.07 %	
3	Ripple	$ 259,153,481	$ 0.007525	34,439,870,367 XRP *	$ 334,116	0.95 %	
4	Litecoin	$ 147,164,496	$ 3.26	45,195,026 LTC	$ 778,374	-0.11 %	
5	Dash	$ 42,470,529	$ 6.70	6,340,721 DASH	$ 229,393	-6.62 %	

Available in ~30 Quote Currencies

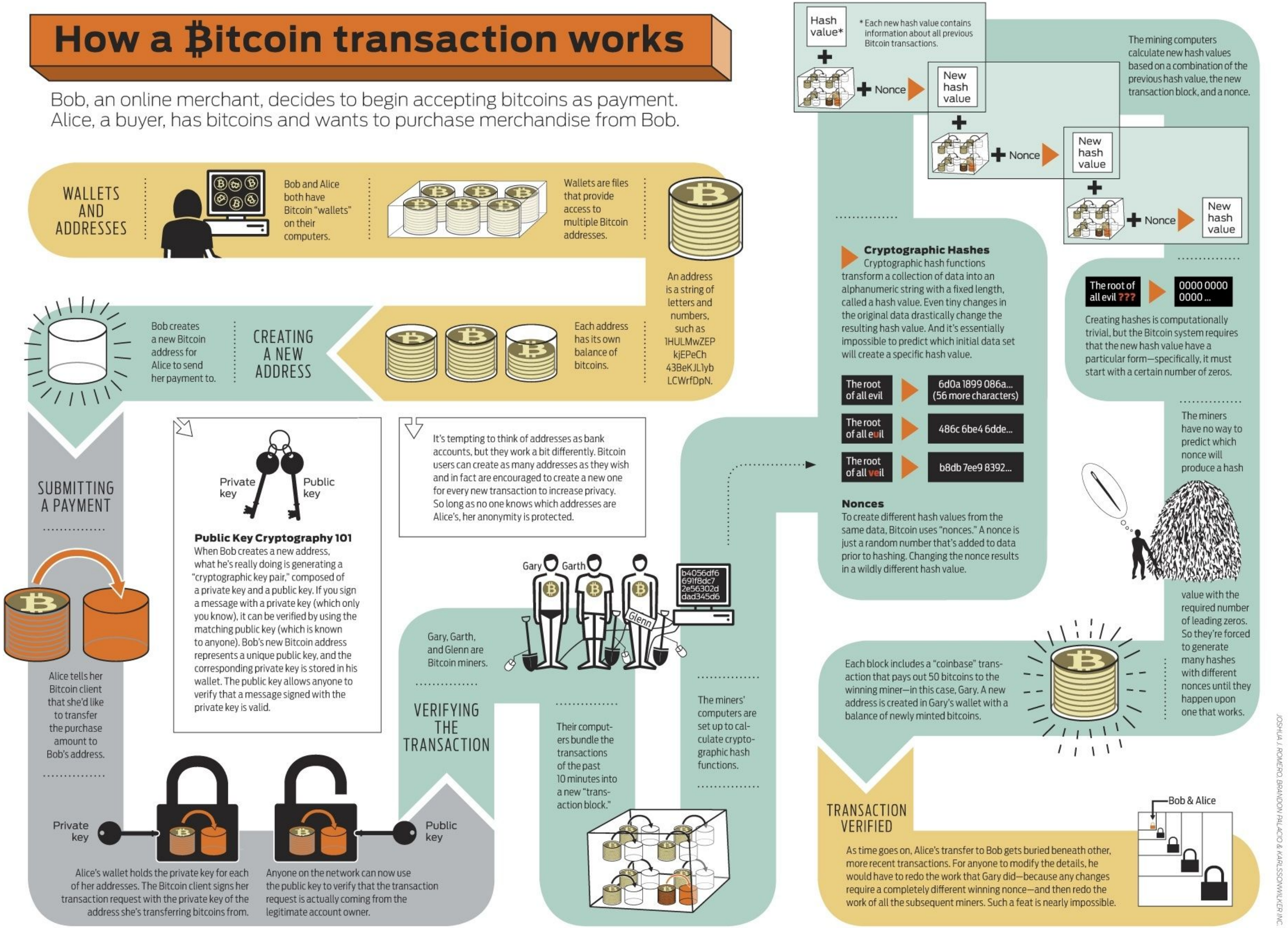
Overview | **Currencies** | **All Markets**

All	KRW	NMC	IDR	RON	ARS	AUD	BGN	BRL	BTC	CAD	CHF	CLP	CNY	CZK	DKK	EUR	GAU
GBP	HKD	HUF	ILS	INR	JPY	LTC	MXN	NOK	NZD	PEN	PLN	RUB	SAR	SEK	SGD	SLL	THB
UAH	**USD**	XRP	ZAR														

Symbol	Latest Price	30 days	Average	Volume	Low/High	Bid	Ask	24h Avg.	Volume	Low/High
▼ Bitfinex USD bitfinexUSD	419.2 2 min ago		412.56 6.64 1.61%	371,878.86 153,423,053.70 USD	389 431	419.2	419.25	419.75 -0.55 -0.13%	2,194.16 920,996.78 USD	418.6 420.79
▼ Coinbase USD coinbaseUSD	419.48 0 min ago		412.62 6.86 1.66%	195,723.11 80,759,360.46 USD	381.09 426.5	419.49	419.55	419.64 -0.16 -0.04%	3,296.23 1,383,240.21 USD	418.51 420.38
▼ BitStamp USD bitstampUSD	417.87 3 min ago		412.27 5.60 1.36%	142,419.21 58,714,707.38 USD	382 427.99	417.87	417.9	418.18 -0.31 -0.07%	1,602.70 670,211.40 USD	416.1 419.5
▲ btc·e USD btceUSD	416.051 0 min ago		413.68 2.37 0.57%	134,150.32 55,495,922.62 USD	393.996 427	415.981	416.192	415.66 0.39 0.10%	2,886.36 1,199,733.16 USD	413.101 416.8
▼ itBit USD itbitUSD	416.41 26 min ago		412.90 3.51 0.85%	81,733.31 33,747,462.50 USD	386 426.33	417.19	417.68	417.88 -1.47 -0.35%	2,960.94 1,237,327.12 USD	416 419.73
▲ LocalBitcoins USD localbtcUSD	474.96 3 min ago		478.09 -3.13 -0.66%	61,061.86 29,193,190.48 USD	271.53 4373.18	41504160417	410.35	471.08 3.88 0.82%	1,321.07 622,324.28 USD	353.86 2099.08
▼ Kraken USD krakenUSD	420.079 1 min ago		412.87 7.21 1.75%	49,662.06 20,503,742.92 USD	387.27 429.999	419.33	420.079	420.49 -0.41 -0.10%	525.75 221,073.11 USD	418.61 422
▼ The Rock Trading Company USD rockUSD	419.34 3 hrs, 45 min ago		415.32 4.02 0.97%	1,536.24 638,026.48 USD	390 429.77	419.01	419.64	420.65 -1.31 -0.31%	9.80 4,122.76 USD	419.34 421.69

How a Bitcoin transaction works

Bob, an online merchant, decides to begin accepting bitcoins as payment.
Alice, a buyer, has bitcoins and wants to purchase merchandise from Bob.



Modern Exchanges Being Built

Real-time order book (level 3)

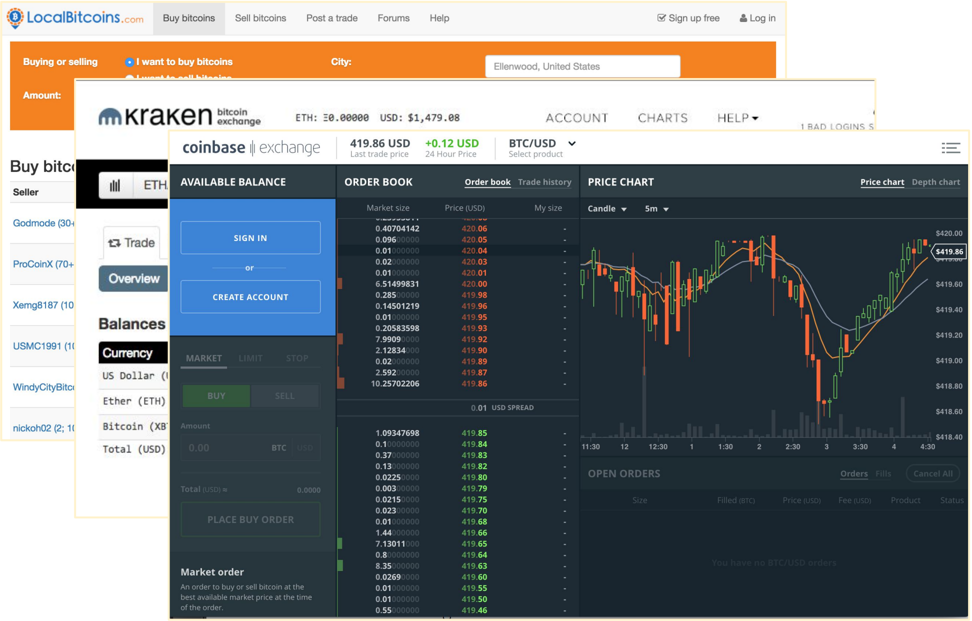
- Websocket Feed
- Colocation (aws region)

REST end-points orders, account, fills

FIX 4.2 support

Reasonable rate limit (3 req. / sec)

Python / Ruby sample code



Fee Structures Encourage Market Making

coinbase

USER 30 DAY VOLUME		TAKER FEE
> 0 %	(~ 0.00 BTC)	0.25 %
> 1 %	(~ 1,925.87 BTC)	0.24 %
> 2.5 %	(~ 4,814.68 BTC)	0.22 %
> 5 %	(~ 9,629.36 BTC)	0.19 %
> 10 %	(~ 19,258.73 BTC)	0.15 %
> 20 %	(~ 38,517.45 BTC)	0.10 %

Note: *Taker trades are charged 0.25% fee at the time of trade but a rebate for the previous 24 hours of trading fees will be*

kraken

Fee Schedule

Maker	Taker	Volume
0.16%	0.26%	< 50,000
0.14%	0.24%	< 100,000
0.12%	0.22%	< 250,000
0.10%	0.20%	< 500,000
0.08%	0.18%	< 1,000,000
0.06%	0.16%	< 2,500,000
0.04%	0.14%	< 5,000,000
0.02%	0.12%	< 10,000,000
0.00%	0.10%	> 10,000,000

Market Making Introduces Execution Risks

Limit orders are harder to time

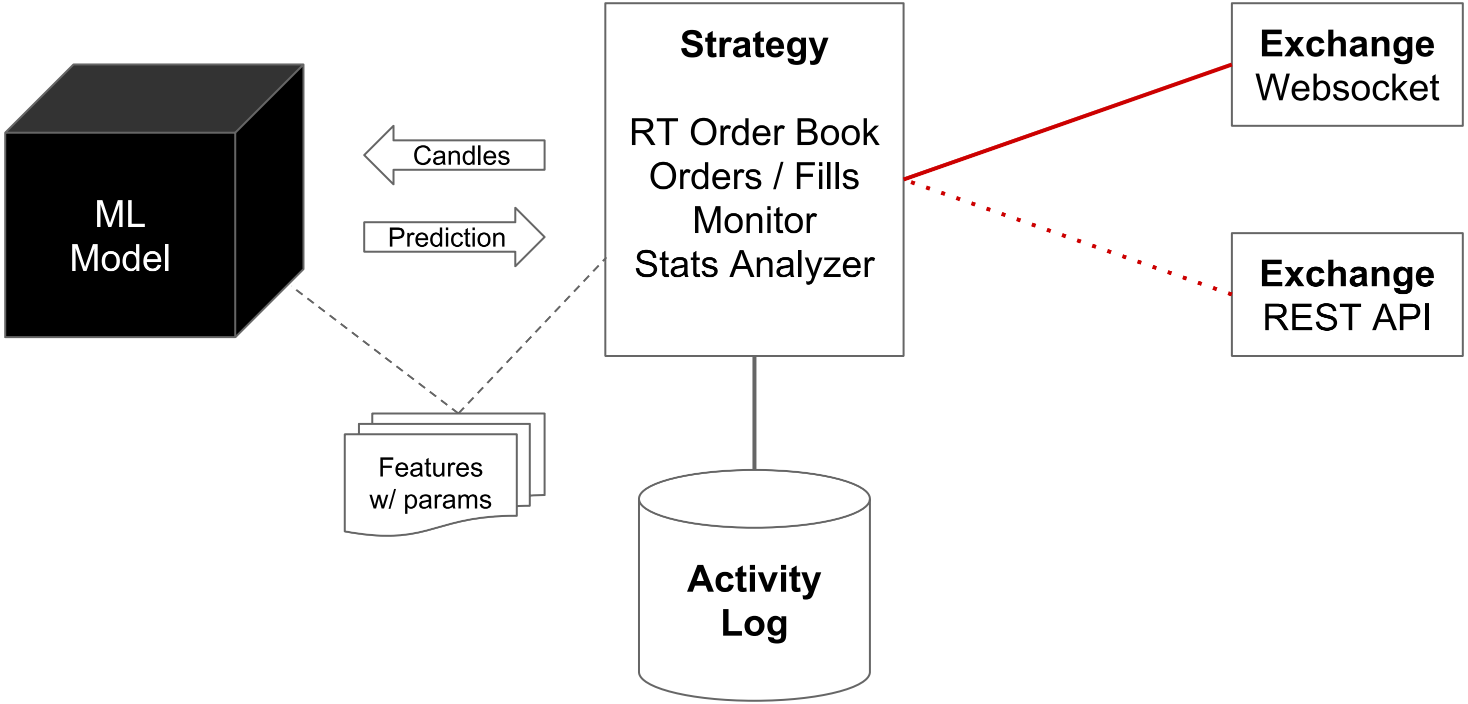
When the prediction is right, execution is particularly challenging

Introduces more parameters that need to be optimized

Taker strategies are easier but may require leverage

- Bitfinex (3.3x), Kraken (5x)
- Beware of clawbacks and exchange risks

Deployment Architecture



ML-Based Strategy Considerations

Input Features

Price differences

Lagged windows

Technical indicators

Order book pressure

Prediction Types

Point regression

Slope regression

Binary classification

Multi-class classification

Parameters

Algorithm params

Candle size

Buy / Sell threshold(s)

Stop loss

Order wait (attempts/time)

Order commit threshold

Max hold

Bet size

Allocation

Picking the Best Modeling Approach

Representation

- instances
  - k-nearest neighbor
  - support vector machine
- hyperplanes
  - naïve bayes
  - logistic regression
- decision trees
- set of rules
  - propositional rules
  - logic programs
- neural networks
- graphical models
  - bayesian networks
  - conditional random fields

Evaluation

- accuracy / error rate
- precision and recall
- squared error
- likelihood
- posterior probability
- information gain
- k-l divergence
- cost/utility
- margin

Optimization

- combinatorial optimization
  - greedy search
  - beam search
  - branch-and-bound
- continuous optimization
  - unconstrained
    - gradient descent
    - conjugate gradient
    - quasi-newton method
- constrained
  - linear programming
  - quadratic programming

ML-Based Strategy Considerations

Input Features

Price differences

Lagged windows

Technical indicators

Order book pressure

Prediction Types

Point regression

Slope regression

Binary classification

Multi-class classification

Parameters

Algorithm params

Candle size

Buy / Sell threshold(s)

Stop loss

Order wait (attempts/time)

Order commit threshold

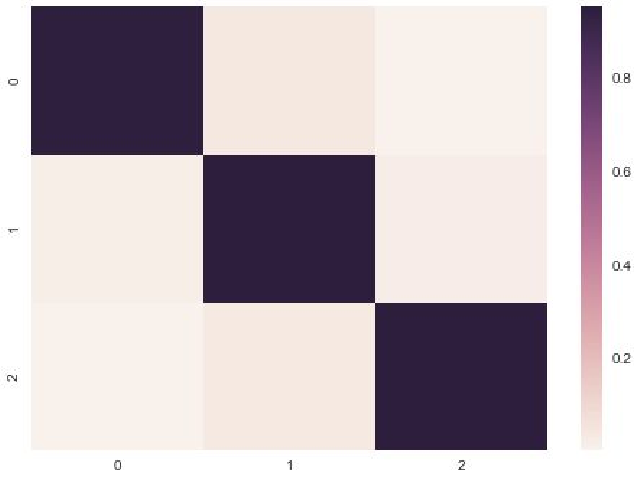
Max hold

Bet size

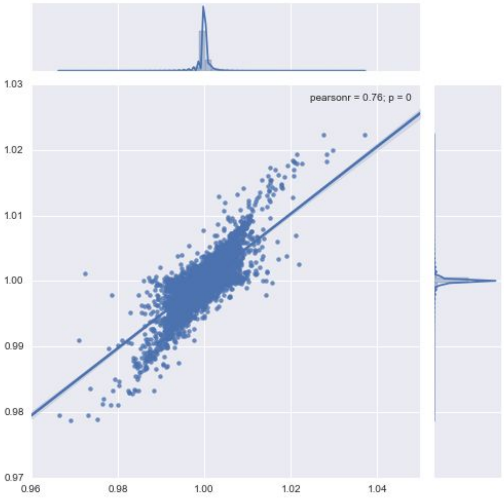
Allocation

Evaluating Model Fit

AUC



R^2



ML-Based Strategy Considerations

Input Features

Price differences

Lagged windows

Technical indicators

Order book pressure

Prediction Types

Point regression

Slope regression

Binary classification

Multi-class classification

Parameters

Algorithm params

Candle size

Long/short threshold(s)

Stop loss

Order wait (attempts/time)

Order commit threshold

Max hold

Bet size

Allocation

Strategy Optimization

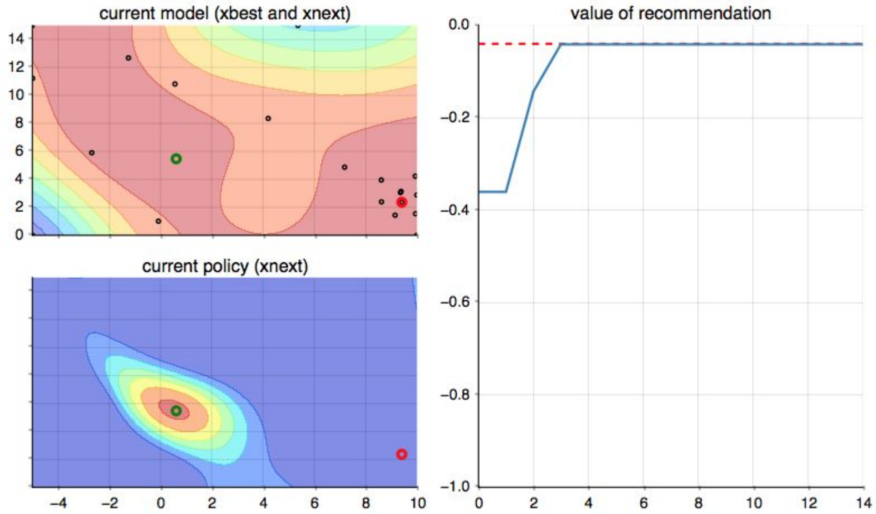
Dozens of parameters

Many difficult to simulate

Hours / days to optimize

Significantly more dynamic than prediction

Next set of machine learning challenges!



Temporal Strategies, Continuous Process

Real-world machine learning applications

Immersive 4 month program

Program almost exclusively hands-on practice

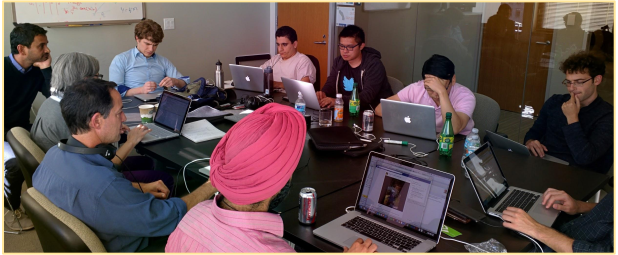
Agile software development methodology

Mentoring by experienced data scientists

1 mentor : 4 fellows ratio

Acceptance rate < 3%

Free to the fellows



Experienced Mentor Team

ARGYLE DATA

ENDECA®

Ucla

A9™

APT Test & Learn®

THE UNIVERSITY OF CALIFORNIA · BERKELEY · 1868 · LET THERE BE LIGHT

coinbase

yelp

RICE Unconventional Wisdom

Fellows Doing Meaningful Data Science Work

UBER

sentient

enlitic

yelp

orange™

Pivotal

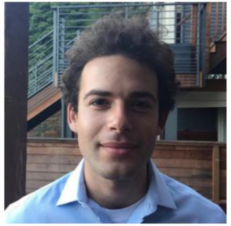

Alex Chao, Uber ATC


Rewon Child, Enlitic


Angie Hsieh


George Manuelpillai, Orange


Philip Margolis


Alex Miller, Yelp


Yevhen Mohylevskyy


Saad Eddin Al Orjany


Shonket Ray


Daniel Saltiel, Engineers Gate


Marjorie Sayer, CloudGenix


Peter Skipper, Sentient Technologies


Layla Tadjpour


Vrushank Vora


Eric Wayman, Pivotal


arshak@startup.ml / conf.startup.ml
May 12 @ Bloomberg West Coast Tech Hub