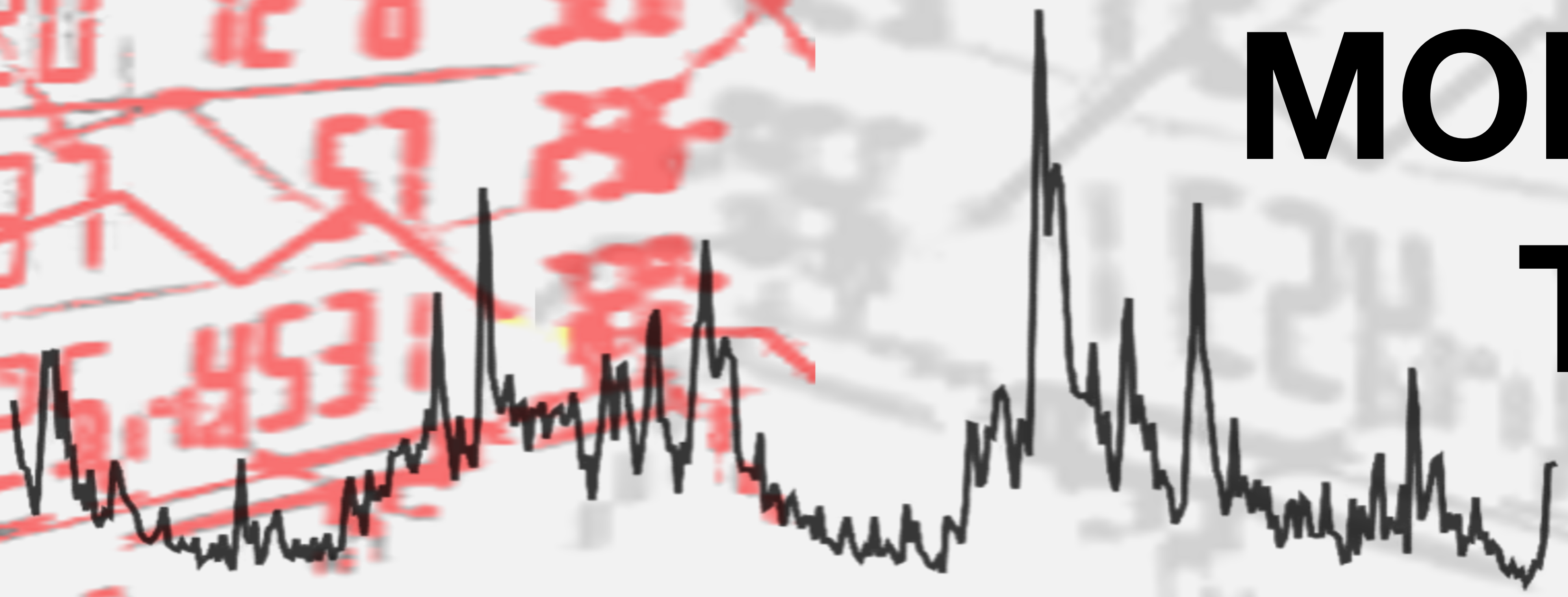


MODELING THE VIX

By OptyX

Kfir Bar
Aarchit Malhotra
Aliza Naqvi
Nicole Roberts
Wilson Rosa

Intro: *What is the VIX?*

The Cboe Volatility Index (VIX):

Real-time index that **represents the market's expectations for the relative strength of near-term price changes or future volatility of the S&P 500 Index (SPX),** which is considered the leading indicator of the broad U.S. stock market.

Being a forward-looking index, it is **constructed using the implied volatilities on S&P 500 index options (SPX options).**

Volatility, or how fast prices change, is often seen as a way to gauge market sentiment, and in particular the **degree of fear** among market participants.

Because it is **derived from the prices of SPX index options with near-term expiration dates,** it generates a *30-day forward projection of* *<u>volatility</u>*.

Intro: *Why Model the VIX?*

Strategizing with VIX Forecasts:

The VIX exhibits **predictive timing properties**, which allows insight into the timing of indicators and assists traders in enhancing long-term fundamentals to better execute their market entries.

Volatility forecasts are **used for risk management**, alpha (risk) trading, and the reduction of trading friction.

Improving the forecasts of future market volatility assists fund managers in adding or reducing risk in their portfolios as well as in **increasing hedges to protect their portfolios** in anticipation of a market sell-off event.

Overview

PART I: *LSTM to Predict VIX*

Predict how VIX will evolve over time using 5.5 years of historical data (2017 through Q2 2022):

- Use **LSTM** to get a predicted VIX curve
- Model **3 days**, **5 days**, and **20 days** out
- Use several approaches:

1. ***Macroeconomic*** *approach:*
    - *Using macroeconomic features, including Sentiment Analysis*
    - *Macro variable-approach is a popular*
    - *method in literature review*
2. ***OptyX approach:***
    - *Using SPX options data as features*

3. Permutations of approaches #1 and #2

OPTYX

Overview

/ Modeling the VIX

PART II: *Traditional VIX Simulations*

- Facebook Prophet
- Monte Carlo

PART III: *Scenario Simulation and OptyX recommendations*

Simulate scenario and study impact on VIX

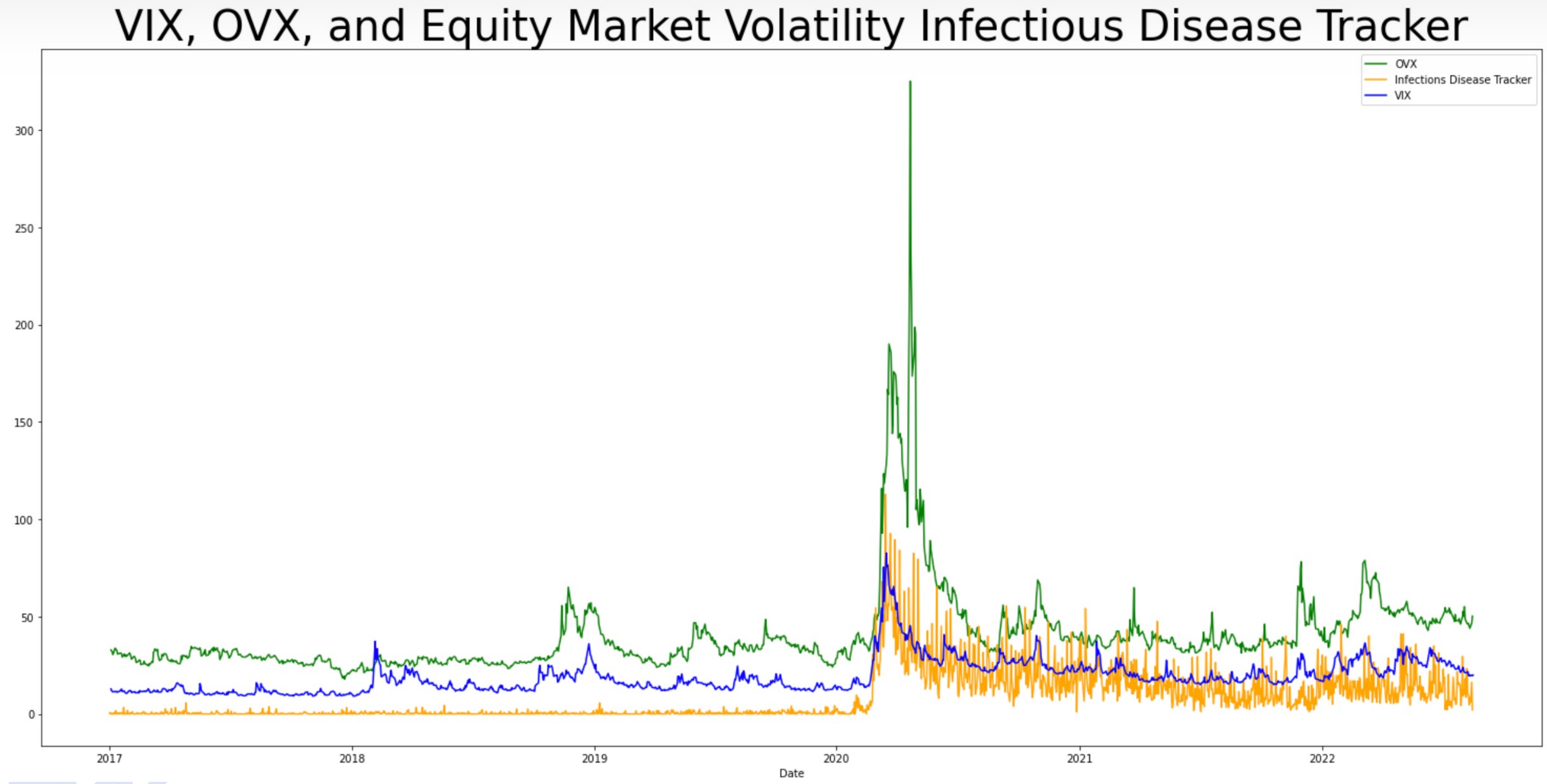
- Period of hyper inflation and pandemic

PART I: LSTM to predict the VIX, Macroeconomic Features

/ Modeling the VIX

- ***VIX***
- *Fed Rate*
- *Inflation – 10 and 5 Yr Breakeven Inflation Rate*
- *Sentiment analysis (details to come)*
- ***CBOE Crude Oil ETF Volatility Index (OVX), 77.2% correlated***
- *US Dollar Index*
- *US Dollar Euro Spread*
- ***ICE BofA US High Yield Index Option-Adjusted Spread , 73.9% correlated***
- *Financial Stress Index*
- ***VVIX, 77% correlated***
- *Change in VIX Futures Contract Volume*
- ***Economic Policy Uncertainty Index for the United States, 65.5% correlated***
- ***Equity Market-Related Economic Uncertainty Index, 68.3% correlated***
- ***Economic Market Volatility: Infectious Disease Tracker, 78.4% correlated***
- *10 Year Yield Inflation-Indexed (market yield on US Treasury securities at 10-yr constant maturity)*
- *Initial Jobless Claims SA and NSA*
- *Day of the week*
- *Day of the month*
- *Global Supply Chain Pressure Index*
- *S&P Cryptocurrency Broad Digital Market Index*
- *SPX Index Volume*

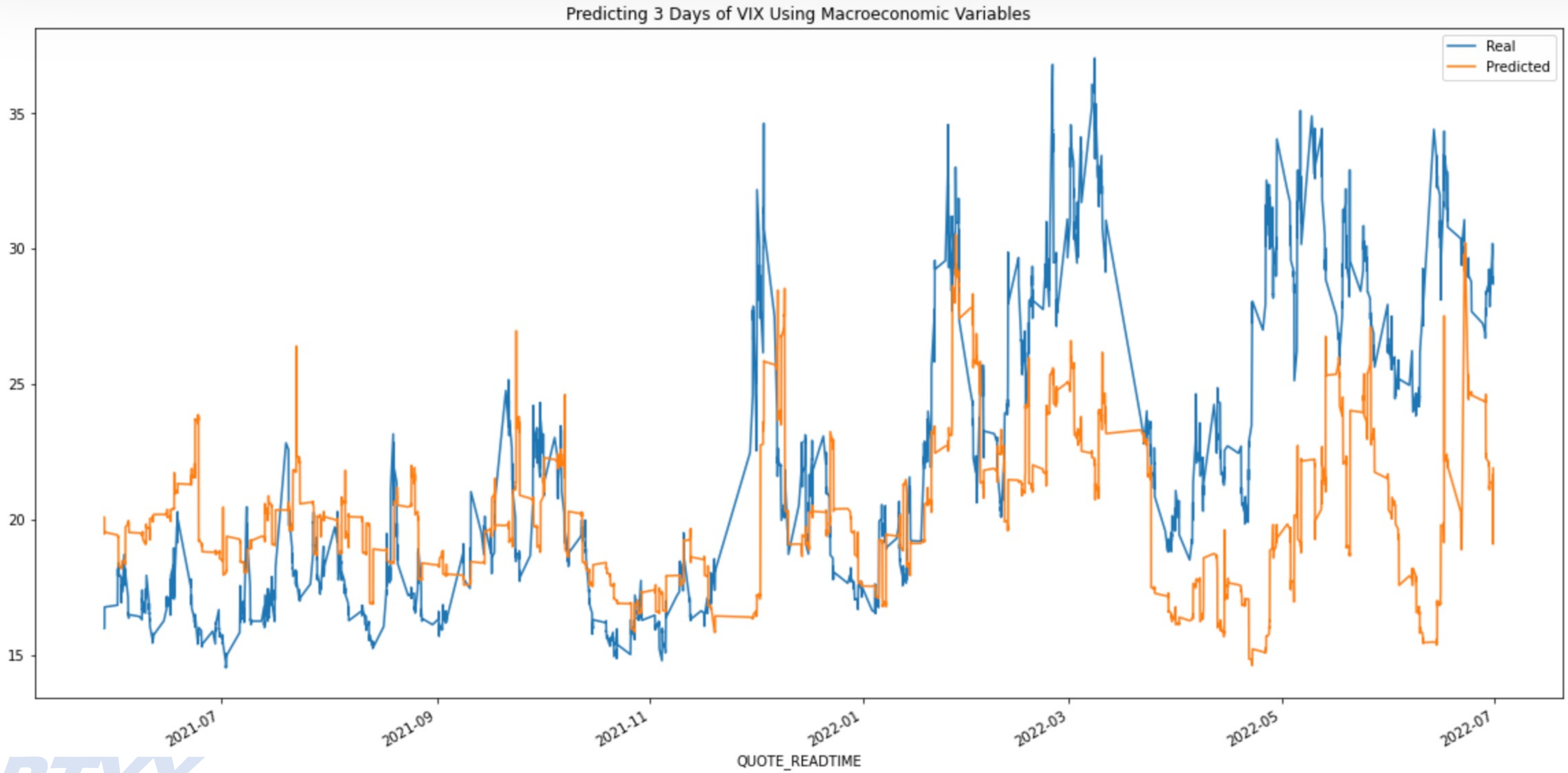
OPTYX


VIX, OVX, and Equity Market Volatility Infectious Disease Tracker
OVX
Infections Disease Tracker
VIX
300
250
200
150
100
50
0
2017
2018
2019
2020
2021
2022
Date

OPTYX

PART I: LSTM to predict the VIX, Macroeconomic Features

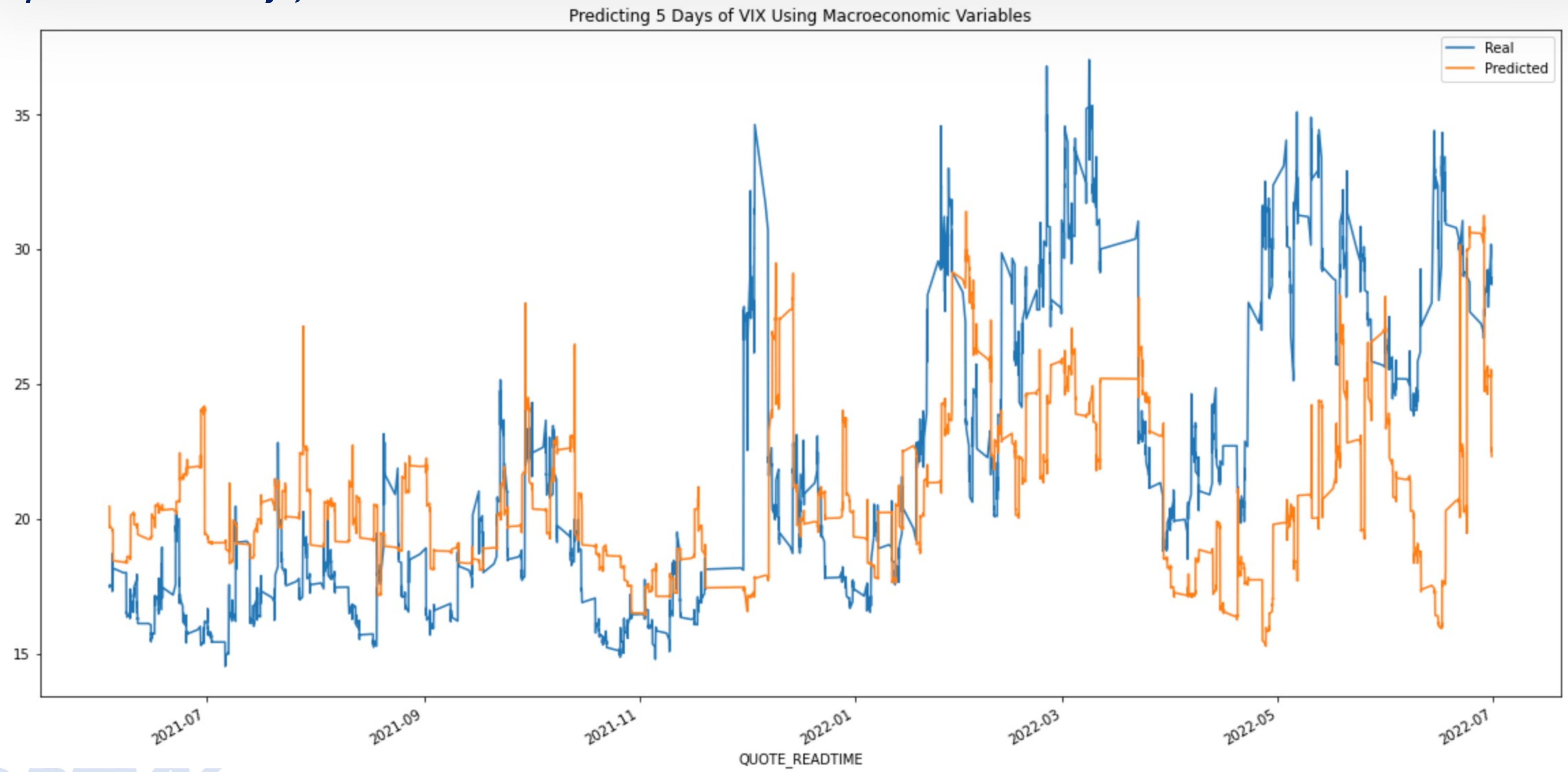
VIX predicted for 3 days, RMSE = 5.380



/ Modeling the VIX

PART I: LSTM to predict the VIX, Macroeconomic Features

VIX predicted for 5 days, RMSE = 5.520



OPTYX

PART I: LSTM to predict the VIX, Macroeconomic Features

Highlight: Sentiment Analysis

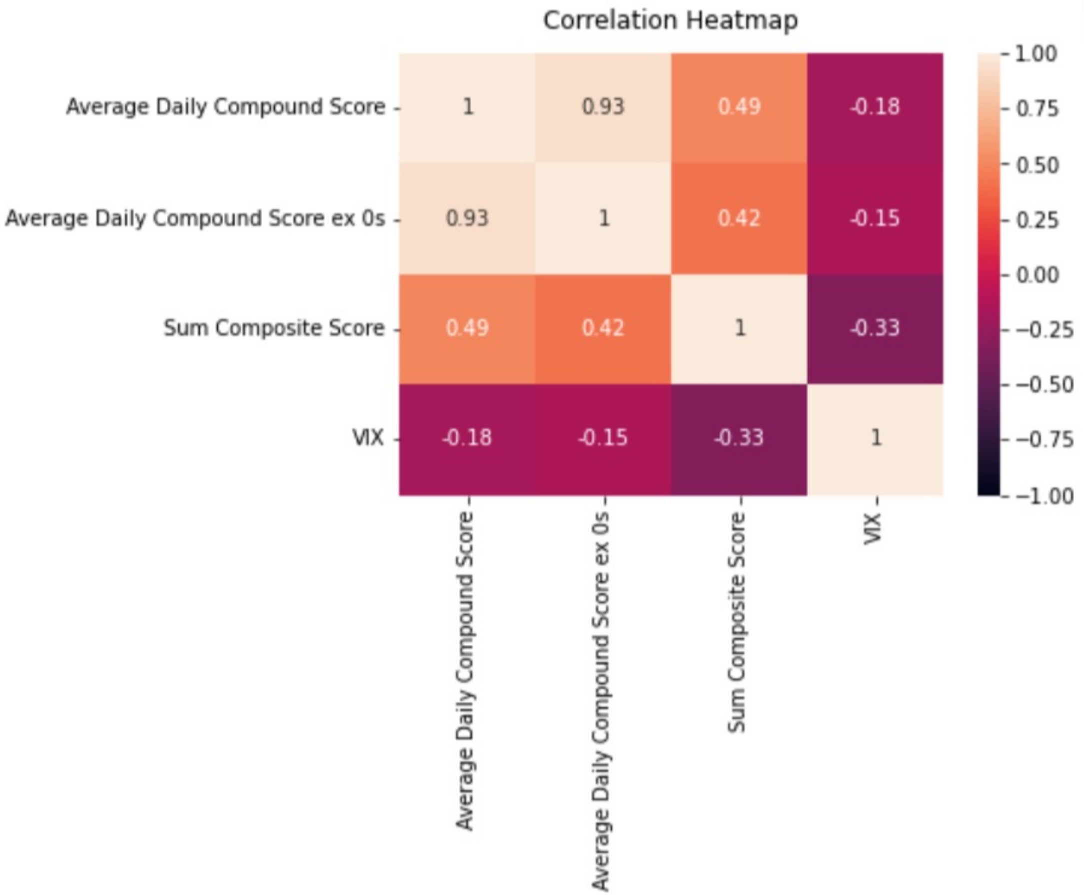
METHODOLOGY:

1. Gathered and cleaned news data from Kaggle, 2008 to 2020.

2. Gathered and cleaned news data from Barron's archive, 2020 to 2022.

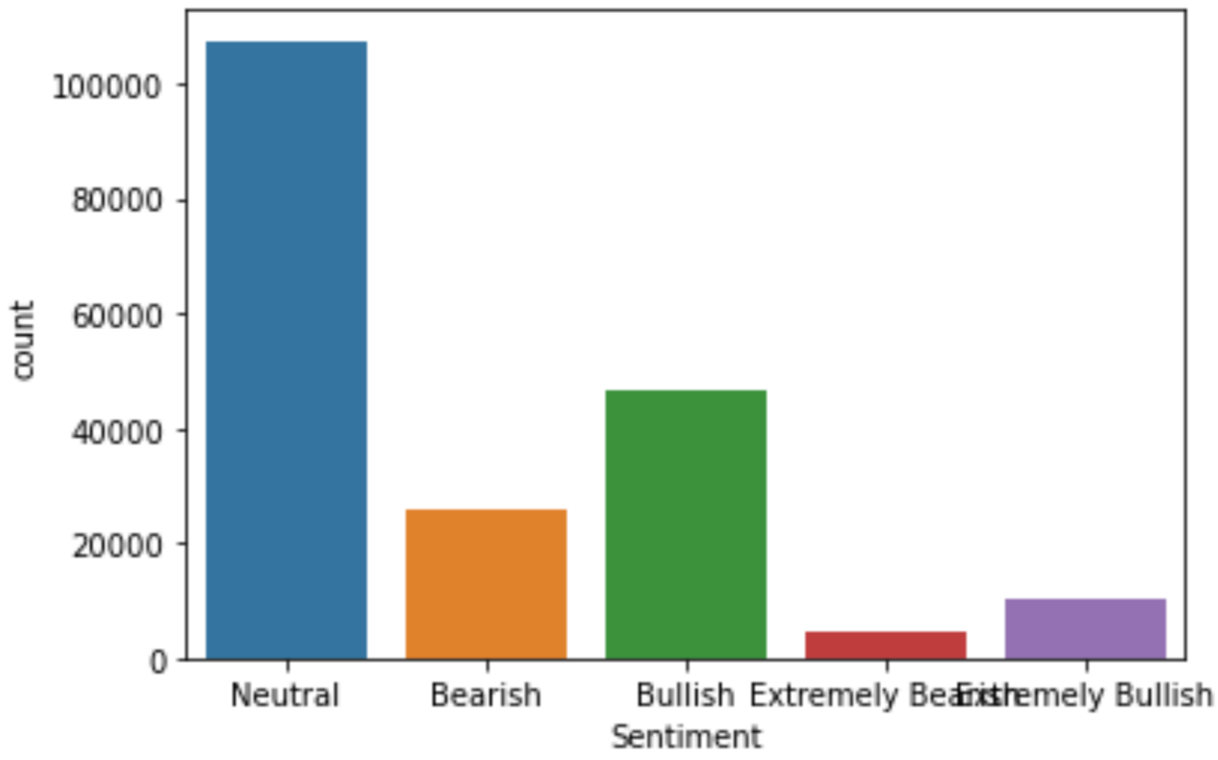
    - *Collectively, gathered ~200k headlines from all sources.*

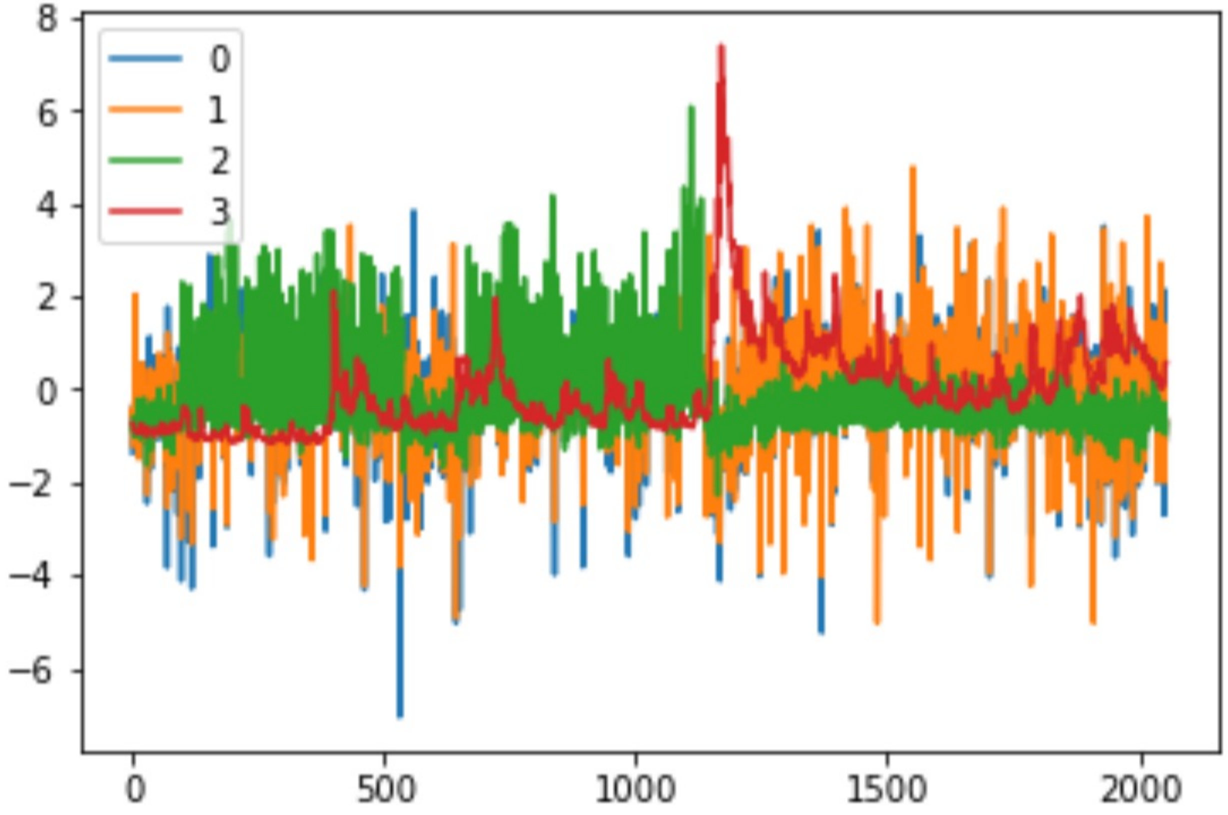
3. Fed the headline data through **NLTK Vader**.



PART I: LSTM to predict the VIX, Macroeconomic Features

Highlight: Sentiment Analysis





PART I: LSTM to predict the VIX, Macroeconomic Features

Highlight: Sentiment Analysis

FINDINGS:

- There's a **slightly negative correlation** between VIX and news based sentiment, which makes sense (*more negative the sentiment, the higher the VIX*).

FUTURE DEVELOPMENTS:

- Feed the headlines through a potentially more precise ML or Deep Learning model.
  - VADER is great because it has a built in Lexicon that is easy to use, but it may not capture every single financial headline as accurately as you would like for a model with predicative capabilities.
- Analyze the content of each article, in addition to the headlines.

PART I: LSTM to predict the VIX, SPX Options Data

Introducing OptyX approach

The **OptyX approach** predicts the VIX on an **intraday scale** using a subset of the **most liquid options.**

OPTYX

METHODOLOGY

All options, at any given point in time, must satisfy the following criteria:

1. ***Expiration date between 27 and 32 days in the future***
2. ***Bid and ask greater than $0.0***
3. ***Strikes within .4% of spot at beginning of day***

PART I: LSTM to predict the VIX, SPX Options Data

OptyX approach

Features consist of **30 min SPX options data for both puts and calls** including:

- *Underlying asset price*
- *Strike price*
- *Strike distance*
- *Greeks (delta, theta, gamma, rho)*
- *Expiration data*
- *Days until expiration*

Each contract is weighted based on their impact on the VIX:

- Every timestamp (30 min intervals) is a weighted average of ~10-20 contracts
- E.g., the *C_DELTA* feature: consists of a **weighted average** of all call option contracts' delta values based on their strike value's impact on the VIX index.

PART I: LSTM to predict the VIX, SPX Options Data

/ Modeling the VIX

How the VIX is calculated: The VIX Formula from CBOE

Weight calculation (each contract's contribution to the VIX)

$$\sigma^2 = \frac{2}{T}\sum_i \boxed{\frac{\Delta K_i}{K_i^2}} e^{RT} Q(K_i) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2$$

σ $\quad VIX/100 \Rightarrow$ **VIX = σ × 100**

T Time to expiration

F Forward index level derived from index option prices

K_0 First strike below the forward index level, F

K_i Strike Price of i th out-of-the-money option; a call if Ki > K0 and a put if Ki M K0; both put and call if Ki =K0.

ΔK_i Interval between strike prices — half the difference between the strike on either side of Ki:

$$\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$$

R Risk-free interest rate to expiration

$Q(K_i)$ The midpoint of the bid-ask spread for option with strike Ki.

OPTYX

PART I: LSTM to predict the VIX, SPX Options Data

Data preparation, getting weights and weighted means

Calculating weights for each strike

```
strikes = vix_spx[['STRIKE']]
strikes
```

```
strikes['Weights'] = (((strikes.shift(-1)) - (strikes.shift(1)))/2) / (strikes**2)
```

$$\boxed{\frac{\Delta K_i}{K_i^2}}$$

```
strikes['Weights'] = strikes['Weights']*1e6
strikes
```

```
vix_spx = vix_spx.merge(strikes, on=['QUOTE_READTIME', 'STRIKE'])
```

Calculating weighted means for each timestamp, 'QUOTE_READTIME'

```
def grouped_weighted_avg(values, weights, by):
    return (values * weights).groupby(by).sum() / weights.groupby(by).sum()
```

```
df_dict = {}

for i in range(1, (col_count)+1):
    col_name = vix_spx.columns[i]
    series = grouped_weighted_avg(vix_spx.iloc[:, i], vix_spx["Weights"], vix_spx["QUOTE_READTIME"])
    df_i = pd.DataFrame(data=[series], index=[col_name])
    df_i = df_i.T
    df_i.index = pd.to_datetime(df_i.index)
    df_dict[col_name] = series
```

```
vix_spx_w = pd.DataFrame(data=df_dict)
```

PART I: LSTM to predict the VIX, SPX Options Data

Data preparation, DataFrame pre-weighted averages

QUOTE_READTIME	UNDERLYING_LAST	EXPIRE_DATE	DTE	C_DELTA	C_GAMMA	C_VEGA	C_THETA	C_RHO	C_IV	C_LAST	...	P_LAST	P_DELTA	P_GAMMA	P_VEGA	P_THETA	P_RHO
2017-01-03 09:30:00	2255.44	2017-02-03	31.27	0.52607	0.00486	2.65495	-0.48291	0.96272	0.12215	26.8	...	40.80	-0.47370	0.00515	2.65583	-0.50894	-0.96124
2017-01-03 09:30:00	2255.44	2017-02-03	31.27	0.50087	0.00500	2.66042	-0.47591	0.92287	0.11972	22.5	...	39.75	-0.50066	0.00528	2.66121	-0.50202	-1.01468
2017-01-03 09:30:00	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.88091	0.11797	22.3	...	45.27	-0.52698	0.00538	2.65484	-0.49214	-1.06999
2017-01-03 09:30:00	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.83645	0.11585	18.2	...	47.97	-0.55569	0.00544	2.63541	-0.48083	-1.12789
2017-01-03 09:30:00	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.79084	0.11377	15.9	...	50.80	-0.58481	0.00546	2.60112	-0.46821	-1.18641
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2019-12-31 16:00:00	3229.70	2020-01-31	31.00	0.58896	0.00326	3.71961	-0.74016	1.61186	0.12644	60.4	...	38.25	-0.40772	0.00337	3.71160	-0.70279	-1.16096
2019-12-31 16:00:00	3229.70	2020-01-31	31.00	0.57516	0.00346	3.74577	-0.70880	1.57883	0.12016	55.0	...	39.82	-0.42403	0.00343	3.74493	-0.69956	-1.20529
2019-12-31 16:00:00	3229.70	2020-01-31	31.00	0.55814	0.00346	3.77303	-0.71525	1.53184	0.11997	53.0	...	41.23	-0.44059	0.00352	3.77242	-0.69912	-1.25170
2019-12-31 16:00:00	3229.70	2020-01-31	31.00	0.54137	0.00353	3.79336	-0.70837	1.48759	0.11813	48.0	...	41.30	-0.45795	0.00353	3.79234	-0.69067	-1.29805
2019-12-31 16:00:00	3229.70	2020-01-31	31.00	0.52373	0.00360	3.80673	-0.69412	1.44155	0.11553	44.3	...	41.20	-0.47562	0.00374	3.80660	-0.66469	-1.34328

Numerous values for each 30 min interval, e.g. 2017-01-03 09:30:00

OPTYX

PART I: LSTM to predict the VIX, SPX Options Data

/ Modeling the VIX

Data preparation, DataFrame post-weighted averages

QUOTE_READTIME	UNDERLYING_LAST	DTE	C_DELTA	C_GAMMA	C_VEGA	C_THETA	C_RHO	C_IV	STRIKE	P_DELTA	P_GAMMA	P_VEGA	P_THETA	P_RHO
2017-01-03 09:30:00	2255.44	31.27	0.527732	0.005013	2.715271	-0.491086	0.970281	0.121867	2250.215999	-0.471096	0.005291	2.725031	-0.519056	-0.955112
2017-01-03 10:00:00	2255.61	31.25	-14.035054	-1.081735	-454.492120	74.610823	-58.860684	-2.624335	754.974000	-16.795234	-0.872504	-489.422488	73.176973	-38.846796
2017-01-03 10:30:00	22[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	2.748752	754.974000	-23.643736	-0.645145	-473.827875	70.614348	-48.802209
2017-01-03 11:00:00	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.442812	754.974000	-13.911224	-0.436781	-533.150381	89.268435	-34.230123
2017-01-03 11:30:00	22[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	6.345671	754.974000	-7.852180	-1.254957	-513.196883	70.581188	-18.914556
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2022-06-30 14:00:00	3811.83	29.08	3.744802	-0.110532	-134.185225	48.428449	-4.880592	0.423606	1289.725807	0.854053	-0.130433	-153.087217	57.531279	1.696495
2022-06-30 14:30:00	3800.73	29.06	4.350282	0.045261	-134.112152	45.172538	-2.472013	1.033898	1289.725807	1.364270	0.002531	-153.900736	57.216623	4.114147
2022-06-30 15:00:00	3785.87	29.04	3.095905	0.006860	-130.687938	45.159502	0.297873	0.764782	1289.725807	2.518978	-0.089376	-158.449023	55.648410	4.406030
2022-06-30 15:30:00	3787.71	29.02	3.924564	-0.087650	-134.345189	44.690222	0.638626	0.193899	1289.725807	3.400569	-0.119250	-150.069003	54.643259	6.308437
2022-06-30 16:00:00	3789.10	29.00	0.290235	0.001364	3.839373	-1.445449	0.851737	0.233583	3932.199403	-0.699461	0.001319	3.872044	-1.743560	-2.315180

Single value for each 30 min interval

OPTYX

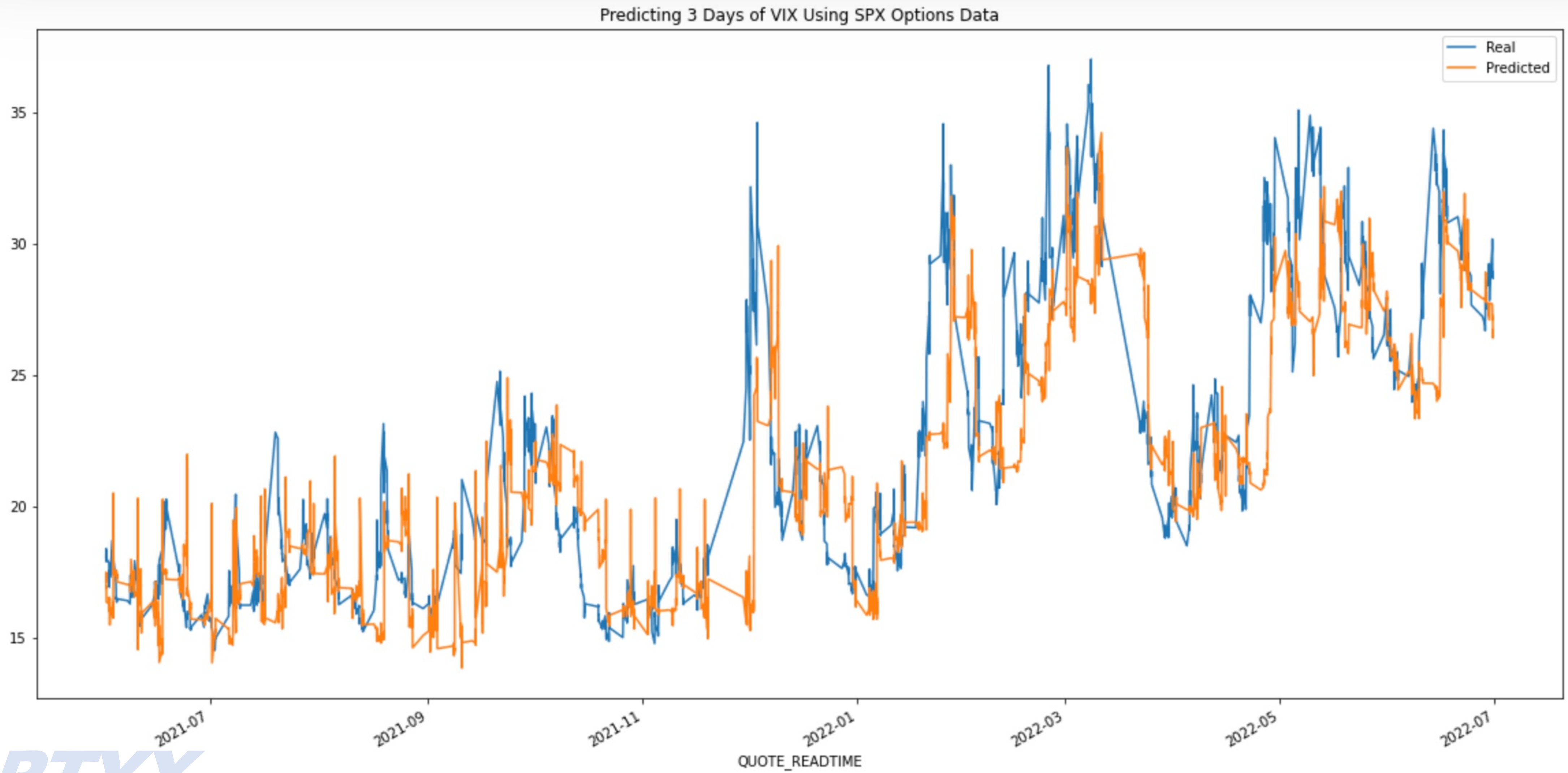
LSTM Model Architecture

- **Multivariate**
- **Multistep (3-5 days)**
- **Stacked**

```
model = Sequential()

model.add(LSTM(units=256, return_sequences=True, input_shape=(a_train.shape[1], a_train.shape[2], )))  #input is (Timesteps, Features):
#E.g. a_train.shape[1] = 60 'steps_to_predict', the X window chunk size used to predict y 'timestep'; E.g. a_train.shape[2] = 5 features

model.add(LSTM(units=64,return_sequences=True))

model.add(Dropout(0.2))  #to prevent overfittingb

model.add(LSTM(units=32))

model.add(Dropout(0.2))   #to prevent overfitting

model.add(Dense(timestep)) #predict N steps into the future, for Multi-step

ADAM=keras.optimizers.Adam(0.0005, beta_1=0.9, beta_2=0.999, amsgrad=False)  #.0005 is the learning rate

model.compile(loss='mse', optimizer=ADAM)  #mean_squared_error  mae

history = model.fit(a_train, b_train, epochs=30, batch_size=64, validation_data=(c_test, d_test), verbose=1, shuffle=False)
```

PART I: LSTM to predict the VIX, SPX Options Data

VIX predicted for 3 days using SPX Options, RMSE = 3.257

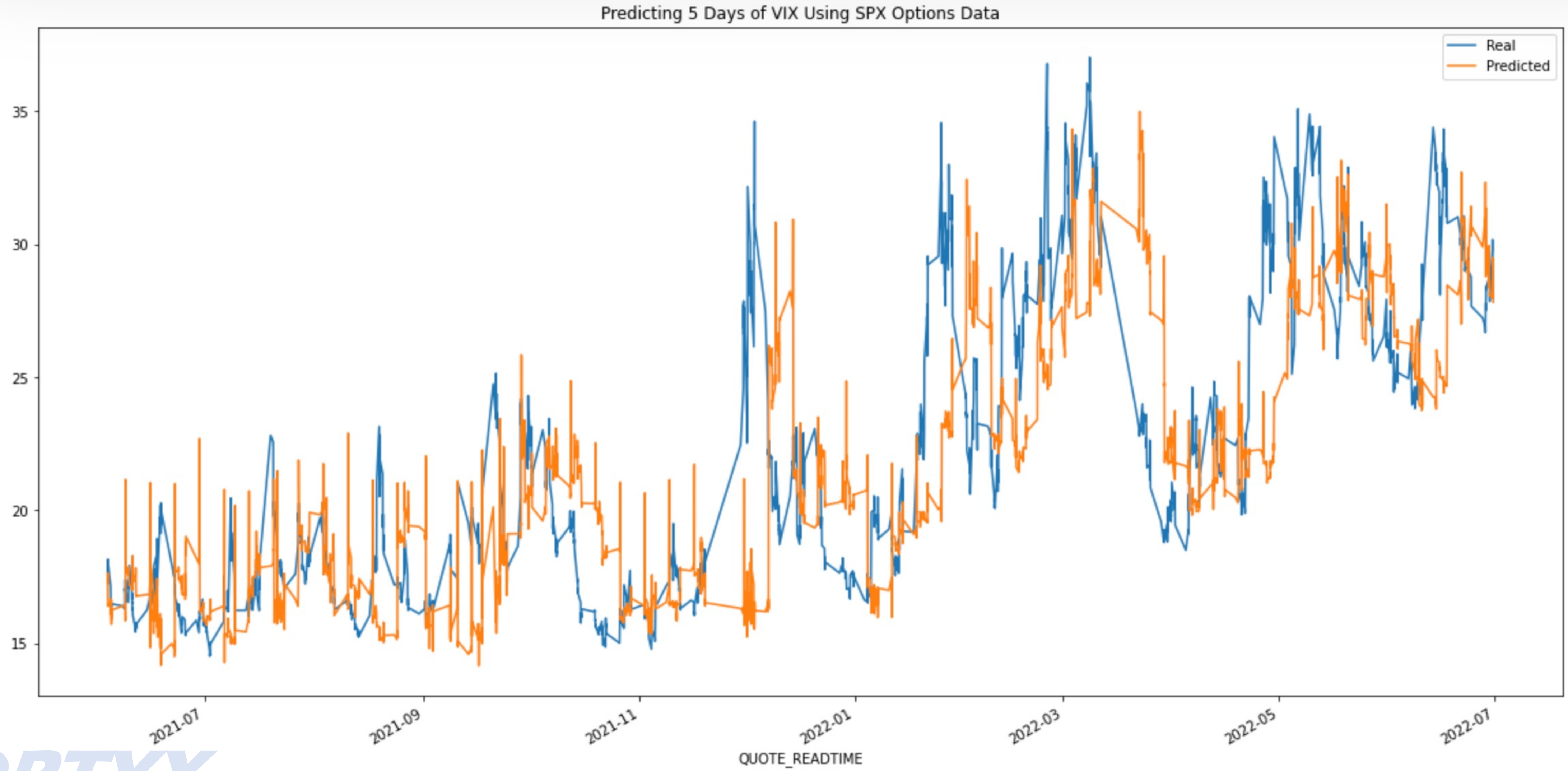


OPTYX

/ Modeling the VIX

PART I: LSTM to predict the VIX, SPX Options Data

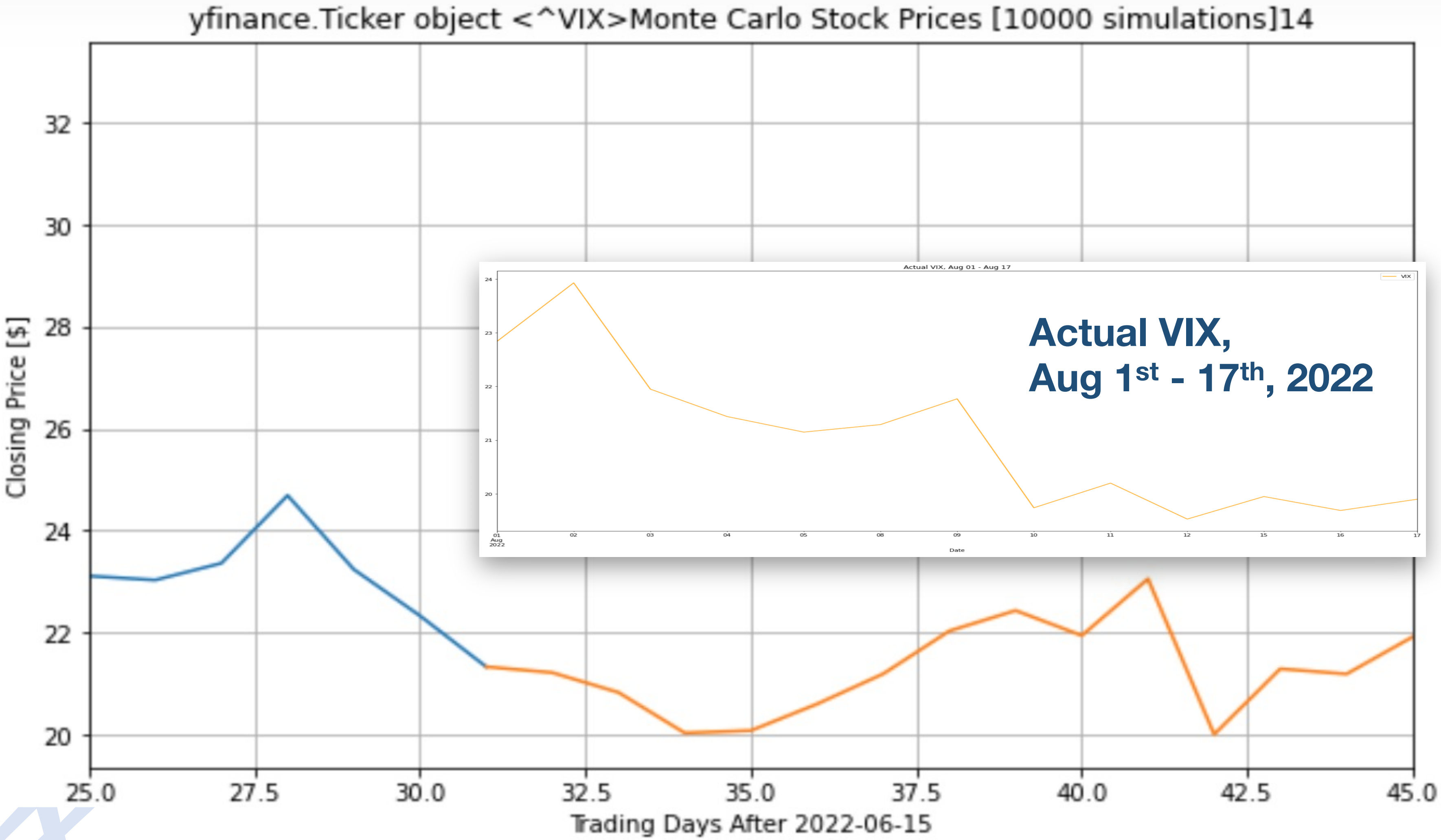
VIX predicted for 5 days using SPX Options, RMSE = 3.957



OPTYX

PART II: Monte Carlo to predict the VIX

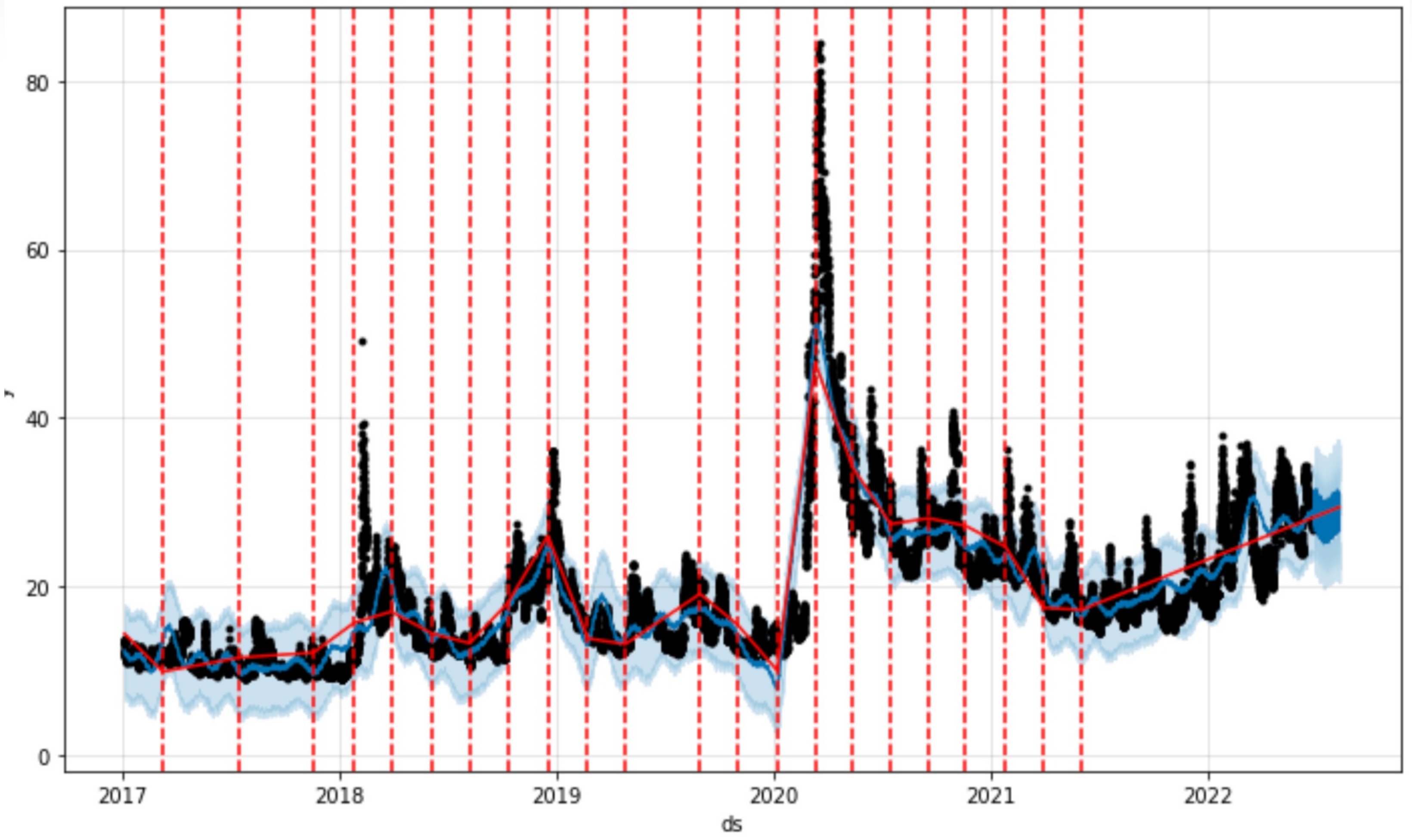
VIX predicted using Monte Carlo Analysis, Aug 1st - 17th, 2022



OPTYX

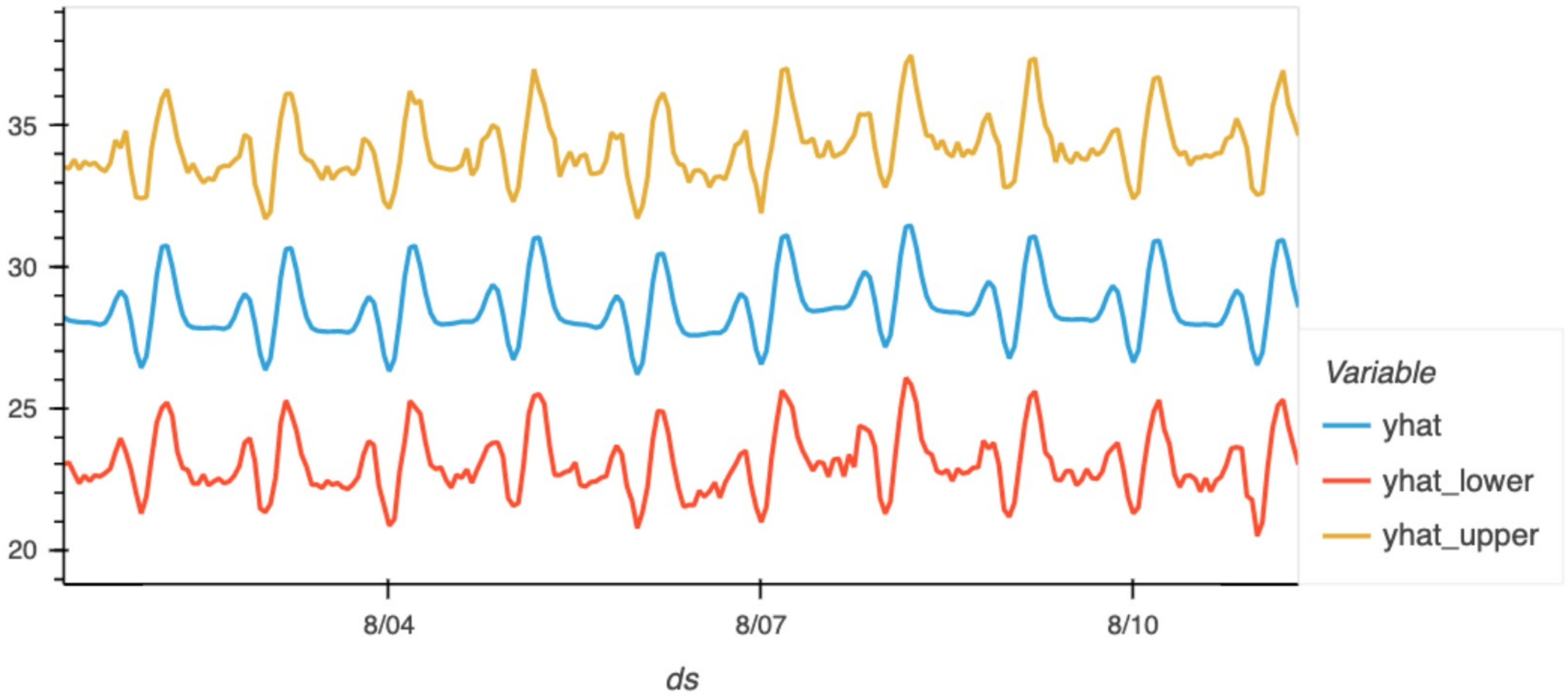
PART II: Facebook Prophet to predict the VIX

VIX predicted for the next 6 months



PART II: Facebook Prophet to predict the VIX

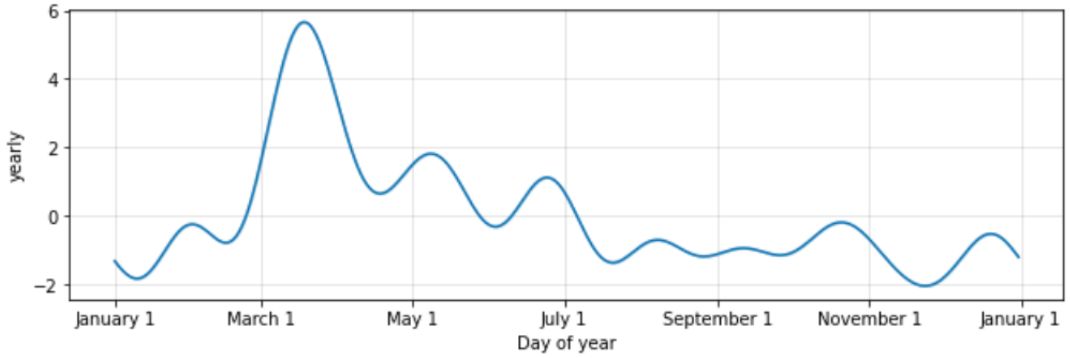
First two weeks in August 2022 prediction

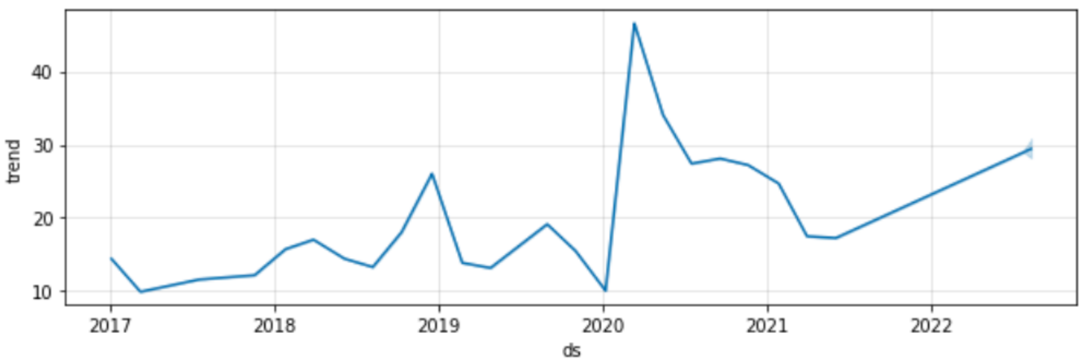


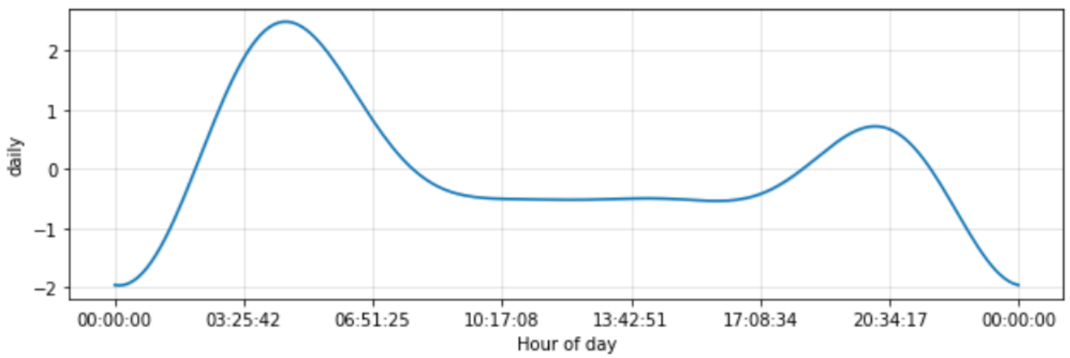
OPTYX

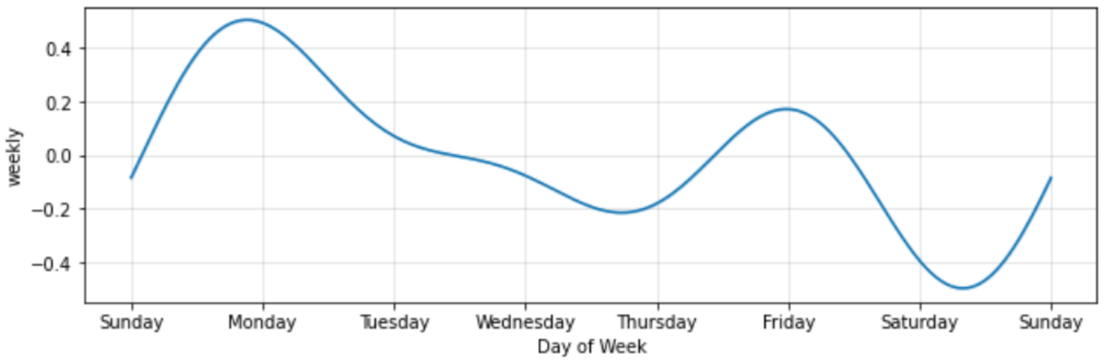
PART II: Facebook Prophet to predict the VIX

Trend analysis





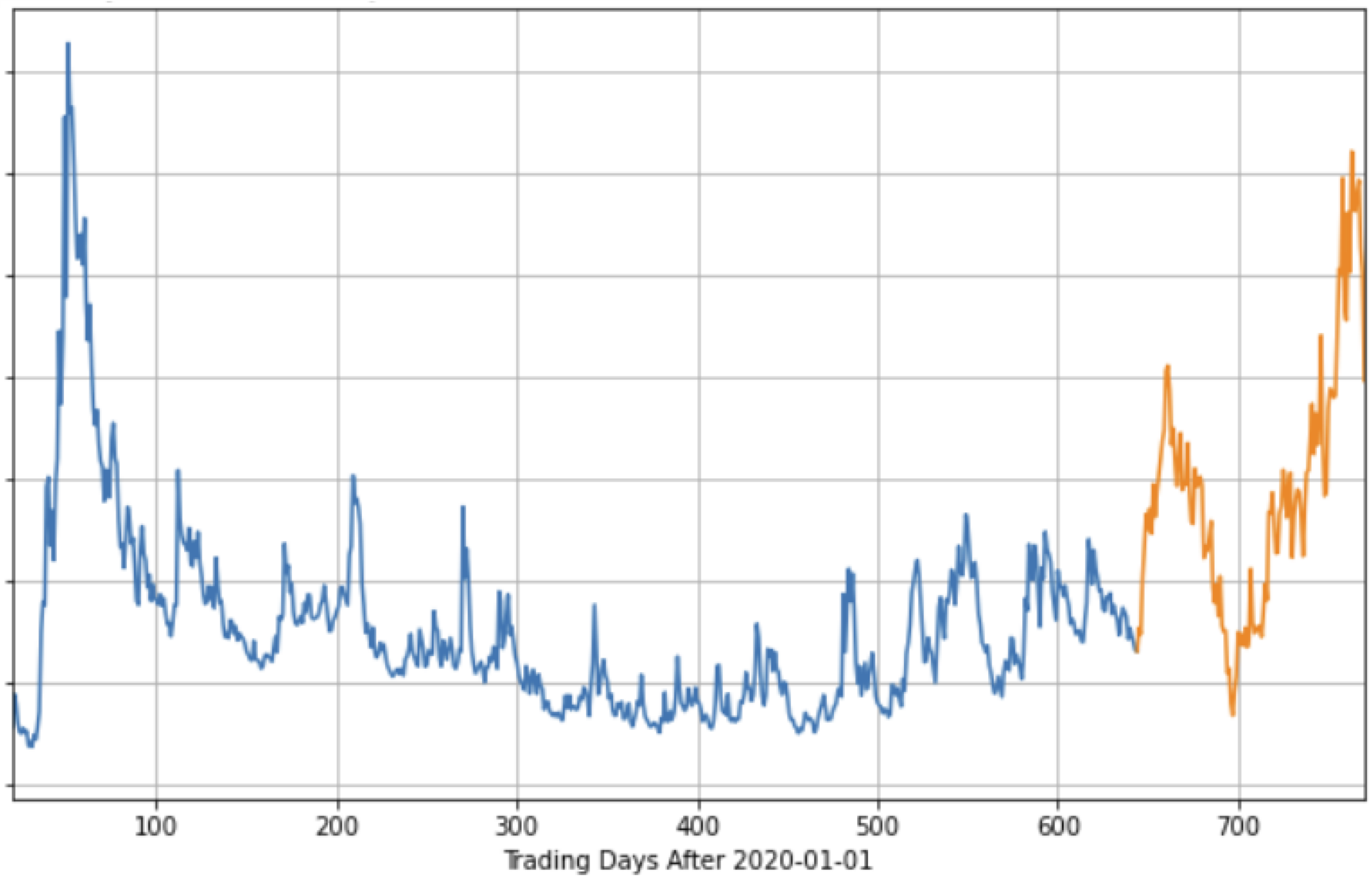




PART III: Scenario Simulation of the VIX

Period of hyper-inflation + pandemic

- Manipulated the following features by scaling them in a step-wise fashion over select periods:
  - ***Inflation***
  - ***Equity Market Volatility Infectious Disease Tracker***
  - ***VVIX***



Thank you.

OPTYX