Beating the Stock Market

(in 4 hours)

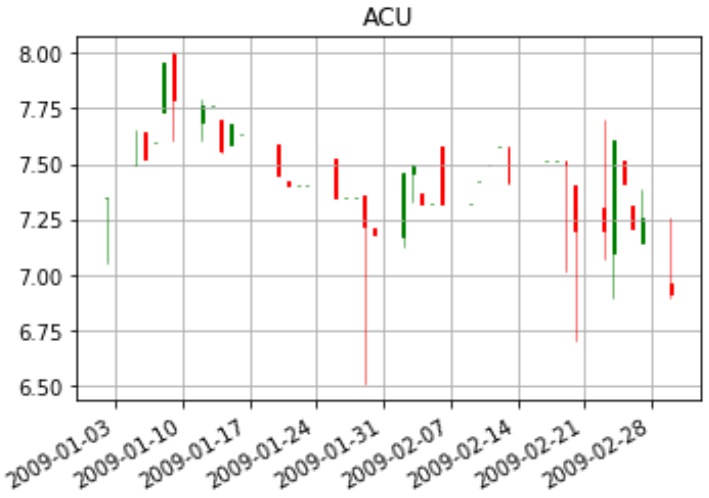
Konstantin Tretyakov

November 19, 2017

Get me a portfolio of stocks that'll take **$100,000** to **$1M** in 3 months

Data

1. Stock prices



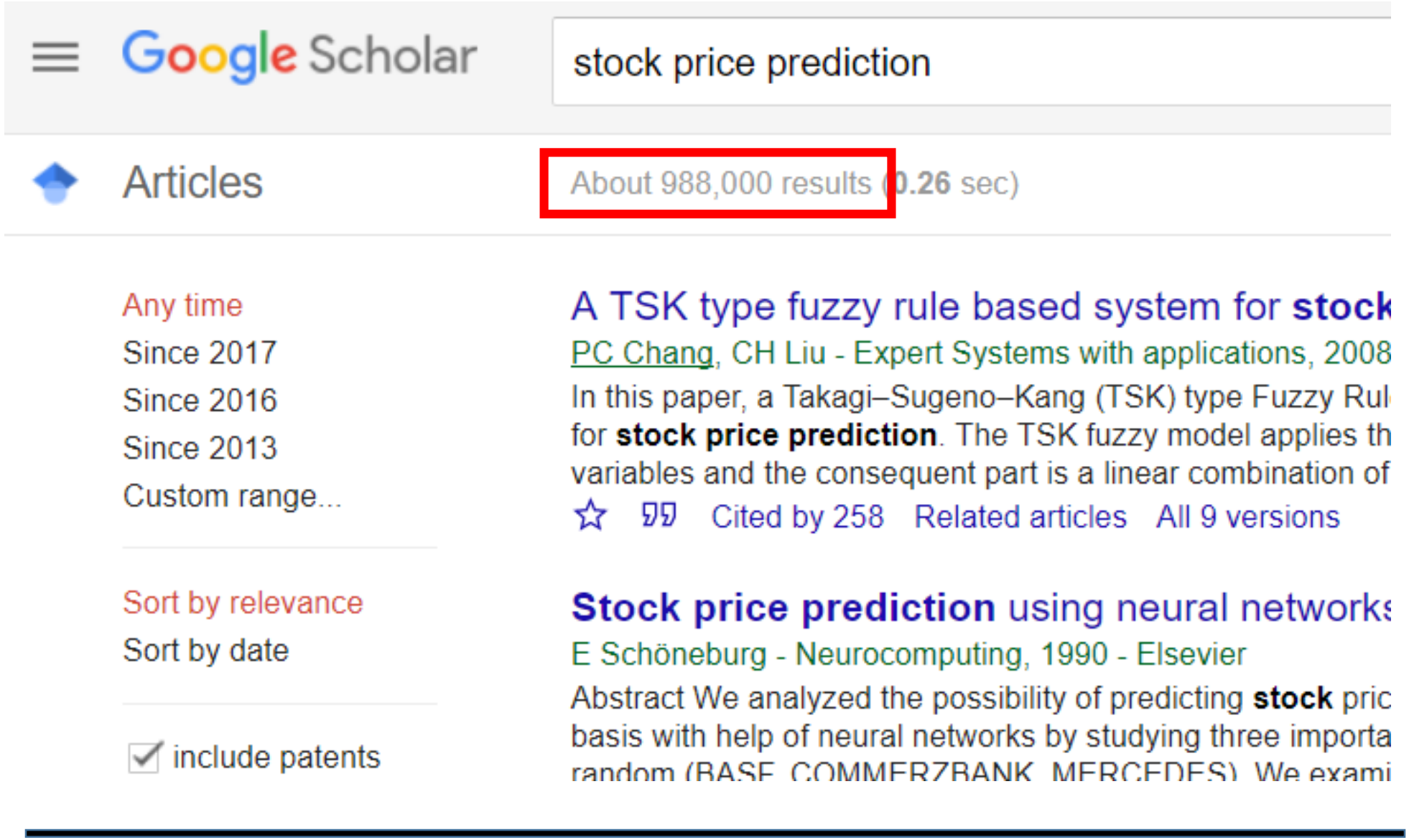
≈7000 stocks over ≈8 years: 11M records

2. Fin. reports

symbol	ACU
period_focus	Q3
end_date	2011-09-30
doc_type	10-Q
revenues	19,036,000.00
op_income	1,120,000.00
net_income	682,000.00
eps_basic	0.22

≈90 000 records

TEXATA™ Big Data Analytics Showdown


Google Scholar
stock price prediction
Articles
About 988,000 results (0.26 sec)
Any time
Since 2017
Since 2016
Since 2013
Custom range...
Sort by relevance
Sort by date
include patents
A TSK type fuzzy rule based system for stock
PC Chang, CH Liu - Expert Systems with applications, 2008
In this paper, a Takagi–Sugeno–Kang (TSK) type Fuzzy Rul
for stock price prediction. The TSK fuzzy model applies th
variables and the consequent part is a linear combination of
Cited by 258 Related articles All 9 versions
Stock price prediction using neural networks
E Schöneburg - Neurocomputing, 1990 - Elsevier
Abstract We analyzed the possibility of predicting stock pric
basis with help of neural networks by studying three importa
random (BASF, COMMERZBANK, MERCEDES). We exami

Google Scholar

stock price prediction

Articles

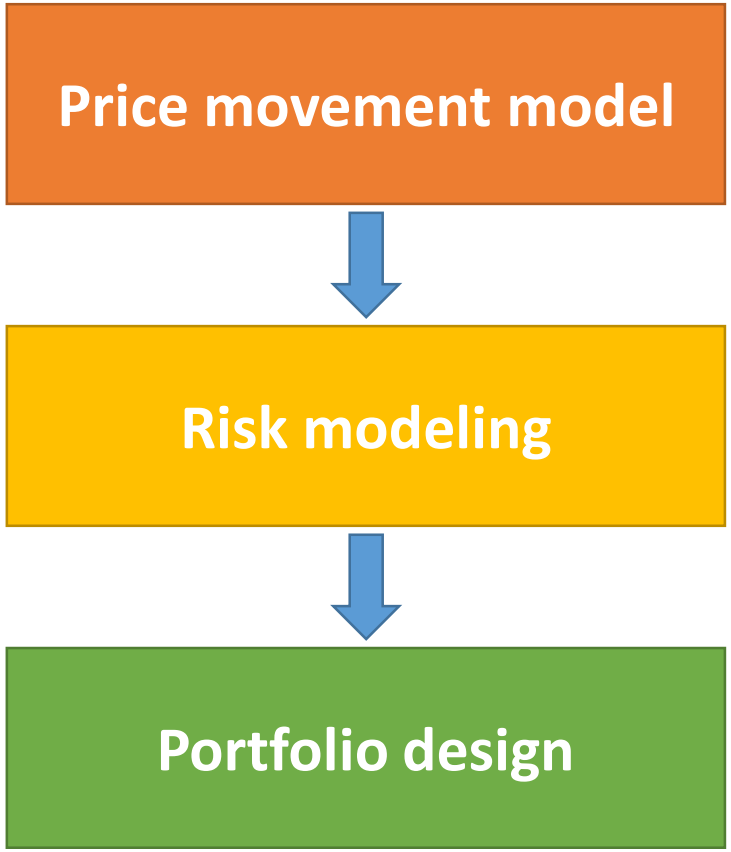
About 988,000 results (0.26 sec)

Any time
Since 2017
Since 2016
Since 2013
Custom range...

Sort by relevance
Sort by date

include patents

The Texata solution

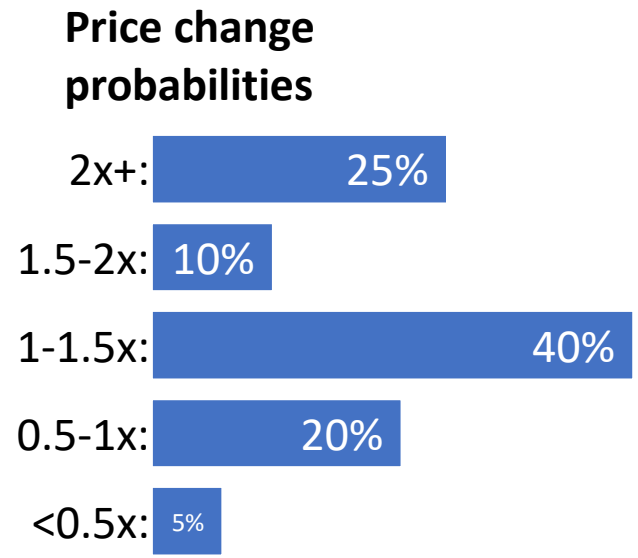


The Texata solution

Price movement model

AAPL

Price movement model



Today

Next month

AAPL

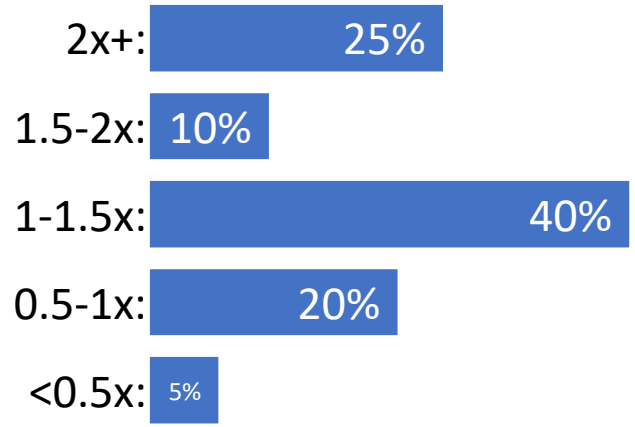
N most recent quarterly report indicators

Stock price over last K weeks

Static company description

→ **Price movement model** →

Price change probabilities



Today

Next month

AAPL

N most recent quarterly report indicators

Stock price over last K weeks

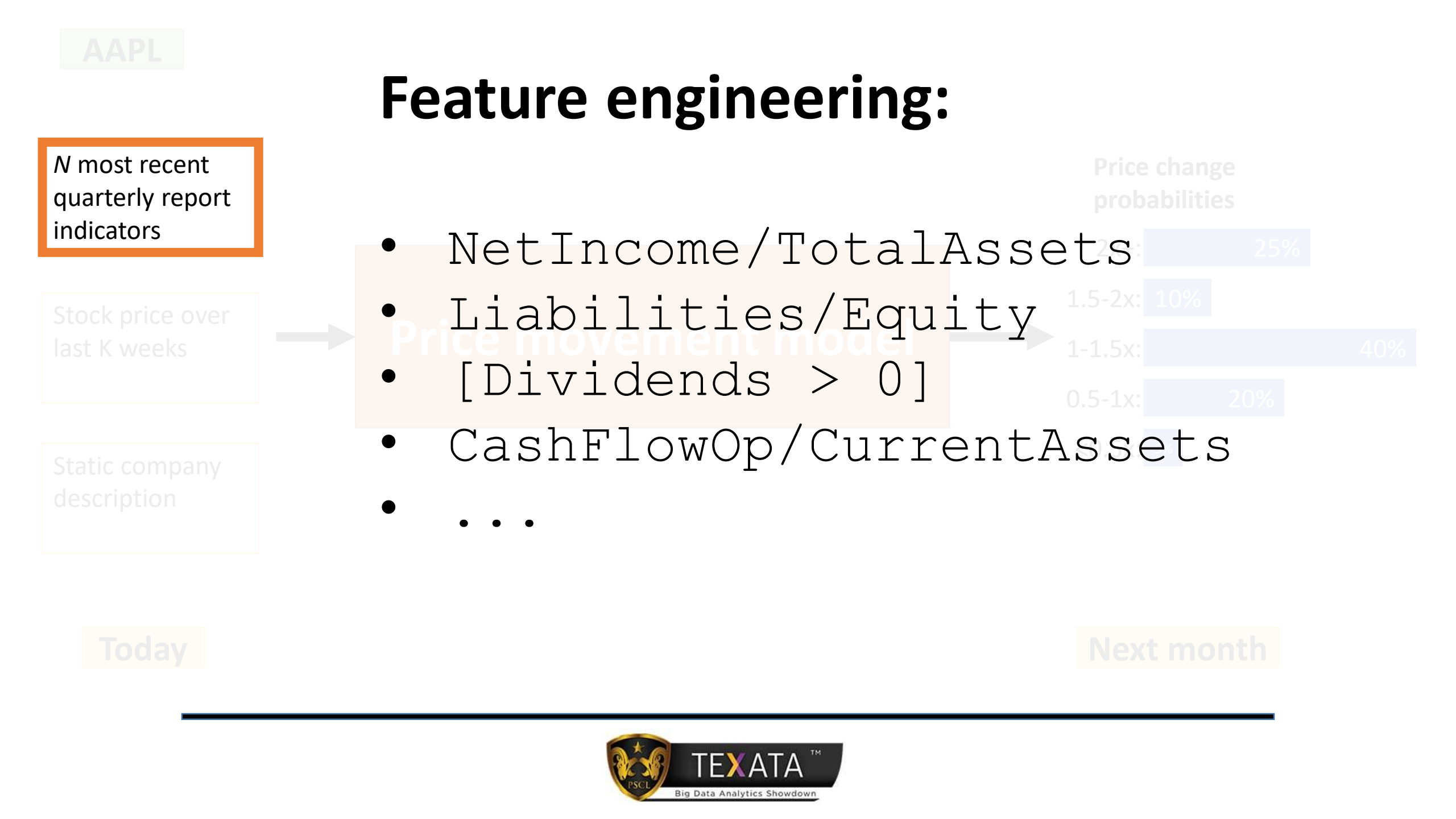
Static company description

Price movement model

Price change probabilities

Today

Next month



Feature engineering:

- NetIncome/TotalAssets
- Liabilities/Equity
- [Dividends > 0]
- CashFlowOp/CurrentAssets
- ...

Feature engineering:

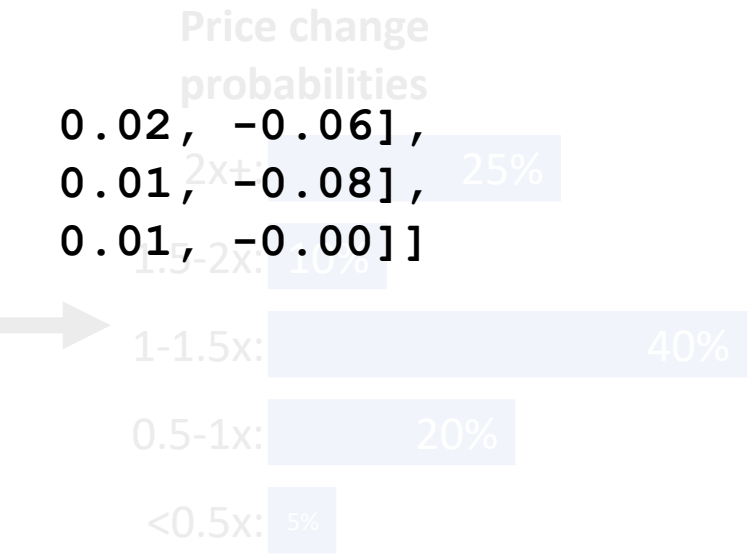
N most recent quarterly report indicators

Stock price over last K weeks

Static company description

```
[[ 0.15,  0.02,  0.06,  0.14, 0.05,  0.00,  0.02, -0.06],
 [ 0.12,  0.01,  0.17,  0.11, 0.16,  0.00,  0.01, -0.08],
 [ 0.26,  0.00,  0.24,  0.27, 0.25,  0.00,  0.01, -0.00]]
```

Price movement model



AAPL

N most recent quarterly report indicators

Stock price over last K weeks

Static company description

```
[[ 3.42,  3.47,  3.38,  3.44],
 [ 3.42,  3.49,  3.40,  3.47],
 [ 3.51,  3.51,  3.41,  3.46],
 [ 3.30,  3.34,  3.27,  3.31]]
```

Price movement model

Price change probabilities

Today

Next month

AAPL

N most recent quarterly report indicators

Stock price over last K weeks

Static company description

Price movement model

Name,
Field of Business,
Location,
…

Price change probabilities

Today

Next month

AAPL

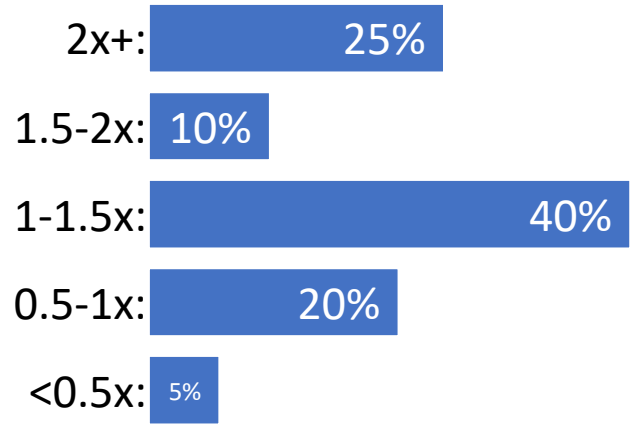
N most recent quarterly report indicators

Stock price over last K weeks

Static company description

→ **Price movement model** →

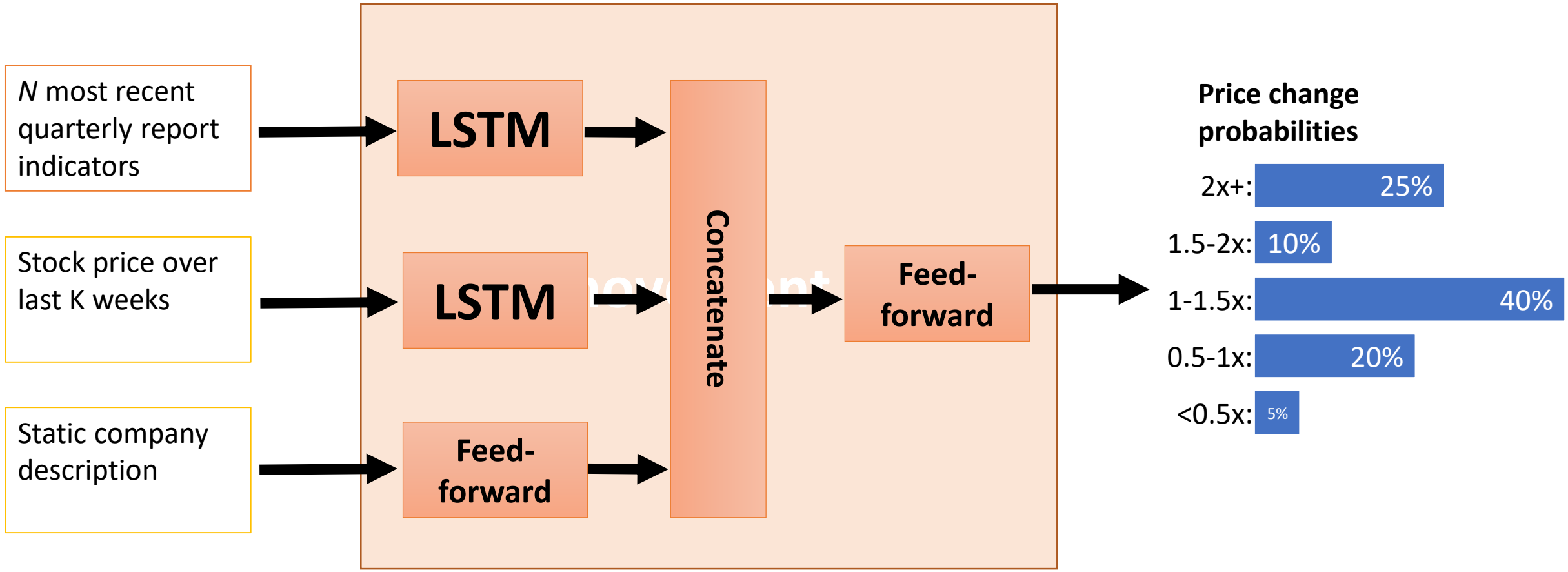
Price change probabilities



Today

Next month

AAPL



Today

Next month

TEXATA™ Big Data Analytics Showdown

AAPL

N most recent quarterly report indicators

Stock price over last K weeks

Static company description

LSTM

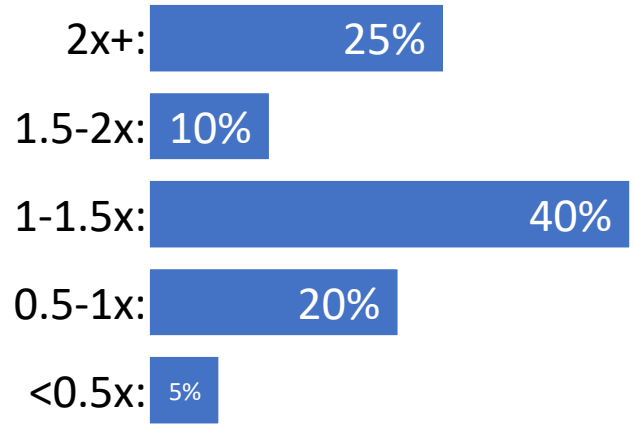
LSTM

Feed-forward

Concatenate

Feed-forward

Price change probabilities

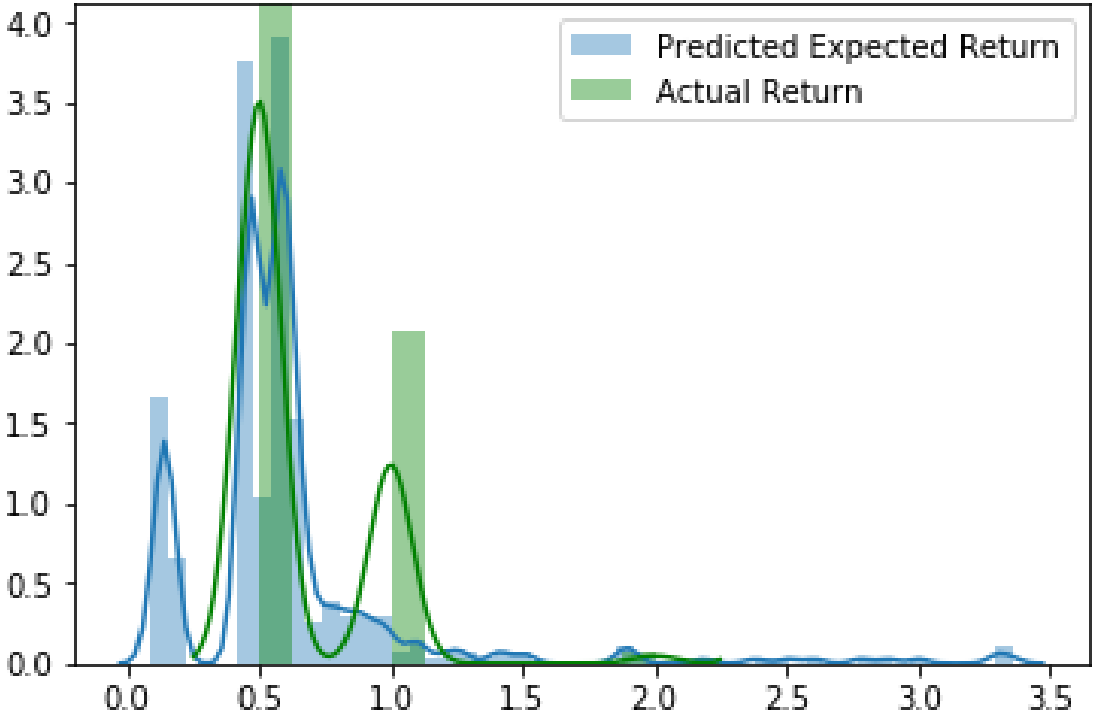


Today

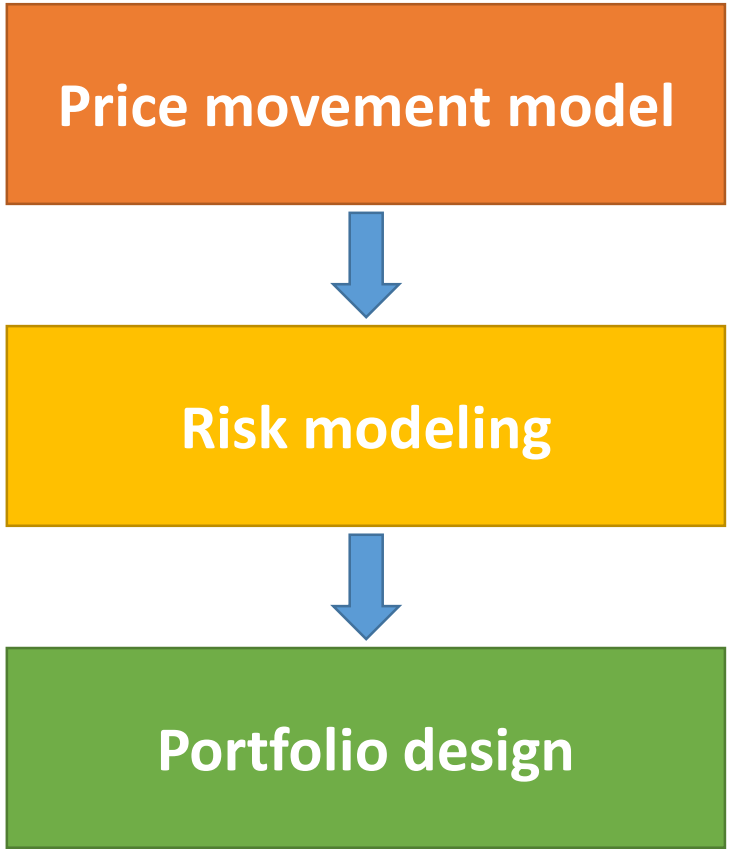
Next month

TEXATA™ Big Data Analytics Showdown

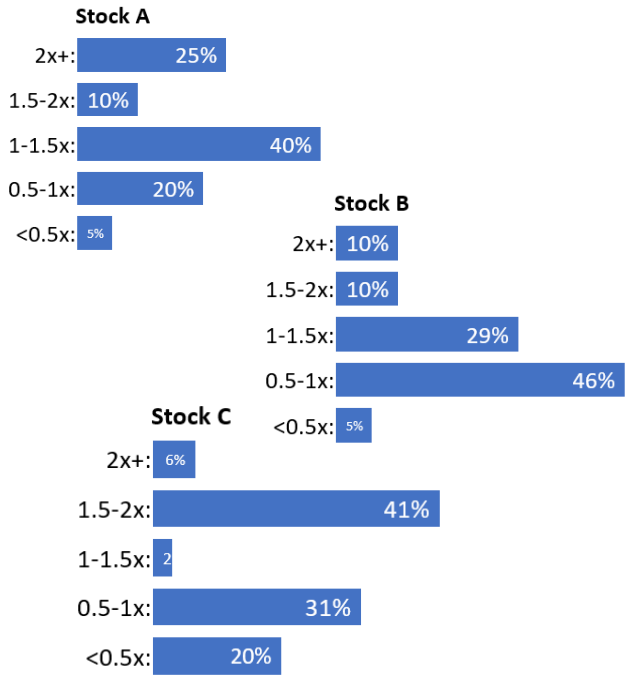
Results


Predicted Expected Return
Actual Return
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5

The Texata solution



The Texata solution



Risk modeling

Expected return

$$ER = E\left\{\sum_i p_i C_i\right\}$$

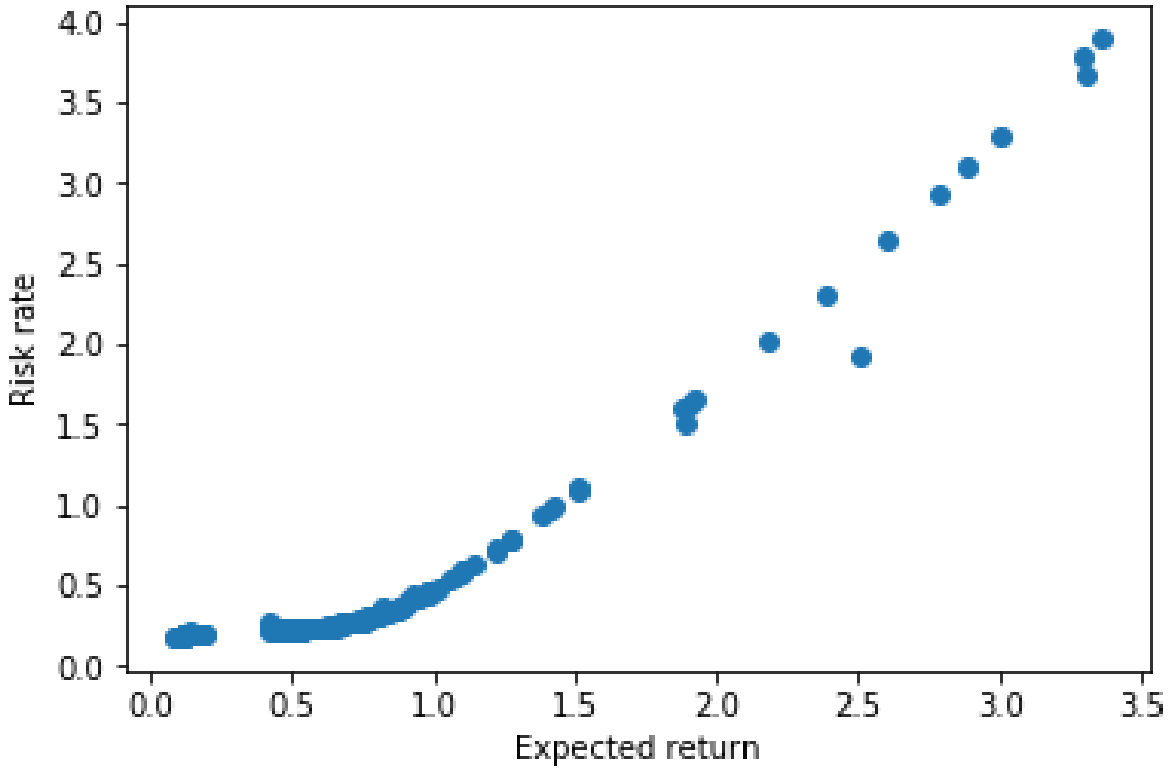
Volatility ("Risk")

$$V = E\{\text{stdev}(R)\}$$

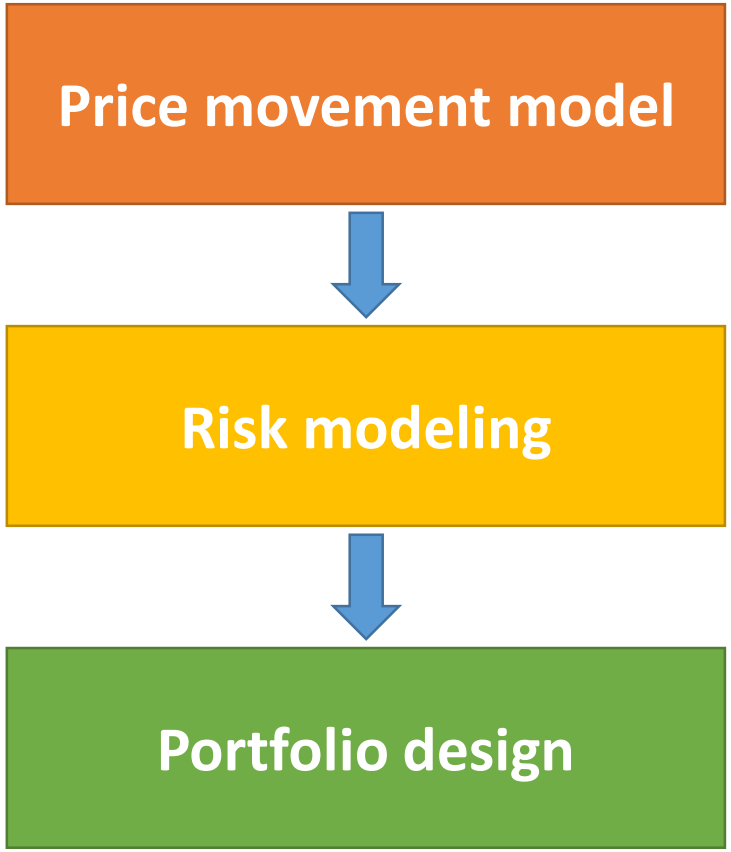
Diversity, Regret, IRR, etc

TEXATA™ Big Data Analytics Showdown

Result: Risk/Return curve



The Texata solution



The Texata solution

Portfolio design

The Texata solution

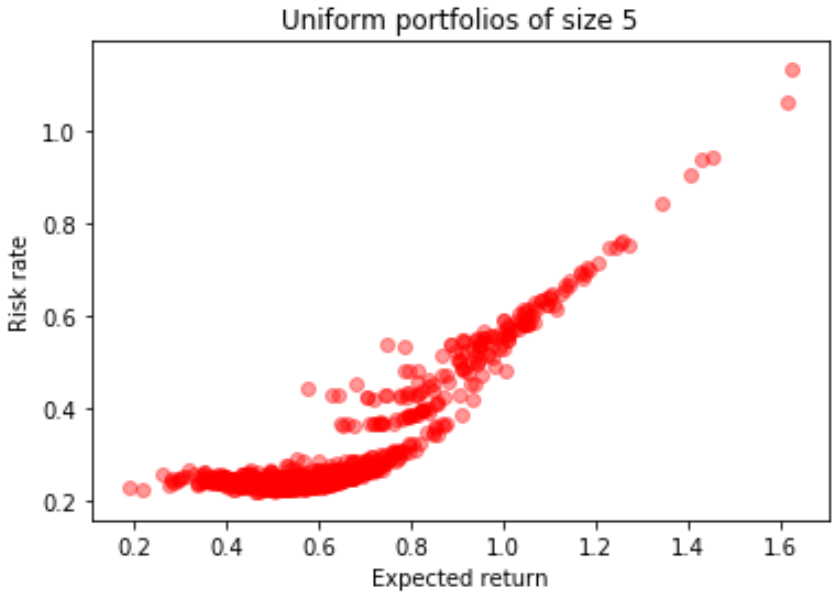
$$\text{argmax}_{\boldsymbol{w}} ER(\boldsymbol{w})$$

$$s.t.$$

$$Risk(\boldsymbol{w}) < T$$

Portfolio design

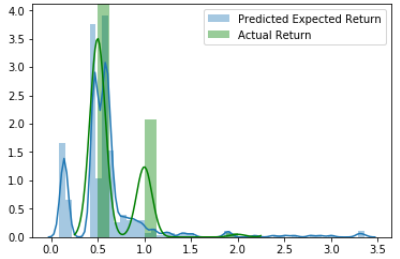
The Texata solution



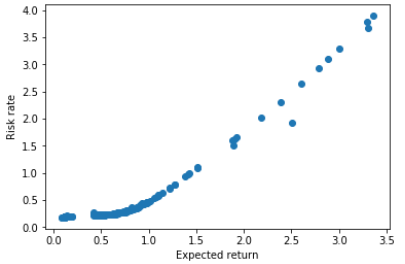
Portfolio design

Summary

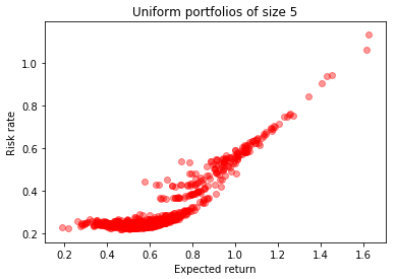
Price movement model



Risk modeling



Portfolio design



Summary

- 1000% in 3 months is not realistic

- 1-5% per month may be achievable, but there's natural risk involved

Tools


K
+
TensorFlow
scikit
learn
SQLAlchemy
matplotlib
jupyter
python
PostgreSQL
BASH
THE BOURNE-AGAIN SHELL
EC2

Links

http://github.com/konstantint/texata-finals-2017/

http://fouryears.eu/