

732A92 TEXT MINING

---

**CLASSIFYING STOCK PRICE MOVEMENTS  
BASED ON 8-K SEC FILINGS**

---

November 25, 2019

Name: Julius Kittler  
Student ID: julki092  
Linköping University

## Abstract

For over a century, fingerprints have been an undisputed personal identifier. Recent court rulings have sparked interest in verifying unique

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	SEC Filings . . . . .	3
1.2	8-K Filings . . . . .	3
1.3	Research Questions . . . . .	3
1.4	Relevance . . . . .	3
<b>2</b>	<b>Theory</b>	<b>3</b>
2.1	Stock Price Forecasting in General . . . . .	3
2.2	Stock Price Forecasting with Text Data . . . . .	3
2.3	State of the Art Classification with Text Data . . . . .	4
<b>3</b>	<b>Data</b>	<b>4</b>
3.1	Retrieval . . . . .	4
3.2	Processing . . . . .	4
3.3	Descriptive Statistics . . . . .	4
3.3.1	General . . . . .	4
3.3.2	Target Variable . . . . .	4
3.3.3	Feature Variables . . . . .	4
<b>4</b>	<b>Method</b>	<b>4</b>
4.1	Evaluation Metric . . . . .	4
4.2	Baseline Models . . . . .	4
4.3	Advanced Models . . . . .	4
4.4	Hyperparameter Tuning . . . . .	4
4.5	Feature Relevance . . . . .	4
<b>5</b>	<b>Results</b>	<b>4</b>
<b>6</b>	<b>Discussion</b>	<b>4</b>
<b>7</b>	<b>Conclusion</b>	<b>4</b>
<b>8</b>	<b>References</b>	<b>4</b>

# 1 Introduction

tbd

## 1.1 SEC Filings

The U.S. Securities and Exchange Commission (SEC) is a government agency in the United States with the mission to "protect investors, maintain fair, orderly, and efficient markets, and facilitate capital formation" [1]. An important task of the SEC is to ensure that publicly traded companies inform their shareholders and the public about their business.

For instance, the SEC requires companies to publish their quarterly and annual results, and inform shareholders about certain relevant events. For each of these purposes, companies have to file specific documents. For instance, the annual report corresponds to the 10-K, the quarterly report corresponds to the 10-Q and another report for specific relevant events corresponds to the 8-K filing.

Importantly, SEC filings are actively used by traders when making investment decisions. Many trading platforms such as Webull and thinkorswim also provide traders with the recent SEC filings of any tradable company (along with other information such as fundamental data, news data and historical prices). SEC filings are interesting for stock price forecasting because they are standardized, publicly accessible for free and because they contain relevant, objective and generally accurate information.

## 1.2 8-K Filings

Companies need to publish an 8-K filing for major events relevant for their business. Such events might be a change in the board of directors, a delisting from a stock exchange and a merger or acquisition [2]. Importantly, 8-K filings are generally due within four business days after the event [3].

Because 8-K filing correspond to major events for the company and because they need to be published shortly after an event occurred, 8-K filings seem interesting for predicting short-term volatility in the stock market. Moreover, the important information in 8-K filings is generally represented in form of text data, whereas other filings such as the annual report often focus on numerical data represented in tabular form.

**TODO: EXAMPLE**

## 1.3 Research Questions

tbd

1. Can we successfully forecast stock prices based on 8-K filings?
2. Which text features are most important when forecasting stock prices based on 8-K filings?

## 1.4 Relevance

tbd

# 2 Theory

Previous work in predicting stock prices with 8-K filings.

## 2.1 Stock Price Forecasting in General

tbd

## 2.2 Stock Price Forecasting with Text Data

tbd

## 2.3 State of the Art Classification with Text Data

tbd

## 3 Data

### 3.1 Retrieval

Data sources, Time period, Exchange Market, Resolution

### 3.2 Processing

### 3.3 Descriptive Statistics

#### 3.3.1 General

#### 3.3.2 Target Variable

#### 3.3.3 Feature Variables

## 4 Method

### 4.1 Evaluation Metric

### 4.2 Baseline Models

### 4.3 Advanced Models

### 4.4 Hyperparameter Tuning

Train vs. test

### 4.5 Feature Relevance

## 5 Results

		Predicted				
		lg_dec	sm_dec	zero	sm_inc	lg_inc
True	lg_dec	107	46	53	32	78
	sm_dec	39	56	117	49	48
	zero	34	54	146	46	33
	sm_inc	55	59	102	65	43
	lg_inc	82	42	41	40	128

## 6 Discussion

## 7 Conclusion

## 8 References

radioactivedecay2 DLS1

## References

- [1] *SEC.gov | What We Do*. [Online]. Available: <https://www.sec.gov/Article/whatwedo.html> (visited on 11/23/2019).
- [2] *SEC.gov | Form 8-K*. [Online]. Available: <https://www.sec.gov/fast-answers/answersform8khtml.html> (visited on 11/23/2019).
- [3] W. Kenton, *8-K (Form 8k)*, en. [Online]. Available: <https://www.investopedia.com/terms/1/8-k.asp> (visited on 11/23/2019).