

پیش بینی سهام با استفاده از الگوریتم بردار ماشین (SVM) با روش NLP با استفاده از

وب کاوی

بهزاد عزیزپور^۱، دکتر محمد علی بالافر^۲

^۱ دانشجوی ارشد مهندسی کامپیوتر - نرم افزار ، دانشگاه آزاد اسلامی واحد گرمی؛ iaugermi.ir@gmail.com
^۲ استادیار - فناوری اطلاعات IT، دانشگاه تبریز؛ balafarila@yahoo.com

چکیده

در این پایان‌نامه، الگوریتم‌های یادگیری ماشینی مورد استفاده در روش NLP به منظور دستیابی به مقصود اظهارنظرهای عمومی در مورد سهام‌های شرکت‌ها در رسانه‌های اجتماعی مورد تحلیل قرار گرفته‌اند تا ارتباط آن‌ها با تغییرات قیمت سهام مشخص شود. رویکرد NLP جهت شناسایی نیت‌ها از بیان اظهار نظرها یک رویکرد دو مرحله‌ای است که در مرحله اول خنثی یا جهت‌دار بودن اظهار نظر را تشخیص می‌دهد و در مرحله دوم مثبت یا منفی بودن آن را شناسایی می‌نماید و ثابت شده است که الگوریتم‌های SVM از جمله بهترین ابزار طبقه‌بندی برای مقاصد مذکور با نرخ دقت به ترتیب ۷۱/۸۴ و ۷۴/۳ هستند. مشخص شده است که فعالیت کاربران در استوک توئیت‌ها در طول شب دارای ارتباط مستقیم و قابل توجهی با حجم معاملات سهام روز کاری بعد دارد. اظهارنظرهای اجتماعی دارای قدرت بالایی برای پیش بینی و تغییر قیمت سهام در روز بعد در ۹ عدد از ۱۵ عدد سهام مطالعه شده توسط روش علی‌گرانگر بوده‌اند و نرخ دقت مجموع پیش بینی تغییرات رو به بالا و پائین سهام‌ها با استفاده از این روش برابر با ۵۸/۹٪ بوده است.

کلمات کلیدی

داده کاوی، وب کاوی، الگوریتم‌های یادگیری ماشینی، روش NLP، استوک توئیت‌ها، توئیت‌ر .

Using algorithms to predict stock vector machines (SVM) with NLP method using web mining

Behzad Azizpour, DR Mohammad Ali Balafar

ABSTRACT

In this thesis, machine learning algorithms are used in NLP to get the public sentiment on individual stocks from social media in order to study its relationship with the stock price change. The NLP approach of sentiment detection is a two-stage process by implementing Neutral v.s. Polarized sentiment detection before Positive v.s. Negative sentiment detection, and SVMs are proved to be the best classifiers with the overall accuracy rates of 71.84% and 74.3%, respectively. It is discovered that users' activity on StockTwits overnight significantly positively correlates to the stock trading volume the next business day. The collective sentiments for afterhours have powerful prediction on the change of stock price for the next day in 9 out of 15 stocks studied by using the Granger Causality test; and the overall accuracy rate of predicting the up and down movement of stocks by using the collective sentiments is 58.9%.

KEYWORDS

Data mining, Web mining, Algorithm SVM, Method NLP, Stocktwits, Twitter.

^۱ نویسنده: بهزاد عزیزپور به همراه نشانی: استان اردبیل، شهرستان گرمی به شماره موبایل: ۰۹۱۴۹۵۶۱۱۶۹