

# PREDICTING **BITCOIN** PRICE FLUCTUATION WITH TWITTER **SENTIMENT** ANALYSIS

PRESENTED BY CITEX TECHNICAL TEAM





Introduction



Background



Method



Results

# INTRODUCTION

Twitter specifically has become known as a location where news is quickly disseminated in a concise format.

Social media has served as platform to express opinions since their inception, and as such tapping into the open APIs provided of the likes of Facebook and Twitter, these arguably biased pieces of information become available with a sea of meta-data.



## RELATED RESEARCH

- Trading on Twitter: Using Social Media Sentiment to Predict Stock Returns by Sul et al.
  - 2.5 million tweets about S&P 500 firms
  - Sentiment quickly is anticipated to be reflected in a stock price on the same trading day.
  - Yielded 11-15% annual gains.
- Algorithmic Trading of Cryptocurrency Based on Twitter Sentiment Analysis by Colianni et al.
  - Sentiment impact on investment decisions specifically on Bitcoin.
  - By using supervised machine learning techniques, They yielded a final accuracy of above 90% hour-by-hour and day by day.



# PROBLEM STATEMENT

- Is there a correlation between Twitter sentiment and BTC price fluctuation?
- Can a naive prediction model based on sentiment changes yield better than random accuracy?

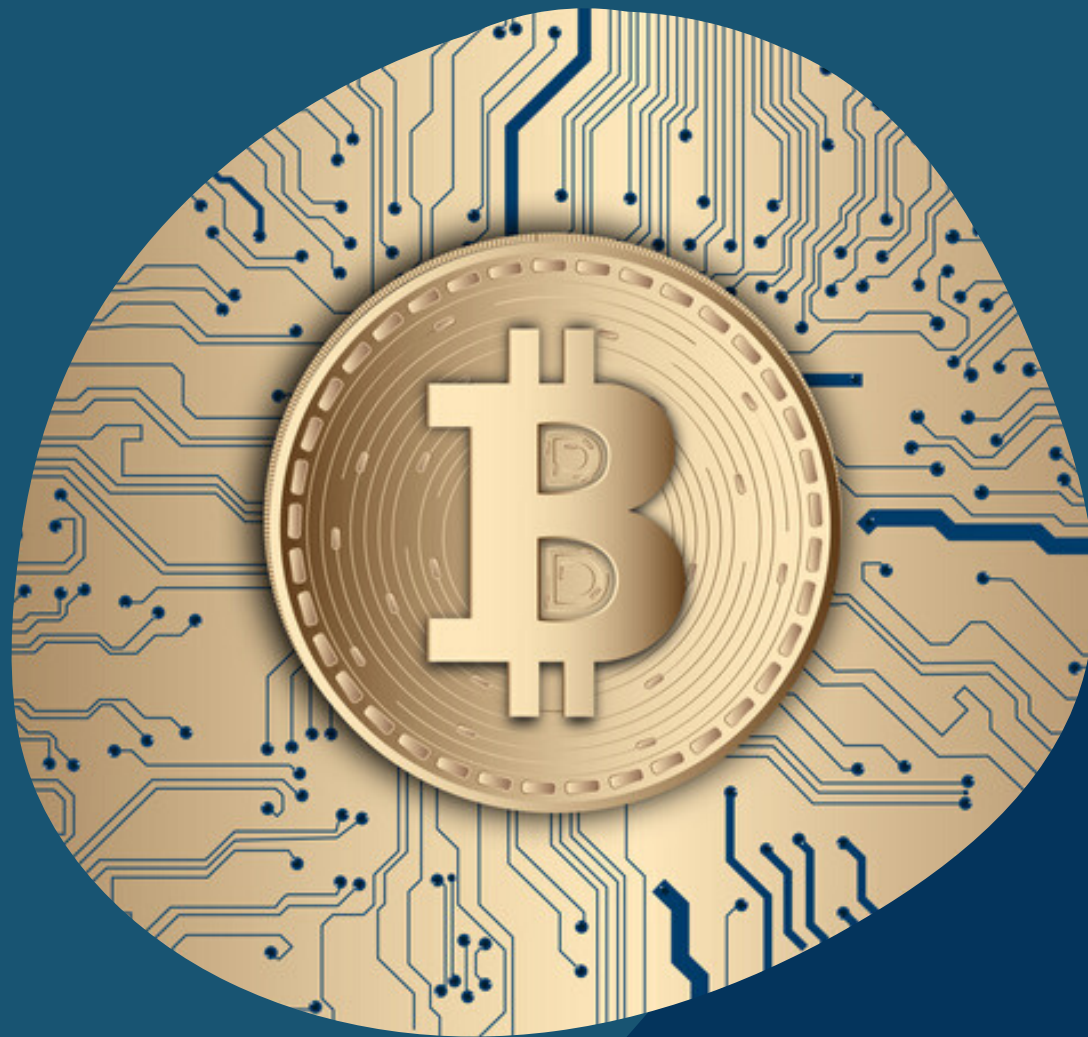


# PURPOSE

## WHY IS IT IMPORTANT TO US?

By comparing sentiment and Bitcoin price at different intervals of time, and optimizing a prediction model given these intervals, a short term analysis of correlation between sentiment and market change can be examined.





## BITCOIN

The value of a currency depends on several factors, the more notable being; public confidence, acceptance, and social expectancy (of value).



## OPINION MINING

Due to the message length restriction and the classifying nature of tweets hashtags, Twitter has become a gold mine for opinionated data, in its semi-structured form. Twitter is widely used as a source when looking for sentiment data sets.



## SENTIMENT ANALYSIS

Sentiment analysis is about finding the underlining opinions, sentiment, and subjectivity in texts , which all are important factors in influencing behaviour.

# METHOD

- 1.Data Collection
- 2.Sentiment Analysis Process
- 3.Deriving Predictions from Sentiment Data





## DATA COLLECTION

1. Historical BTC/USD exchange rate data
  2. Tweets
- ! The datasets were collected using a dedicated server, allowing for uninterrupted continuous data gathering.

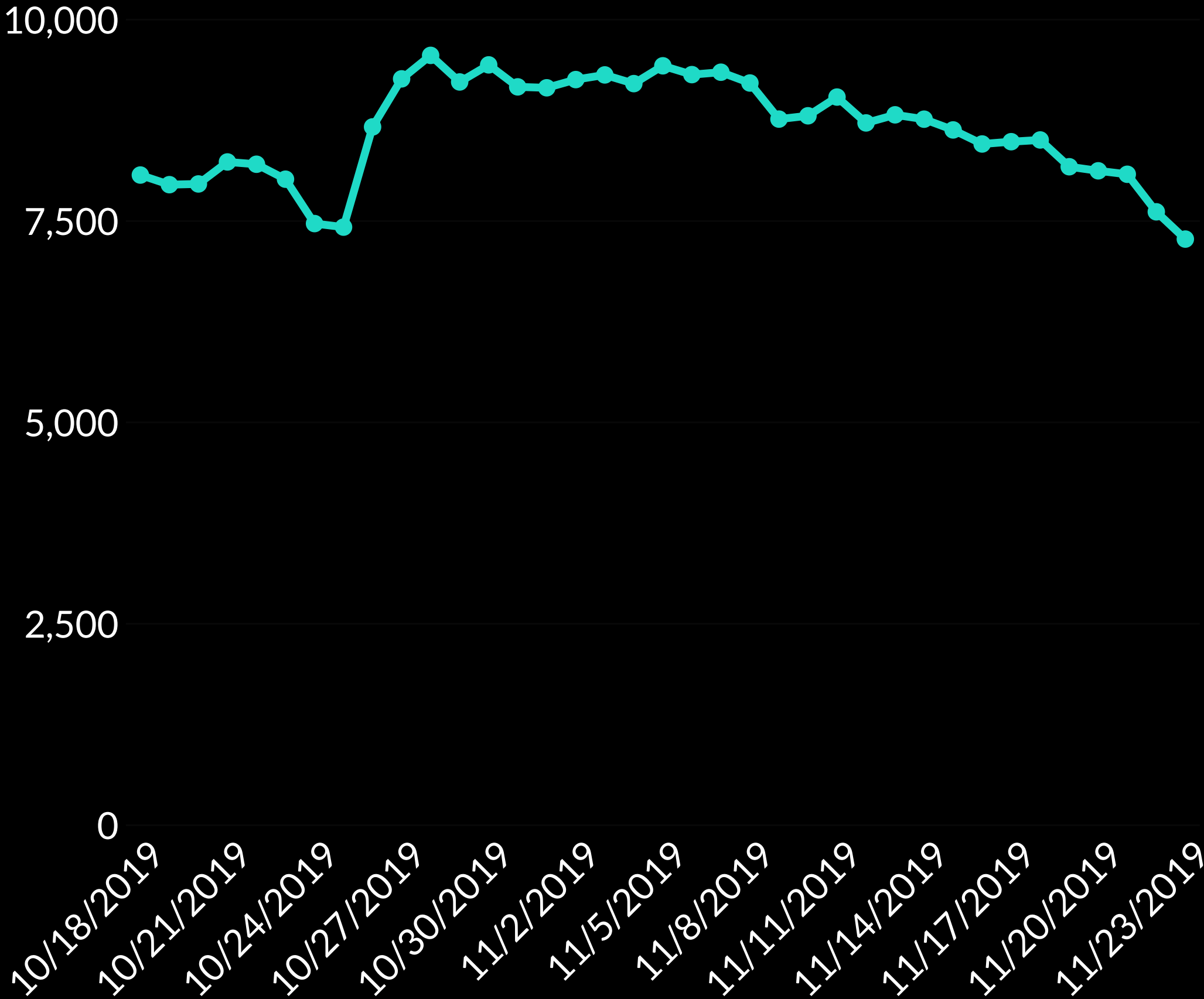


BITCOIN PRICE DATA

OUR CURRENT  
HISTORICAL BITCOIN  
PRICE DATA

*Daily Interval :1526 instances*  
*Hourly Interval :36506 instances*  
*5 Minutes Interval :438027 instances*  
*1 Minute Interval :2109031 instances*

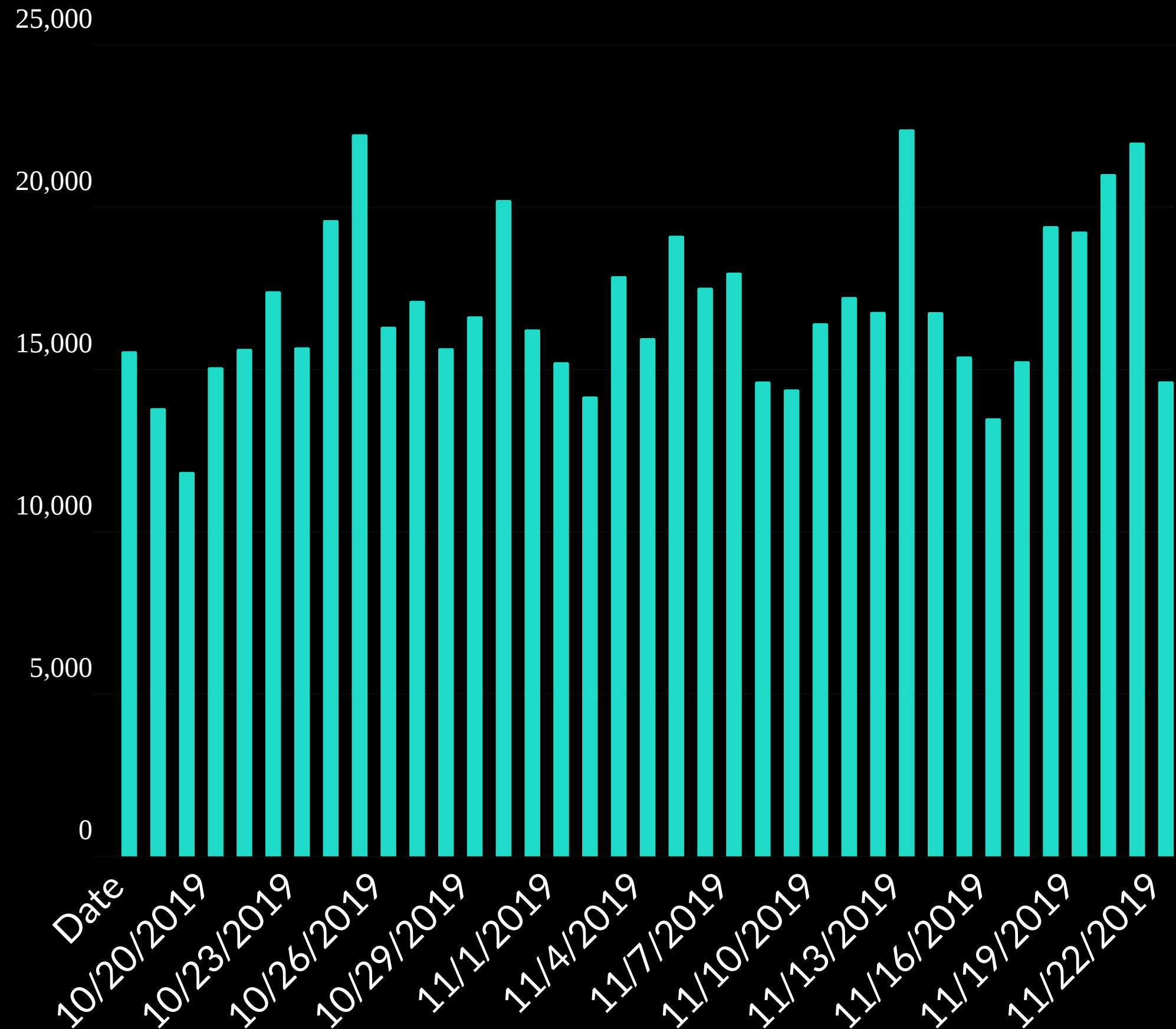
Historical price points for Bitcoin were gathered daily from Bitmex publicly available API.



OUR CURRENT  
TWEET NUMBERS

625231

To collect data for the sentiment analysis Twitter’s streaming API was used in combination with Tweepy. Tweepy is an open source framework which is written in Python, facilitates tweet collection from Twitter’s API.



# SENTIMENT ANALYSIS PROCESS

- Scrubbing bot generated content
- Sentiment analysis of individual tweets with VADER
- Aggregation of individual tweet sentiment score into a combined score for each time series interval.





# REDUCING NOISE IN THE TWITTER DATASET

tweets were scrubbed from bots and any non-alphabetic symbols (excluding "#" & "@").

## AGGREGATING SENTIMENT SCORES

With the sentiments returned by VADER, the individual tweet sentiment scores are grouped into time-series.

For each group the sentiment mean is taken on the underlying tweets to indicate the average sentiment.

## INDIVIDUAL TWEET SENTIMENT ANALYSIS

VADER is used to derive a sentiment score from each tweet. VADER provides a compound sentiment score between -1.0 and 1.0 for the text fed to it.



## RESULTS

Prediction performance

Changes in sentiments were modeled against changes in Bitcoin prices.

## FACTS AND FIGURES

# 83%

Bitcoin price fluctuation prediction accuracy based on Twitter sentiment stated in an article named Predicting Bitcoin price fluctuation accuracy with Twitter sentiment analysis

# MACHINE LEARNING ALGORITHMS

Accuracy Using Sentiment as Feature Vector

**95%**

NAIVE BAYES

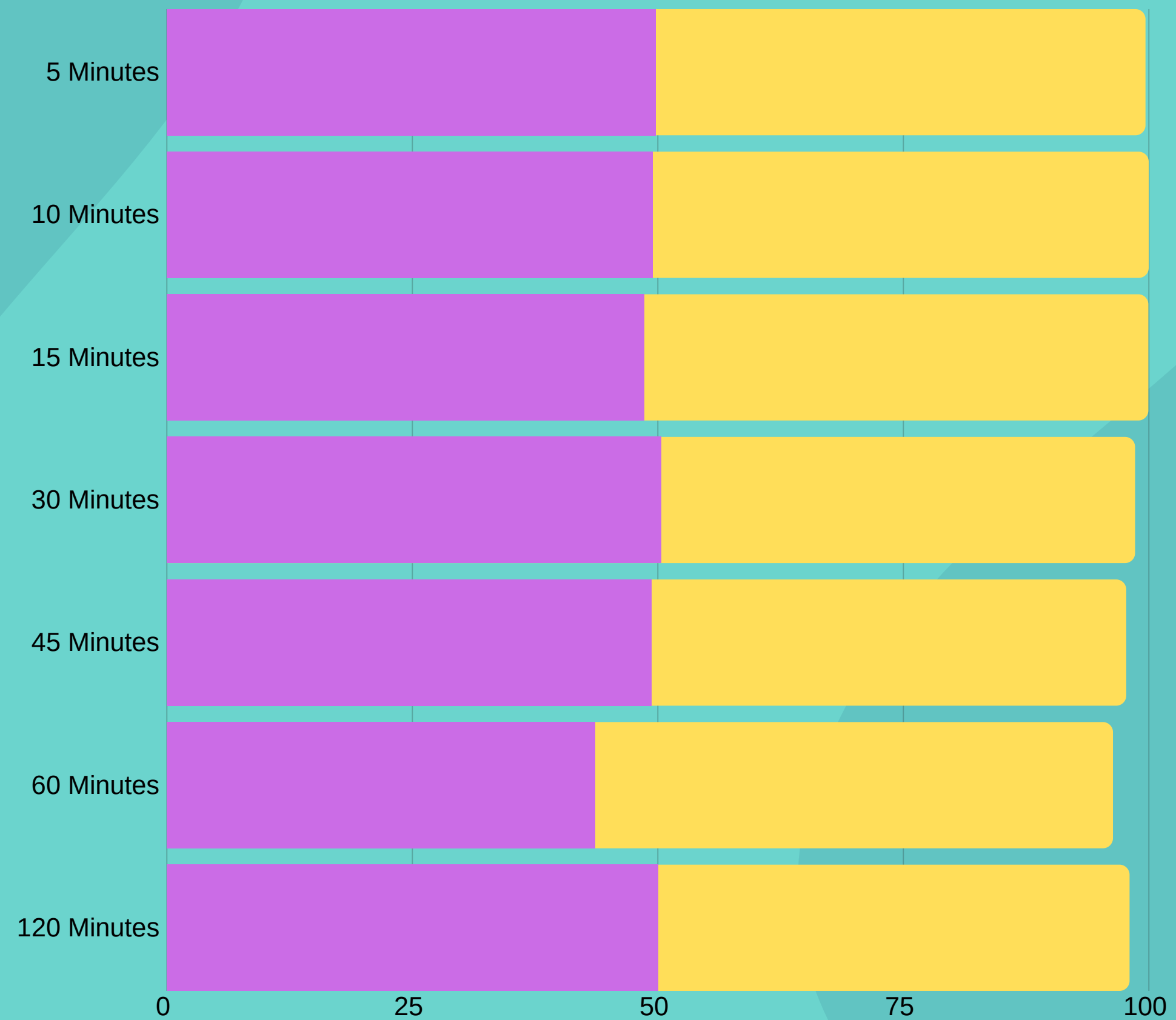
Accuracy Using Tweet as Feature Vector

**86%**

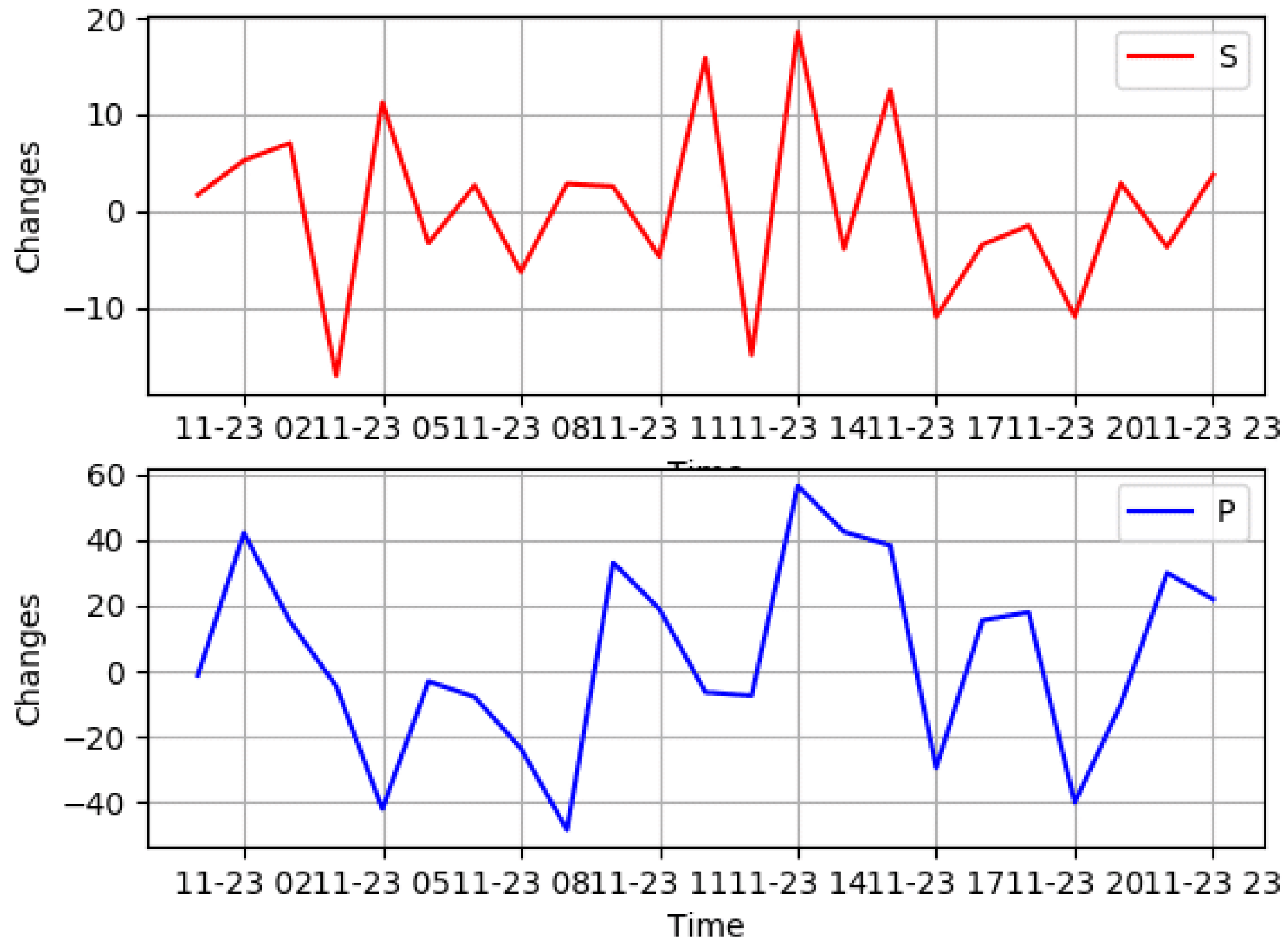
LOGISTIC REGRESSION



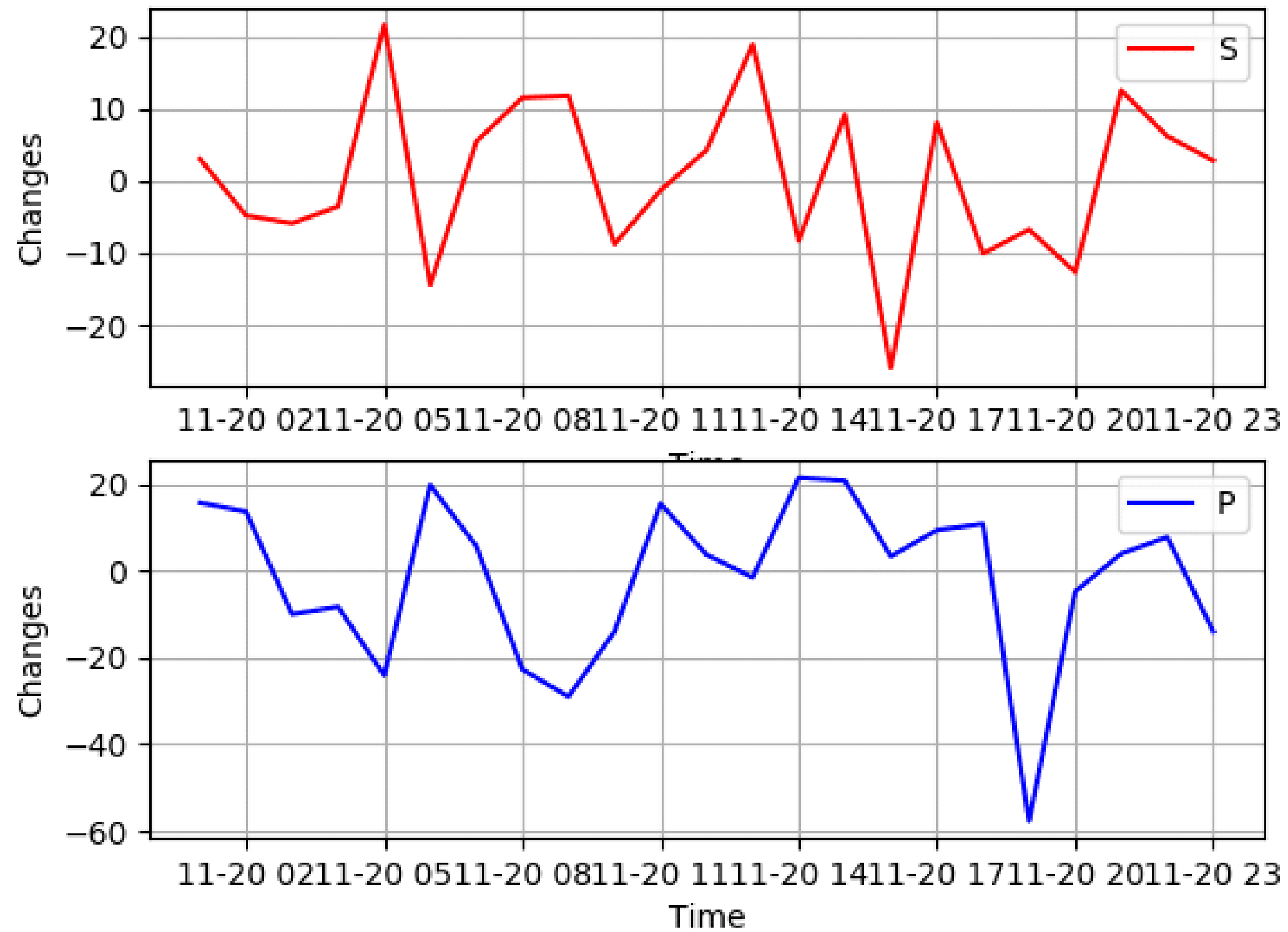
# 1 Frequency Shift Accuracy Vs 2 Frequency Shift Accuracy



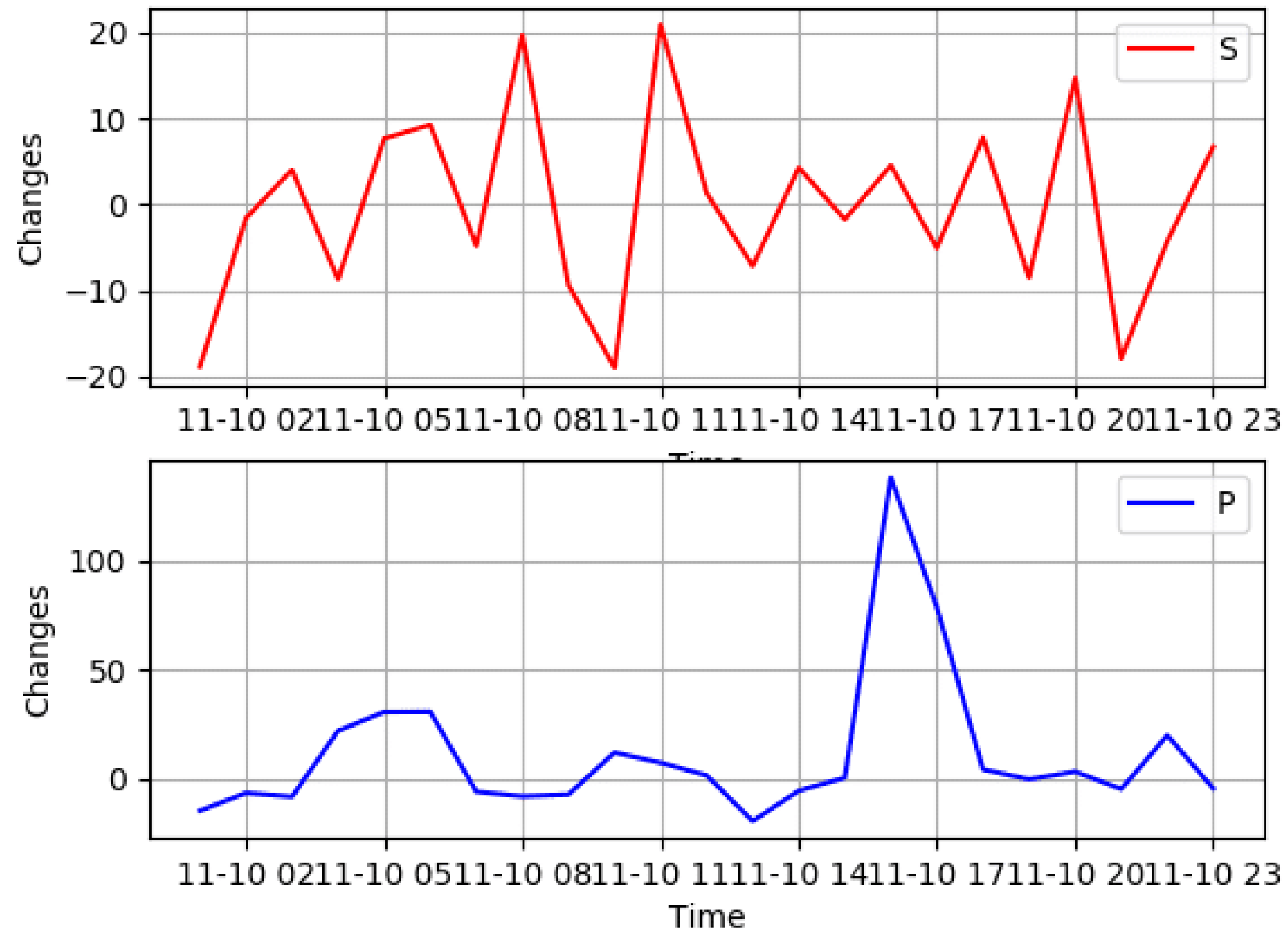
Sentiment~Price 2019-11-23



Sentiment~Price 2019-11-20

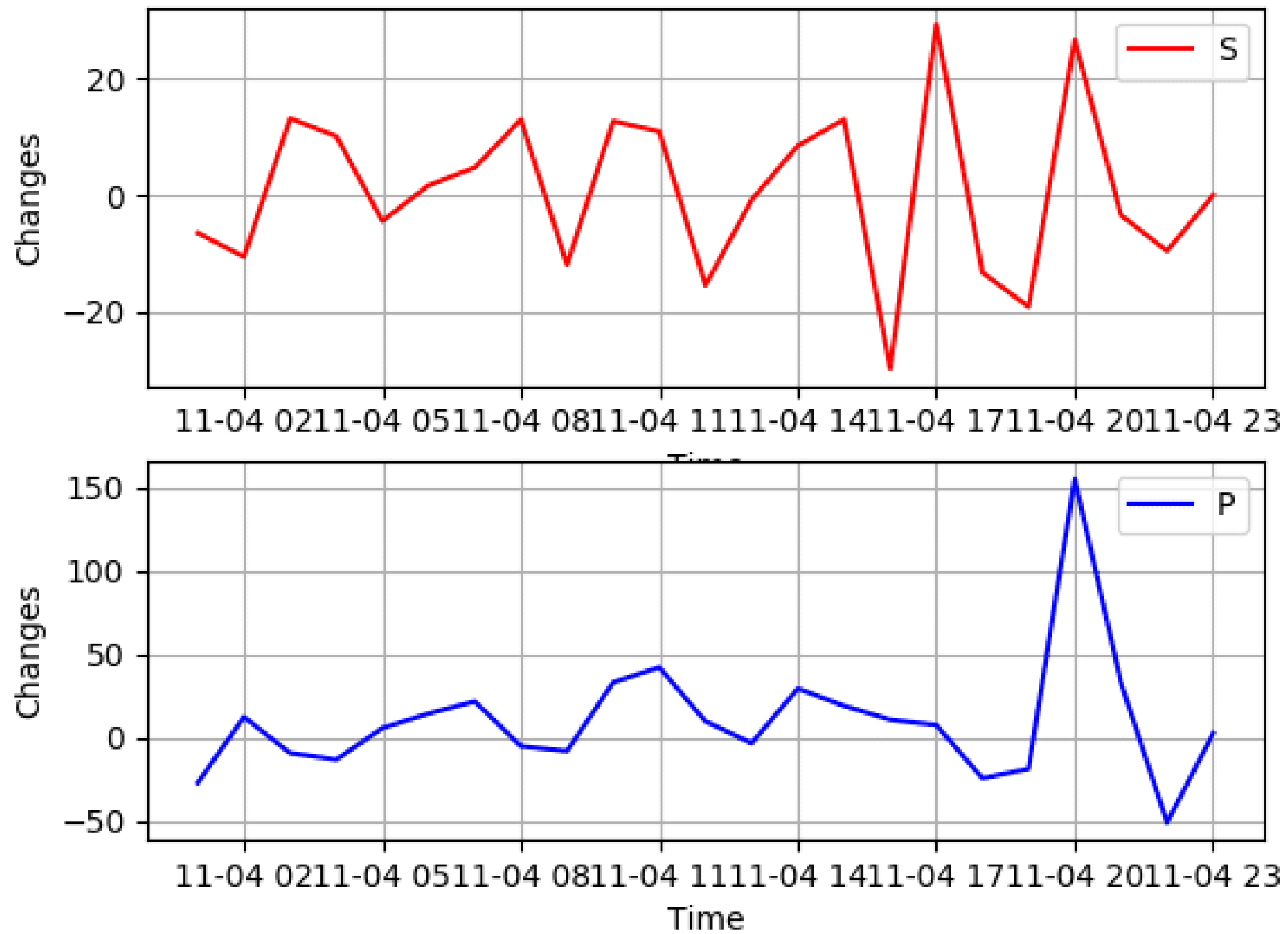


Sentiment~Price 2019-11-10

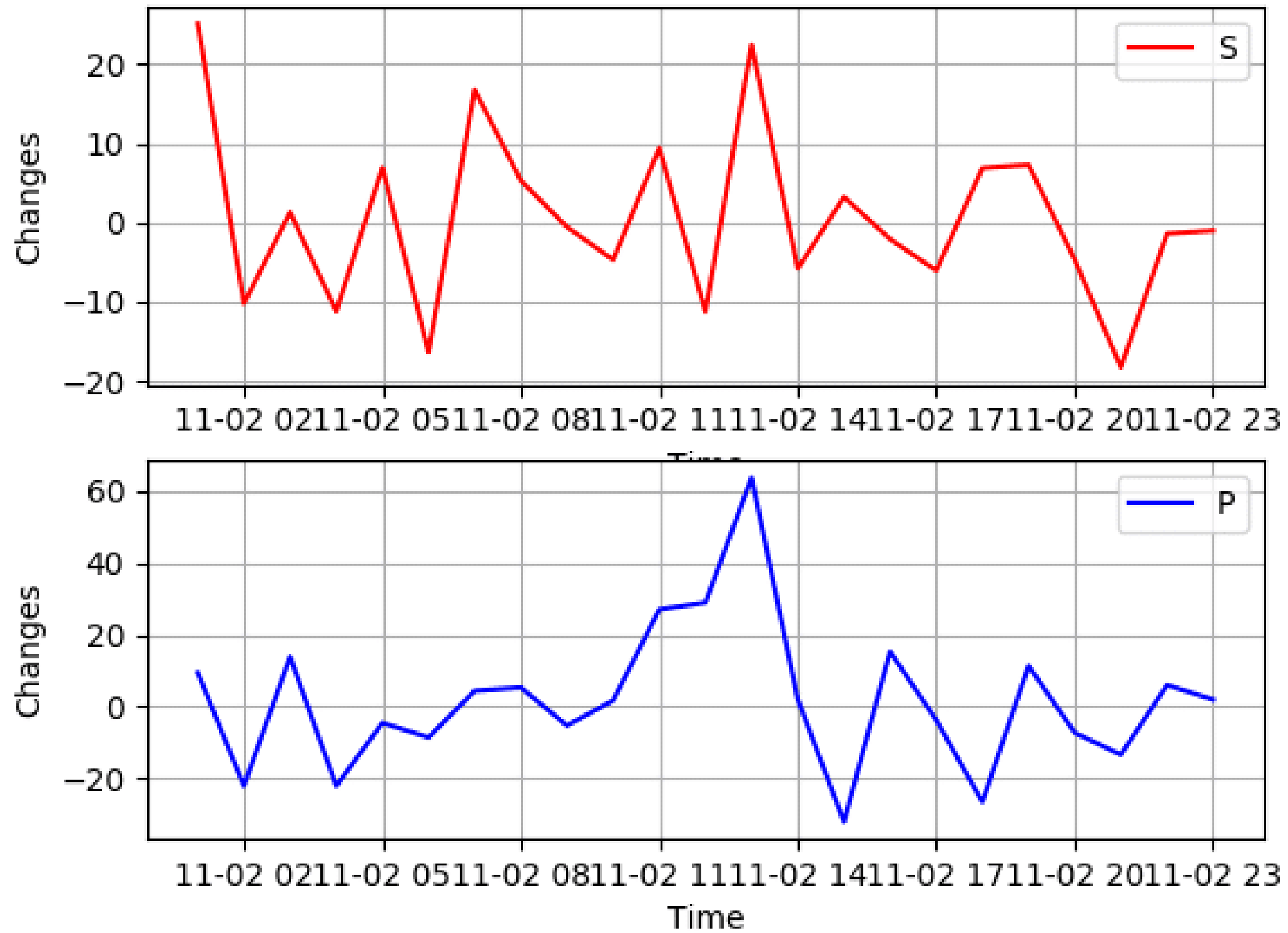




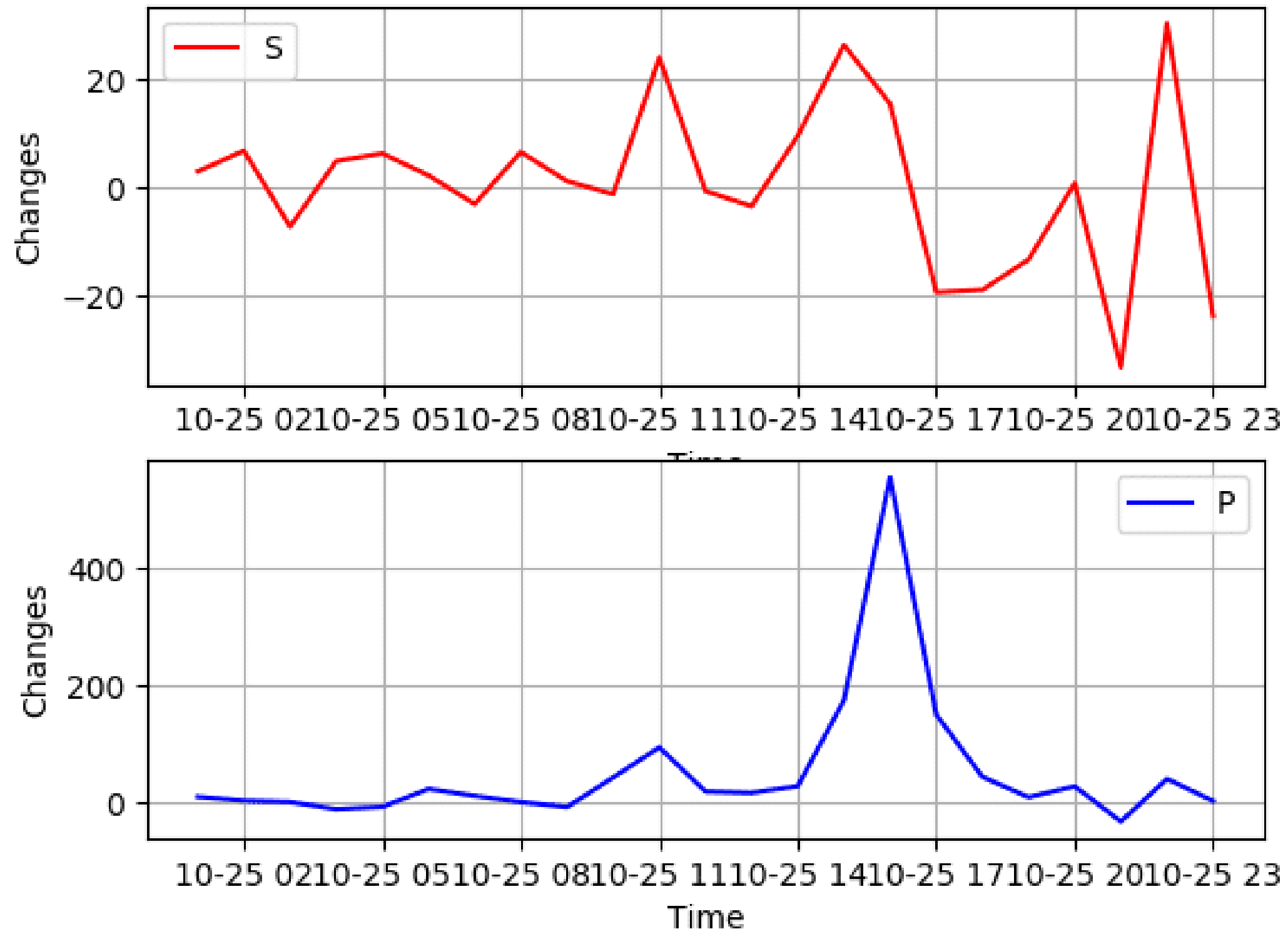
Sentiment~Price 2019-11-04



Sentiment~Price 2019-11-02



Sentiment~Price 2019-10-25





ROOZBEH  
DADASHZADEH  
RESEARCHER



BENYAMIN  
ABBASI  
LEAD RESEARCHER