Is it Priced In? Classifying Efficient Markets

1. **Neophytos Charalambides**
2. **Anuvetha Govindarajan**
3. **Tanuj Mathur**
4. **Ram Purandhar Reddy Sudha**
5. **Pinky Thomas**

What is our problem?

Can we predict surprise of an event using:

1. 60-day history of the stock before and after the event
2. 60-day history of the overall market before and after the event

Why should we care?

- Is a stock "Priced In"?
- Efficient Market Hypothesis: do markets reflect current knowledge?
- Application to investment strategies
- Kahneman — 2002 Nobel Prize

# Data Pipeline and Modeling - Overview
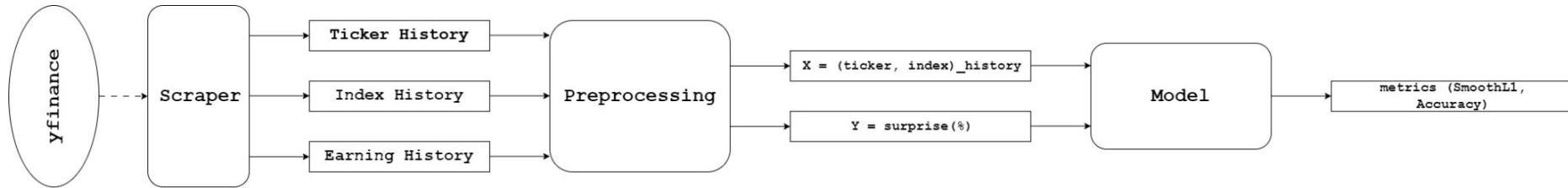
| DATA	Scrapping Source	Type	Time/Scale/Length
The current members of NASDAQ100	Wikipedia	List of symbols	100 members
The five biggest indices in the US stock market	Yahoo Finance	List of symbols	5 members
Historical data of the former list of stocks.	Yahoo Finance	Dictionary of dictionary of time series	5 years(daily)
Historical data of the latter list of indices.	Yahoo Finance	Dictionary of dictionary of time series	5 years(daily)
Earnings event data of the former list of stocks	Yahoo Finance	Dictionary of dictionary of time series	5 years(quarterly)

Data Parsing and Cleaning

| PROBLEMS	SOLUTIONS
Newly IPOed companies	Iterated through the dictionary and popped timeline where data was incomplete.
Stock splits and dividends(rare)	Iterated through dictionary of event to see announcements and popped such cases.
Outliers where companies turned EPS from negative to positive.	Data kept in because it was a relevant event and can be used to observe more such changes.
Choosing between prices each day.	Closing price used across the board.
Minor event data on Yahoo Finance for 2024	Iterated through the timeseries to spot and pop duplicates and extra (minor) events.

Data Visualization
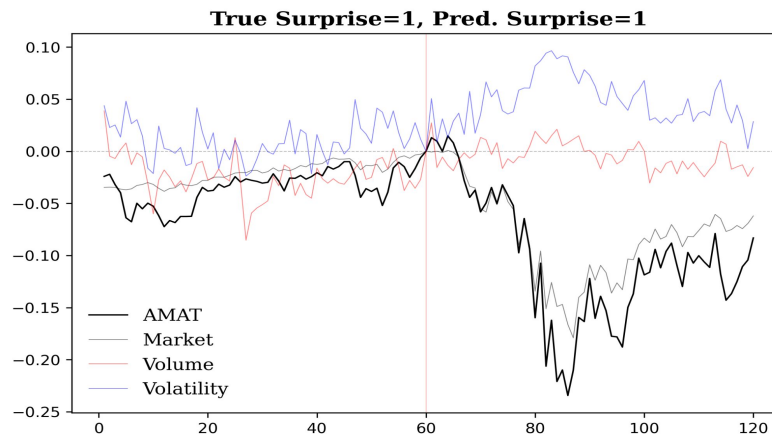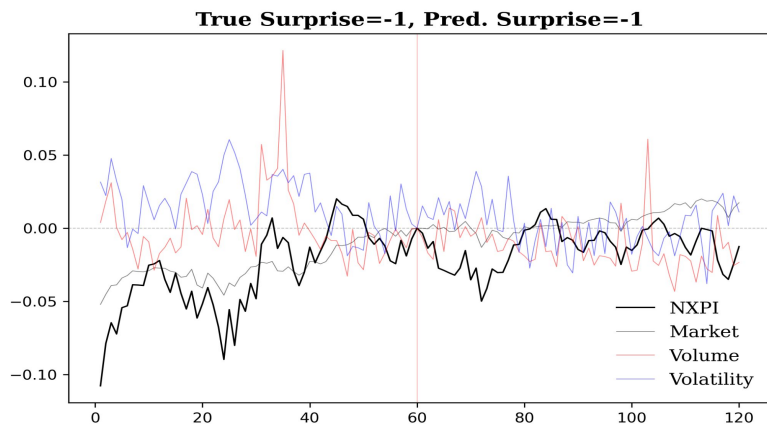
Surprise > 0
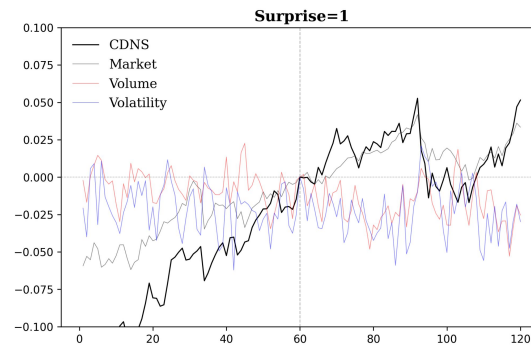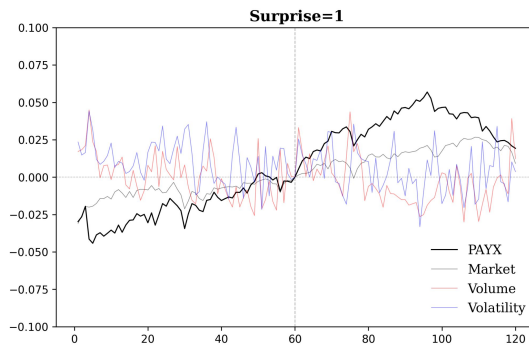
Surprise < 0
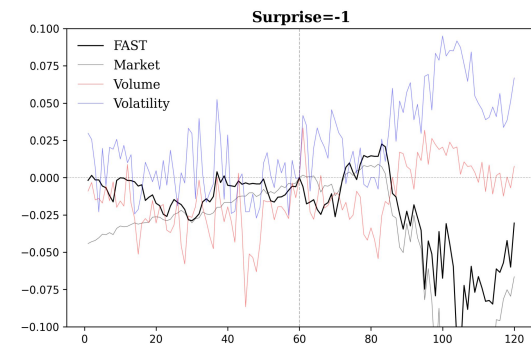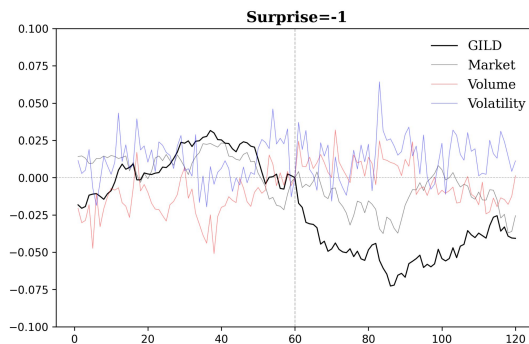
Model

- Best results using an RNN Model, treating the inputs as a time-series of 4 variables (ticker closing, volatility, volume, market closing).

- Architecture: 4 layers with 32 units each, `tanh` nonlinearity.
- `Adam` optimizer with `SmoothL1Loss`.
- Data split into Train, Val, Test ~ (0.7, 0.15, 0.15)
- Early Stoppage based on the Validation Set (patience = 25).
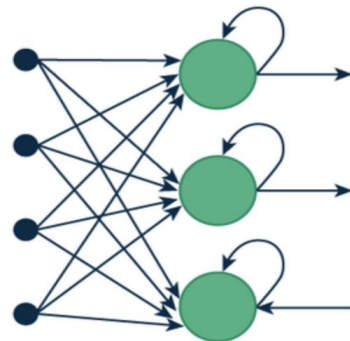
Results & Conclusions

- <u>Overall testing data accuracy</u>: achieved 80% accuracy in predicting surprises
- <u>Model insights</u>: captures temporal dependencies

- <u>Accurate prediction</u>: alignment between predicted and actual surprises
- <u>Future direction</u>: use large datasets, sentiment analysis

Acknowledgements

We would like to thank:

Alec Clott

Steven Gubkin

Sayantan Khan

Roman Holowinsky