Classification of news headlines with impact on the probability of stock prices changes

Maximilian Franke, Sean Mondesire

St. Thomas University

Miami, FL, USA

mfranke@stu.edu, smondesire@stu.edu

*Abstract*—Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

# Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

# Background

Since there is a large number of potentially relevant sources of varying quality, the discussion in this paper of the identified problem will be based on an analysis of the literature which follows a standard framework for literature reviewing, developed by vom Brocke et al.[1] The procedure, which comprises 5 phases, is transformed into 3 major phases in terms of result orientation, focusing on identifying the research gap. The transformed procedure of literature analysis used in this thesis is summarized in the following figure.

Fig. 1: Transformed framework for literature reviewing

1. Phase 1

First, the scope of review is defined, in a structured and documented form. The requirement of a structured definition of review scope is fulfilled by the taxonomy according to Cooper.[2] This is divided into five properties: (1) focus; (2) goal; (3) organization; (4) target group; (5) framework. The focus defines the "type of sources", meaning that research results, research methods and applications of stock prediction models are included in the research. The aim is to integrate them in order to identify research gaps. The procedure in the research is to be regarded as conceptual. The target group is to be defined as the general professional audience, professionals, and practitioners. The framework of the literature analysis is complete and selective, in the sense of a complete research in German and English language.

Second, the conceptualization of the topic is performed via a concept map[1] with the topic’s stocks and model, in order to identify key terms and understand the subject matter. As a result, a list of keywords for the research is created, focusing on stocks and prediction which is summarized in the following search word list:

|  | Search word list | |
| --- | --- | --- |
| Related Stocks | Related method |
| 1 | Headlines / news | Classification |
| 2 | stocks | prediction |

1. Phase 2

The literature search is carried out in the publication databases Karlsruher Virtueller Katalog, ScienceDirect, and Web of Science in the stages defined by the search word list.[3] The results are verified or falsified in a first iteration by restricting the abstract, title and keyword list with respect to research relevance. In a second iteration, in the form of a full text analysis, the selected results are examined more closely and relevant articles for the classification of news headlines are included in the analysis so that a total of **X** relevant articles are identified.

A first structural analysis of the researched literature results shows that the number of relevant articles has increased in recent years. The high proportion of results from recent years in this research area is an indication of the novelty and relevance of the topic.

1. Phase 3

In this section, the results of the literature analysis are combined with the problem selection to generate a discussion of a research agenda, which is the starting point for further action.

# Methodology

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

# Results and discussion

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

##### Acknowledgment

##### References

Schmidt, C.W. “Metabolomics: what's happening downstream of DNA.” In: *Environmental Health Perspectives* 112.7 (2004). DOI: 10.1289/ehp.112-a410.

Rawat, W. and Wang, Z. “Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review”. In: *Neural Computation* 29.9 (2017). DOI: 10.1162/neco a 00990.