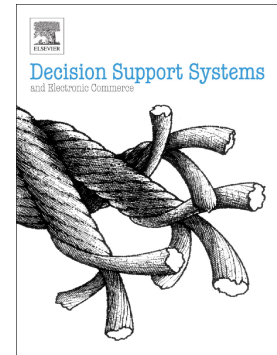


## Accepted Manuscript

Long-term stock index forecasting based on text mining of regulatory disclosures

Stefan Feuerriegel, Julius Gordon



PII: S0167-9236(18)30105-2  
DOI: doi:[10.1016/j.dss.2018.06.008](https://doi.org/10.1016/j.dss.2018.06.008)  
Reference: DECSUP 12966  
To appear in: *Decision Support Systems*  
Received date: 31 December 2017  
Revised date: 20 June 2018  
Accepted date: 21 June 2018

Please cite this article as: Stefan Feuerriegel, Julius Gordon , Long-term stock index forecasting based on text mining of regulatory disclosures. Decsup (2018), doi:[10.1016/j.dss.2018.06.008](https://doi.org/10.1016/j.dss.2018.06.008)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Long-term stock index forecasting based on text mining of regulatory disclosures

Stefan Feuerriegel<sup>a,\*</sup>, Julius Gordon<sup>b</sup>

<sup>a</sup>*ETH Zurich, Weinbergstr. 56/58, 8092 Zurich, Switzerland*

<sup>b</sup>*Chair for Information Systems Research, University of Freiburg, Platz der Alten Synagoge,  
79098 Freiburg, Germany*

---

## Abstract

Share valuations are known to adjust to new information entering the market, such as regulatory disclosures. We study whether the language of such news items can improve short-term and especially long-term (24 months) forecasts of stock indices. For this purpose, this work utilizes predictive models suited to high-dimensional data and specifically compares techniques for data-driven and knowledge-driven dimensionality reduction in order to avoid overfitting. Our experiments, based on 75,927 ad hoc announcements from 1996–2016, reveal the following results: in the long run, text-based models succeed in reducing forecast errors below baseline predictions from historic lags at a statistically significant level. Our research provides implications to business applications of decision-support in financial markets, especially given the growing prevalence of index ETFs (exchange traded funds).

*Keywords:* Text mining, Natural language processing, Financial news, Financial forecasting, Stock index, Predictive analytics

---

## 1. Introduction

The efficient market hypothesis formalizes how financial markets process and respond to new information [1]. Its semi-strong form states that asset prices fully

---

\*Corresponding author.

*Email addresses:* [sfeuerriegel@ethz.ch](mailto:sfeuerriegel@ethz.ch) (Stefan Feuerriegel), [juliusgordon89@gmail.com](mailto:juliusgordon89@gmail.com) (Julius Gordon)

Download English Version:

<https://daneshyari.com/en/article/6948329>

Download Persian Version:

<https://daneshyari.com/article/6948329>

[Daneshyari.com](https://daneshyari.com)