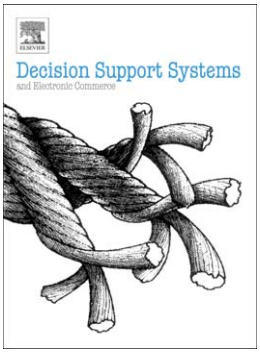
Accepted Manuscript



Sentiment Analysis in Financial Texts

Samuel W.K. Chan, Mickey W.C. Chong

PII: S0167-9236(16)30182-8
DOI: doi:10.1016/j.dss.2016.10.006
Reference: DECSUP 12783

To appear in: *Decision Support Systems*

Received date: 23 September 2015
Revised date: 20 September 2016
Accepted date: 31 October 2016

Please cite this article as: Samuel W.K. Chan, Mickey W.C. Chong, Sentiment Analysis in Financial Texts, *Decision Support Systems* (2016), doi:10.1016/j.dss.2016.10.006

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.

ACCEPTED MANUSCRIPT

Sentiment Analysis in Financial Texts

Samuel W.K. Chan and Mickey W.C. Chong
Department of Decision Sciences
The Chinese University of Hong Kong
Shatin, Hong Kong
{swkchan, mickey_chong}@cuhk.edu.hk

Abstract

The growth of financial texts in the wake of big data has challenged most organizations and brought escalating demands for analysis tools. In general, text streams are more challenging to handle than numeric data streams. Text streams are unstructured by nature, but they represent collective expressions that are of value in any financial decision. It can be both daunting and necessary to make sense of unstructured textual data. In this study, we address key questions related to the explosion of interest in how to extract insight from unstructured data and how to determine if such insight provides any hints concerning the trends of financial markets. A sentiment analysis engine (SAE) is proposed which takes advantage of linguistic analyses based on grammars. This engine extends sentiment analysis not only at the word token level, but also at the phrase level within each sentence. An assessment heuristic is applied to extract the collective expressions shown in the texts. Also, three evaluations are presented to assess the performance of the engine. First, several standard parsing evaluation metrics are applied on two treebanks. Second, a benchmark evaluation using a dataset of English movie review is conducted. Results show our SAE outperforms the traditional bag of words approach. Third, a financial text stream with twelve million words that aligns with a stock market index is examined. The evaluation results and their statistical significance provide strong evidence of a long persistence in the mood time series generated by the engine. In addition, our approach establishes grounds for belief that the sentiments expressed through text streams are helpful for analyzing the trends in a stock market index, although such sentiments and market indices are normally considered to be completely uncorrelated.

Keywords: Text analysis; financial time series; decision support systems.

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
✓ امکان دانلود نسخه ترجمه شده مقالات
✓ پذیرش سفارش ترجمه تخصصی
✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
✓ دانلود فوری مقاله پس از پرداخت آنلاین
✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات