Capstone Project:
Predicting the Stock Market

__Problem Statement__

Is it more profitable for a trading/investing strategy to predict the current market regime (bear or bull market) or predict the forward price returns in advance?

Presented on 27 April 2023 by Tan Jun Jie

Agenda

1. The extreme unpredictability of stock markets
2. A short intro to Hidden Markov Model
3. Part 1: HMM for Daily Market Regime Prediction
4. Part 2: HMM for Daily Price Return Prediction
5. Conclusion & Recommendations
6. Limitations & Caveats

Stock markets are extremely hard to predict due to “**Reflexivity**”

“In situations that have thinking participants, there is a two-way interaction between the participants’ thinking and the situation in which they participate.”

~ George Soros, billionaire hedge fund manager who nearly bankrupted the Bank of England

★MOST SUCCESSFUL

CHIMPANZEE ON WALL STREET

Raven, a six-year-old chimpanzee, became the 22nd most successful money manager in the USA after choosing her stocks by throwing darts at a list of 133 internet companies. The chimp created her own index, dubbed MonkeyDex, and in 1999 delivered a 213% gain – outperforming more than 6,000 professional brokers on Wall Street. "She quadrupled the performance of the Dow and doubled the performance of the Nasdaq composite," said Roland Perry, editor of the *Internet Stock Review*.

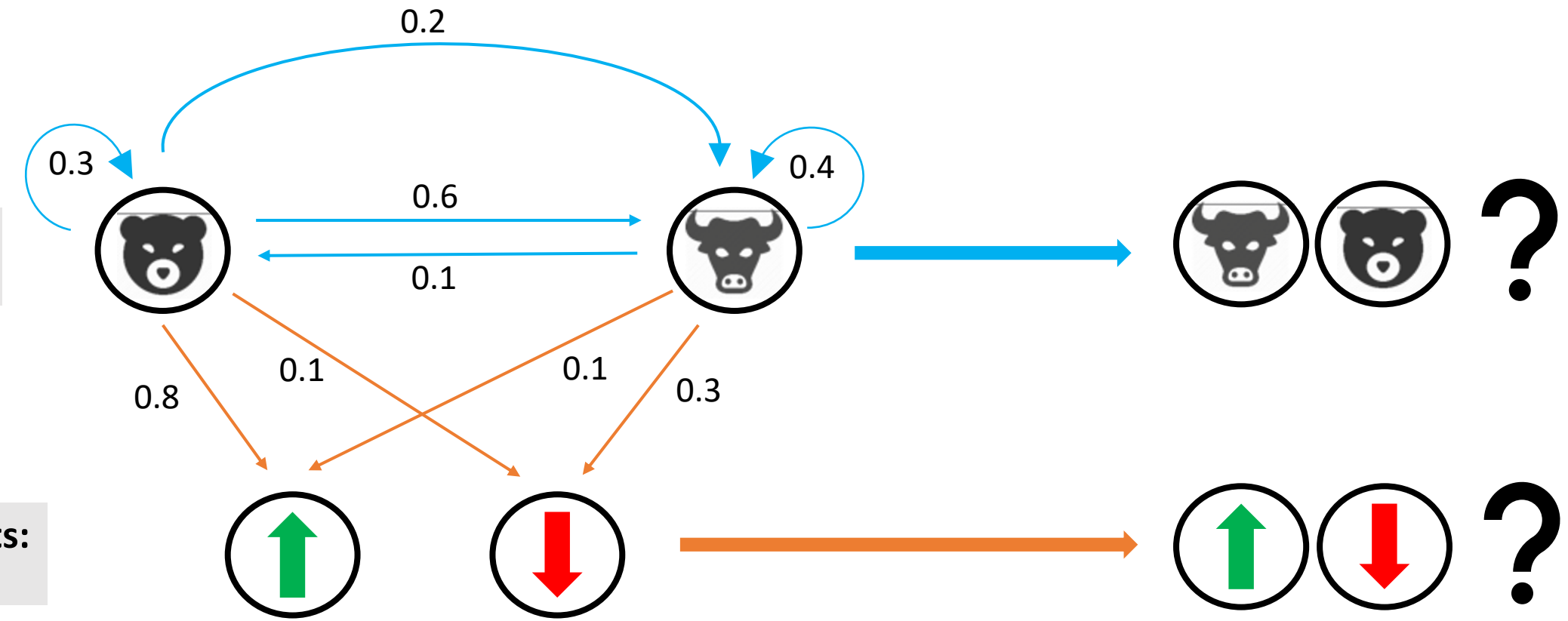

RG. I. PR. . . . . A. AJ. . . . . SS. I. . . . . ST. . . . . SF. I.I. PR. . . .
. . . . . 200. 8 1/2. . . . . 66. 92 3/4. . . . . 20. 99. . . . 16 5/8. . . . . 76.
GU. . . . KM. . . . . APR. . . . . U. . . . . SF. . . . . I.I. PR. . . . Q. . . . . .
. . . 45 3/4. . . . 35 3/4 @ 6. . . . . 97 1/8. . . 100 3/4. . . 64 1/8. . . . . 76. . . 4S. 14. 9
“It’s a bull market.”
Mr Partridge, from The Reminiscence of the Stock Operator

But what defines a Bull or Bear market?

Hidden Markov Model

Hidden States:
2-state Market Regimes

Observable Inputs:
Price Changes



Emission Probabilities

Transition Probabilities

Part 1

Hidden Markov for Market Regime Prediction

Research Methodology

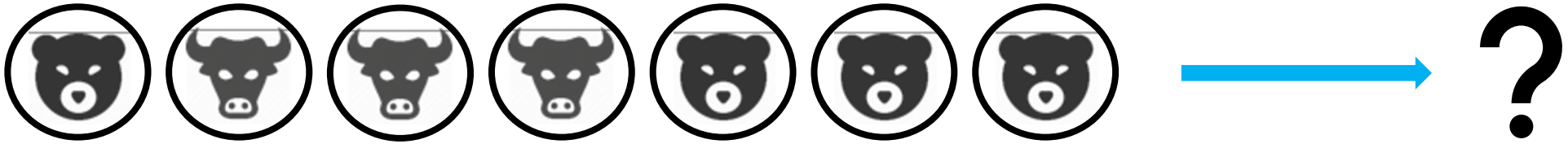
Asset	SPY ETF (underlying index: S&P500 Index)
Model Training Period	Since inception 1993-2009
Model Testing Period	2010-2023/1
Backtest Period	2018-2023/1
Model Evaluation Metric	Cumulative %Returns during Backtest Period

Backtesting Strategy:

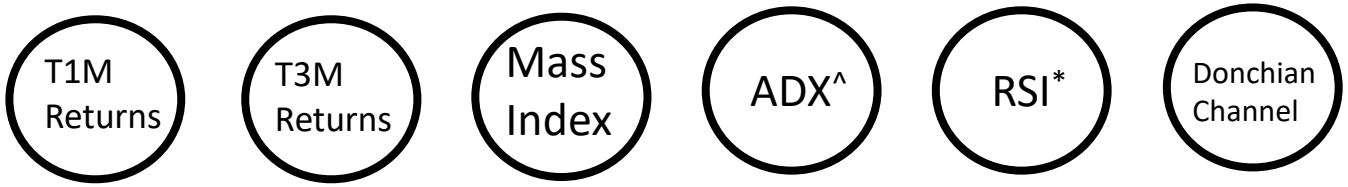
- Daily trading strategy
- Long SPY when last regime=BULL; Short SPY when last regime=BEAR

Converting Price Observations into Technical Indicators

Hidden States:
Market Regimes

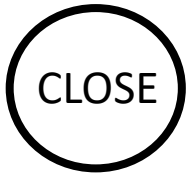


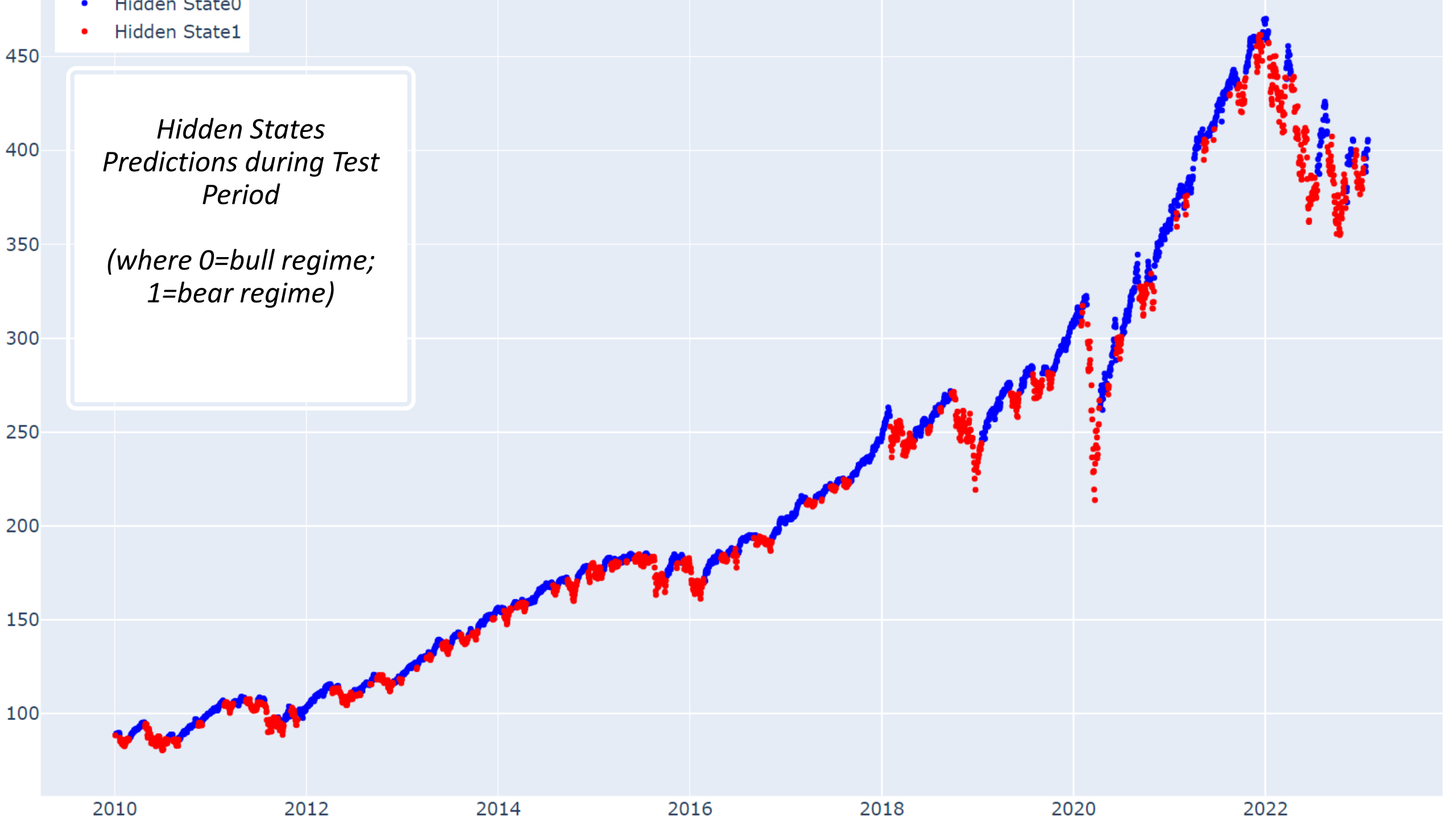
Observable Inputs:
Technical Indicators



^ Average Directional Index
** Relative Strength Index*

Observable Output:
Prices




Hidden State0
Hidden State1
Hidden States Predictions during Test Period
(where 0=bull regime; 1=bear regime)
450
400
350
300
250
200
150
100
2010
2012
2014
2016
2018
2020
2022

Strategy Backtest during 2018-2023/1

Start date	2018-01-02
End date	2023-01-27
Total months	60

	Backtest
Annual return	5.5%
Cumulative returns	31.1%
Annual volatility	19.1%
Sharpe ratio	0.38
Calmar ratio	0.19
Stability	0.36
Max drawdown	-28.6%
Omega ratio	1.07
Sortino ratio	0.55
Skew	NaN
Kurtosis	NaN
Tail ratio	1.02
Daily value at risk	-2.4%
Alpha	0.13
Beta	-0.39

Part 2

Hidden Markov for Daily Price Returns Prediction

Research Methodology

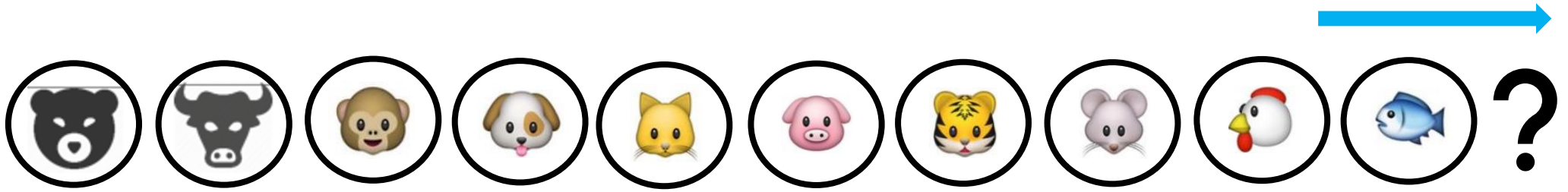
Asset	SPY ETF (underlying index: S&P500 Index)
Model Training Period	Since inception 1993-2009
Sampling Period for "Possible_Outcomes"	2010-2017
Model Testing Period	2018-2023/1
Backtest Period	2018-2023/1
Model Evaluation Metric	Cumulative %Returns during Backtest Period

Backtesting Strategy:

- Day trading strategy (to trade on the OPEN)
- If Predicted Close > Actual Open: Buy SPY at OPEN price and Sell at CLOSE
- If Predicted Close < Actual Open: Sell SPY at OPEN price and Buy at CLOSE

Using Augmented Price Features

Hidden States:
10 states



Observable Inputs:
Augmented Price Features

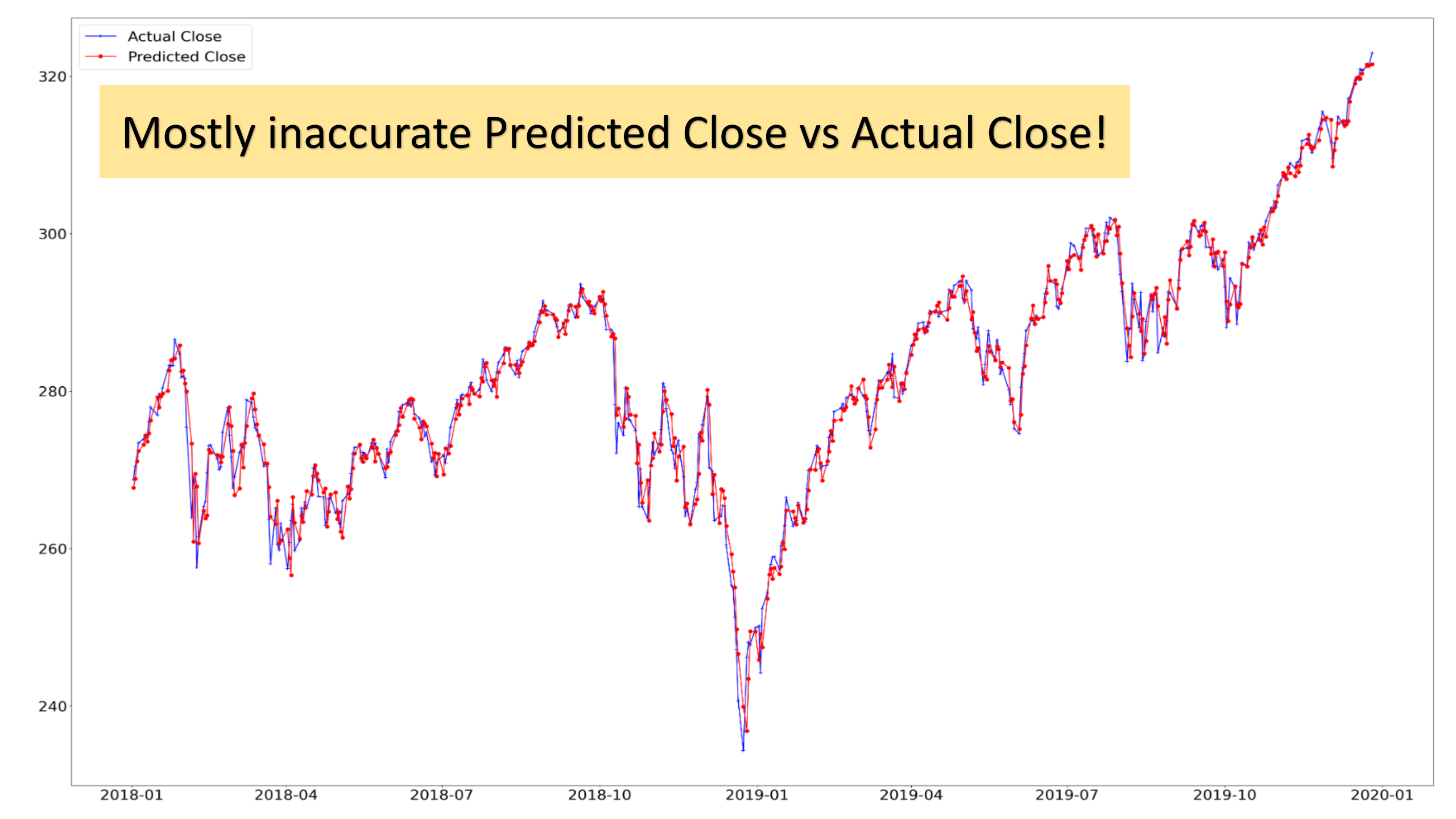
fracohp frachp fraclp

fracohp = (Close-Open)/Open
frachp = (High-Open)/Open
fraclp=(Open-Low)/Open

Observable Output:
Prices

OPEN HIGH LOW CLOSE

CLOSE ?


Actual Close
Predicted Close
Mostly inaccurate Predicted Close vs Actual Close!
320
300
280
260
240
2018-01
2018-04
2018-07
2018-10
2019-01
2019-04
2019-07
2019-10
2020-01

Strategy Backtest during 2018-2023/1

Start date	2018-01-03
End date	2023-01-27
Total months	60

	Backtest
Annual return	0.8%
Cumulative returns	3.9%
Annual volatility	15.5%
Sharpe ratio	0.13
Calmar ratio	0.02
Stability	0.26
Max drawdown	-32.0%
Omega ratio	1.02
Sortino ratio	0.17
Skew	-0.27
Kurtosis	3.54
Tail ratio	0.95
Daily value at risk	-1.9%
Alpha	0.02
Beta	0.00

Conclusions & Recommendations

- Hidden Markov models proved to be reasonably capable of predicting current market regimes but underperformed considerably when trying to predict forward price returns.

- For the best alpha strategy, it is recommended to employ HMM models to **predict current market regime** just like Mr Partridge and not trade unnecessarily.

But market regime prediction strategy worked beautifully for some assets!

Gold: “GLD”

Energy Stocks: “XLE”

Note: This Model was fitted on a walk-forward rolling 5-year period of daily Close data. Strategy has also been enhanced with annual volatility targeting of 20% to improve risk/reward

Limitations & Caveats

1. Trading Costs such as short-financing costs, trading commissions, slippage and market impact costs have not been simulated in the backtesting.
2. Use of non-timeseries feature inputs may not be sufficient to capture underlying statistical patterns
3. More relevant features could be added to enhance model's accuracy
4. Model Uncertainty
    1. While HMM-clustering of SPY ETF market regimes enabled the model to outperform its benchmark up till Aug20, it failed to predict the market regime correctly following the global fiscal and monetary easing that fueled the unprecedented stock market recovery from Aug20-Dec21. From 2022, the model started to produce accurate regime classifications again.
    2. While the HMM regime-clustering model did not work as well for SPY ETF, it worked phenomenally well for more volatile assets such as oil and gold when a walk-forward training window was used to fit the HMM model.
    3. The choice of training window (static window, expanding window or walk-forward window) can result in vastly different backtesting profitability depending on the asset – this risks explicit overfitting.
    4. The inherent mathematical randomness of the HMM model can also yield very different results resulting in overfitting. To reduce this variance, extra care must be exercised during the cross-validation process.
    5. In all assets backtested with similar models (not all shown here in this slides), there are varying lengths of time periods when the model underperformed and posed a drag on profitability. This underscores the fact that HMM regime-clustering model, like most quantitative strategies, is no Holy Grail to consistent profitability. Rigorous empirical analysis needs to be conducted to reduce overfitting variance and ensure the statistical significance of regime signals.
    6. As with all trading strategies, past performance is NOT an indicator of future performance (backtesting over longer periods of time easily proves this)
5. Due to the inconsistencies of quantitative strategies, one needs to be especially clear about the characteristics of an asset that allowed the strategy to work and re-train the model if such characteristics changes in the future.

What's Next

- "Best model" forecasts as of 26 April 2023:
- **BULL** market for SPY
- **BULL** market for GLD
- **BULL** market for XLE

Bonus Slides

(not part of my presentation)

Maximum backtest periods which I can get with similar walk-forward fitting methodology and trading strategy with 20% volatility-targeting.

For reference purposes only.

Even longer backtest for GLD

Start date	2010-01-04
End date	2023-04-25
Total months	159

	Backtest
Annual return	12.522%
Cumulative returns	379.876%
Annual volatility	20.507%
Sharpe ratio	0.68
Calmar ratio	0.51
Stability	0.95
Max drawdown	-24.695%
Omega ratio	1.12
Sortino ratio	1.01
Skew	0.17
Kurtosis	2.67
Tail ratio	1.09
Daily value at risk	-2.528%
Alpha	0.15
Beta	-0.06

SLV

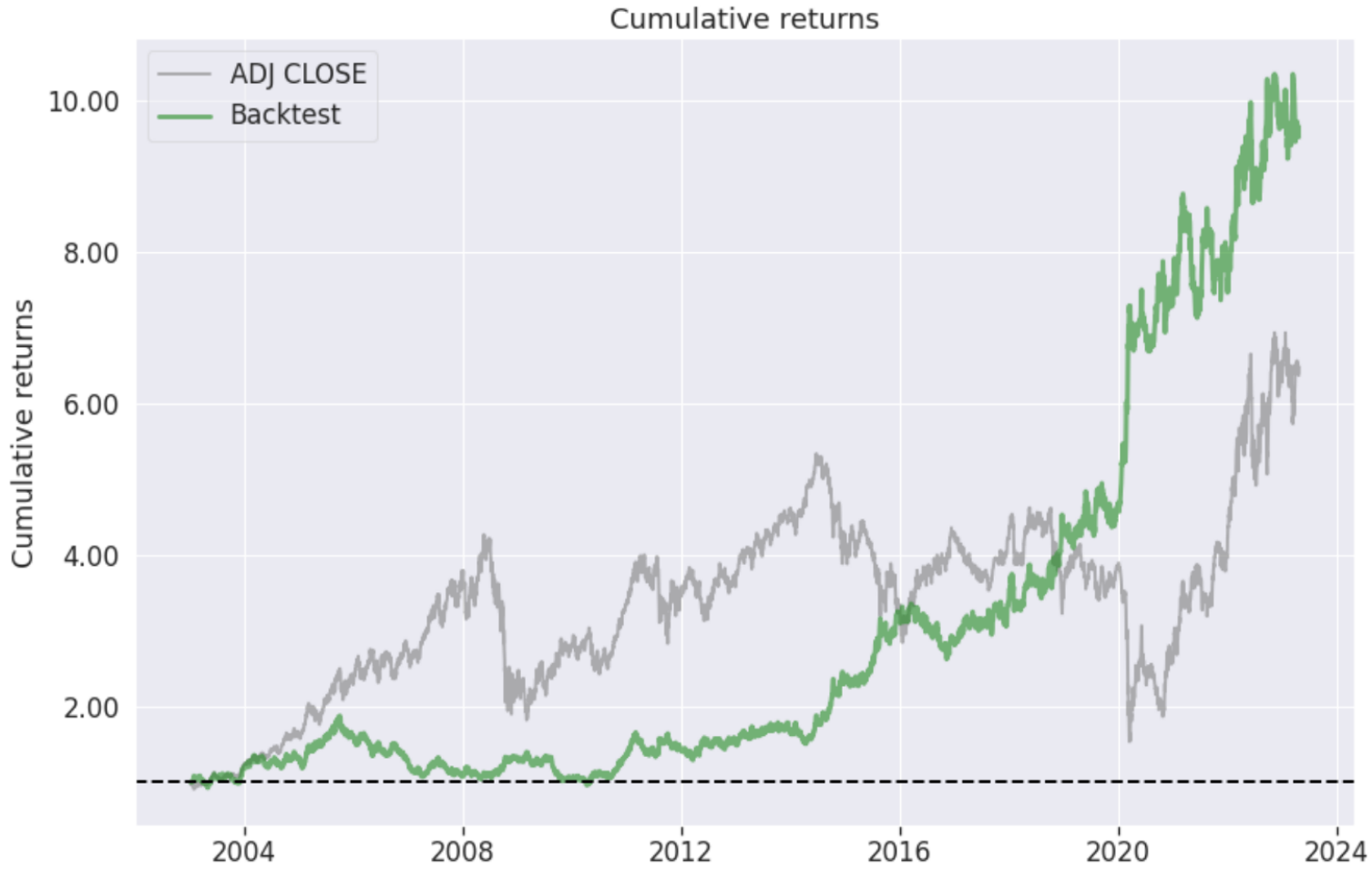
Start date	2012-01-03
End date	2023-04-25
Total months	135

	Backtest
Annual return	17.109%
Cumulative returns	495.144%
Annual volatility	20.569%
Sharpe ratio	0.87
Calmar ratio	0.76
Stability	0.88
Max drawdown	-22.607%
Omega ratio	1.16
Sortino ratio	1.31
Skew	0.20
Kurtosis	2.83
Tail ratio	1.06
Daily value at risk	-2.52%
Alpha	0.20
Beta	-0.14

XLE (did not work until 2018-onwards)

Start date	2003-01-02
End date	2023-04-25
Total months	243

	Backtest
Annual return	11.748%
Cumulative returns	852.35%
Annual volatility	20.341%
Sharpe ratio	0.65
Calmar ratio	0.24
Stability	0.83
Max drawdown	-48.743%
Omega ratio	1.11
Sortino ratio	0.96
Skew	0.18
Kurtosis	1.36
Tail ratio	1.09
Daily value at risk	-2.511%
Alpha	0.17
Beta	-0.16



Some assets which didn’t work

XLV

Start date	2012-01-03
End date	2023-04-25
Total months	135

	Backtest
Annual return	4.327%
Cumulative returns	61.343%
Annual volatility	20.526%
Sharpe ratio	0.31
Calmar ratio	0.09
Stability	0.25
Max drawdown	-48.64%
Omega ratio	1.05
Sortino ratio	0.45
Skew	0.02
Kurtosis	1.37
Tail ratio	1.02
Daily value at risk	-2.561%
Alpha	0.11
Beta	-0.28

SPY

Start date	1999-01-04
End date	2023-04-25
Total months	291

	Backtest
Annual return	0.759%
Cumulative returns	20.132%
Annual volatility	20.488%
Sharpe ratio	0.14
Calmar ratio	0.02
Stability	0.67
Max drawdown	-43.601%
Omega ratio	1.02
Sortino ratio	0.20
Skew	-0.07
Kurtosis	3.01
Tail ratio	1.01
Daily value at risk	-2.57%
Alpha	0.07
Beta	-0.41

VEU

Start date	2012-01-03
End date	2023-04-25
Total months	135

	Backtest
Annual return	-1.883%
Cumulative returns	-19.321%
Annual volatility	20.452%
Sharpe ratio	0.01
Calmar ratio	-0.04
Stability	0.03
Max drawdown	-44.016%
Omega ratio	1.00
Sortino ratio	0.01
Skew	0.01
Kurtosis	2.09
Tail ratio	0.99
Daily value at risk	-2.576%
Alpha	0.01
Beta	-0.15