



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0100433

(43) 공개일자 2015년09월02일

(51) 국제특허분류(Int. Cl.)

G06F 19/00 (2011.01)

(21) 출원번호 10-2014-0022247

(22) 출원일자 2014년02월25일

심사청구일자 2014년02월25일

기술이전 희망 : 기술양도

(71) 출원인

창원대학교 산학협력단

경상남도 창원시 의창구 창원대학교 20(퇴촌동)

(72) 발명자

박태호

대전광역시 서구 도솔로245번길 46-10 (내동)

김중환

경상남도 창원시 마산회원구 석전동로 17 (석전동)

(뒷면에 계속)

(74) 대리인

김종선, 이형석

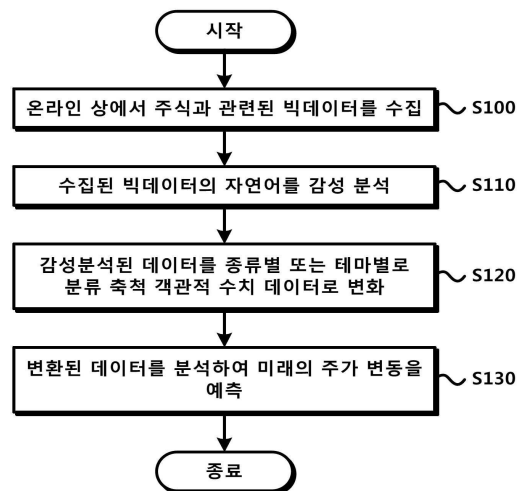
전체 청구항 수 : 총 5 항

(54) 발명의 명칭 빅데이터를 이용한 실시간 주식 예측 방법

(57) 요약

본 발명은 온라인 네트워크 상의 대규모 데이터를 효율적으로 수집 분석하여 더욱 정확히 주가 변동을 예측할 수 있도록 하는 빅데이터를 이용한 실시간 주식 예측 방법에 관한 것으로, 온라인 네트워크 상에서 주식과 관련된 빅데이터를 수집하고, 상기 수집된 빅데이터의 목표 단어를 감성 분석하고, 상기 감성 분석된 데이터를 종류별 또는 테마별로 분류 축척하고, 이를 객관적 수치 데이터로 변환하고, 상기 변환된 데이터를 분석하여 미래의 주가 변동을 예측하여서, 온라인 네트워크 상의 SNS를 통해 주가 관련 글을 수집하고 감성 분석을 통하여 더욱 정확한 주가 변동 예측이 가능하다.

대표도 - 도1



(72) 발명자

김지옥

경상남도 진주시 주약약골길75번길 18 (주약동)

박성재

경상북도 경산시 중앙초등길5길 8-8, 2동 508호 (중방동, 광명아파트)

신창욱

경상남도 김해시 가락로 301, 가동 306호 (구산동, 유신장미타워아파트)

특허청구의 범위

청구항 1

온라인 네트워크 상에서 주식과 관련된 빅데이터를 수집하는 수집단계;

상기 수집된 빅데이터의 목표 단어를 감성 분석하는 감성 분석단계;

상기 감성 분석된 데이터를 종류별 또는 테마별로 분류 축척하고, 이를 객관적 수치 데이터로 변환하는 데이터 변환단계; 및

상기 변환된 데이터를 분석하여 미래의 주가 변동을 예측하는 주가변동예측단계;를 포함하는 빅데이터를 이용한 실시간 주식 예측 방법.

청구항 2

제 1항에 있어서,

상기 빅데이터는 SNS 데이터를 포함하는 빅데이터를 이용한 실시간 주식 예측 방법.

청구항 3

제 2항에 있어서,

상기 SNS 데이터는 트위터, 블로그 및 페이스북 데이터중 적어도 어느 하나를 포함하는 빅데이터를 이용한 실시간 주식 예측 방법.

청구항 4

제 1항에 있어서,

상기 감성 분석단계는

상기 수집된 빅데이터에서 목표 단어를 일련의 체계에 따라 분석하여 긍정적 문장과 부정적 문장을 분류하는 단계를 포함하는 빅데이터를 이용한 실시간 주식 예측 방법.

청구항 5

제 1항에 있어서,

상기 주가변동예측단계는

국내외 및 해외 증시 변동 정보 및 과거 주가 변동내용을 수집하고, 같은 기간에 상승, 하락한 종목을 그룹화하고, 기업을 종목별, 테마별 그룹화한 데이터를 축척하여 주가변동을 예측하는 것을 특징으로 하는 빅데이터를 이용한 실시간 주식 예측 방법.

명세서

기술분야

[0001]

본 발명은 빅데이터를 이용한 실시간 주식 예측 방법에 관한 것으로, 더욱 상세하게는 온라인 네트워크 상의 대규모 데이터를 효율적으로 수집 분석하여 더욱 정확히 주가 변동을 예측할 수 있도록 하는 빅데이터를 이용한 실시간 주식 예측 방법에 관한 것이다.

배경 기술

- [0002] 최근 들어 인터넷이 활성화되어 감에 따라 인터넷 통신을 이용한 증권 정보를 제공하는 증권 정보 제공 사이트가 늘어가고 있다.
- [0003] 이러한 증권 정보 제공 사이트는 현재 증권 거래소의 시황관에 디스플레이되는 증권 정보를 인터넷 홈페이지를 통해 실시간으로 전송할 수 있도록 된 것으로, 사용자는 증권 거래소의 객장을 방문하기 않고도 실시간으로 증권 정보를 확인할 수 있으므로, 주식 투자의 자료로 활용할 수 있다.
- [0004] 이와 관련하여, 대한민국 공개특허공보 특2000-0054759호에는 사용자가 구입할 증권에 대해 예측한 정보를 제공하여 사용자가 구입할 증권을 용이하게 선택할 수 있도록 된 인터넷을 이용한 실시간 증권 정보 예측 프로그램이 개시되어 있다.
- [0005] 개시된 종래 기술의 인터넷을 이용한 실시간 증권 정보 예측 프로그램에 따르면, 증권 정보 제공 사이트가 로드된 서버와, 인터넷을 통해 상기 서버에 접속할 수 있도록 된 사용자의 인터넷 단말기를 포함하여 구성되며, 상기 서버에는 미리 입력된 시계열 분석을 통한 주식 시세 추론 방법에 의한 함수에 따라 모든 종목의 실시간 증권 정보를 예측하는 프로그램과 일정 개수의 권장 종목을 선정하는 실시간 증권 정보 예측 프로그램이 내장되며, 상기 증권 정보 제공 사이트에는 상기 모든 종목의 실시간 증권 정보를 예측하는 프로그램에 의한 정보를 디스플레이하는 표시부와 상기 실시간 증권 정보 예측 프로그램에 의해 선정된 권장 종목을 디스플레이하는 권장 종목 표시부가 배치되어, 상기 모든 종목의 실시간 증권 정보를 예측하는 프로그램에 의한 정보를 디스플레이하는 종목과 상기 실시간 증권 정보 예측 프로그램에 의해 선정된 권장 종목이 선정되면, 이 모든 종목과 권장 종목이 상기 증권 정보 제공 사이트에 실시간으로 디스플레이되어 고객은 미리 예측된 가격에 따라 주식을 투자할 수 있도록 한 것이다.
- [0006] 그러나, 이러한 종래의 증권 정보 제공 사이트는 어떠한 분석이나 추천 없이, 사용자에게 주가에 관한 정보만을 제공함으로 어떠한 주식을 어떠한 시점에 매매해야 하는지 정확한 판단을 하기 어려운 점이 있었다. 따라서, 주식 중개인이 몇 가지 종목을 추천하여 사용자가 추천 종목의 주식을 매매하도록 하는 방법이 개발되었으나, 이러한 시스템은 주식 동향에 따른 추천 종목을 실시간으로 판단하기 어려우며, 온라인 네트워크 상의 대규모 데이터를 효율적으로 운용하지 못하여 정확히 주식 변동을 예측할 수 없는 문제점이 있었다.

선행기술문헌

특허문헌

- [0007] (특허문헌 0001) 대한민국 공개특허공보 특2000-0054759호(2000년09월05일)

발명의 내용

해결하려는 과제

- [0008] 본 발명은 상술한 종래기술의 문제점을 극복하기 위한 것으로서, 온라인 네트워크 상의 대규모 데이터를 효율적으로 수집 분석하여 더욱 정확히 주가 변동을 예측할 수 있도록 하는 빅데이터를 이용한 실시간 주식 예측 방법을 제공하는 데에 그 목적이 있다.
- [0009] 또한, 온라인 네트워크 상의 SNS를 통해 주가 관련 글을 수집하고 감성 분석을 통하여 더욱 정확한 주가 변동 예측이 가능한 빅데이터를 이용한 실시간 주식 예측 방법을 제공하는 데에 그 목적이 있다.

과제의 해결 수단

- [0010] 상기 목적을 달성하기 위해 본 발명에 따른 빅데이터를 이용한 실시간 주식 예측 방법은, 온라인 네트워크 상에서 주식과 관련된 빅데이터를 수집하는 수집단계; 상기 수집된 빅데이터의 목표 단어를 감성 분석하는 감성 분

석단계; 상기 감성 분석된 데이터를 종류별 또는 테마별로 분류 축적하고, 이를 객관적 수치 데이터로 변환하는 데이터변환단계; 및 상기 변환된 데이터를 분석하여 미래의 주가 변동을 예측하는 주가변동예측단계;를 포함한다.

[0011]

상기 빅데이터는 SNS 데이터를 포함할 수 있다.

[0012]

상기 SNS 데이터는 트위터, 블로그 및 페이스북 데이터중 적어도 어느 하나를 포함하는 것이 바람직하다.

[0013]

상기 감성 분석단계는, 상기 수집된 빅데이터에서 목표 단어를 일련의 체계에 따라 분석하여 긍정적 문장과 부정적 문장을 분류하는 단계를 포함한다.

[0014]

또한, 상기 주가변동예측단계는, 국내외 및 해외 증시 변동 정보 및 과거 주가 변동내용을 수집하고, 같은 기간에 상승, 하락한 종목을 그룹화하고, 기업을 종목별, 테마별 그룹화한 데이터를 축적하여 주가변동을 예측하는 것이 더욱 바람직하다.

발명의 효과

[0015]

상기와 같이 구성된 본 발명에 따른 빅데이터를 이용한 실시간 주식 예측 방법에 의하면, 온라인 네트워크 상의 대규모 데이터를 효율적으로 수집 분석하여 더욱 정확히 주가 변동을 예측할 수 있는 효과가 있다.

[0016]

또한, 온라인 네트워크 상의 SNS를 통해 주가 관련 글을 수집하고 감성 분석을 통하여 더욱 정확한 주가 변동 예측이 가능한 효과도 있다.

도면의 간단한 설명

[0017]

도 1은 본 발명의 바람직한 일 실시예에 따른 빅데이터를 이용한 실시간 주식 예측 방법을 나타낸 순서도.

도 2는 본 발명의 바람직한 일 실시예에 따른 빅데이터를 이용한 실시간 주식 예측 방법을 구현하기 위한 네트워크 시스템 구성도이다.

발명을 실시하기 위한 구체적인 내용

[0018]

본 발명은 다양한 변형 및 여러 가지 실시예를 가질 수 있는바, 특정 실시예들을 도면에 예시하고 상세한 설명에 보다 상세하게 설명하고자 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0019]

이하, 본 발명의 바람직한 실시예를 첨부된 도면을 참조하여 상세히 설명하기로 한다.

[0020]

도 1은 본 발명의 바람직한 일 실시예에 따른 빅데이터를 이용한 실시간 주식 예측 방법을 나타낸 순서도이다.

[0021]

도 1에 도시된 바와 같이 본 발명에 따른 빅데이터를 이용한 실시간 주식 예측 방법은, 온라인 네트워크 상에서 주식과 관련된 빅데이터를 수집하는 수집단계(S100); 상기 수집된 빅데이터의 목표 단어를 감성 분석하는 감성 분석단계(S110); 상기 감성 분석된 데이터를 종류별 또는 테마별로 분류 축적하고, 이를 객관적 수치 데이터로 변환하는 데이터변환단계(S120); 및 상기 변환된 데이터를 분석하여 미래의 주가 변동을 예측하는 주가변동예측단계(S130)를 포함한다.

[0022]

상기 빅데이터는 SNS 데이터를 포함할 수 있다. 상기 SNS 데이터는 트위터, 블로그 및 페이스북 데이터중 적어도 어느 하나를 포함하는 것이 바람직하다.

[0023]

상기 감성 분석단계(S110)는, 상기 수집된 빅데이터에서 목표 단어를 일련의 체계에 따라 분석하여 긍정적 문장과 부정적 문장을 분류하는 단계를 포함한다.

[0024]

또한, 상기 주가변동예측단계(S130)는, 국내외 및 해외 증시 변동 정보 및 과거 주가 변동내용을 수집하고, 같은 기간에 상승, 하락한 종목을 그룹화하고, 기업을 종목별, 테마별 그룹화한 데이터를 축적하여 주가변동을 예측하는 것이 더욱 바람직하다.

[0025]

도 2는 본 발명의 바람직한 일 실시예에 따른 빅데이터를 이용한 실시간 주식 예측 방법을 구현하기 위한 네트워크 시스템 구성도로서, 도면부호 10은 SNS 데이터, 20은 인터넷, 30은 증권 정보 서버, 40은 운영 서버를 나

타낸다.

- [0026] 상기 SNS 데이터(10)는 온라인 네트워크, 특히 인터넷 상에서 수집할 수 있는 트위터, 블로그 및 페이스북 데이터중 적어도 어느 하나를 포함하는 것이 바람직하다.
- [0027] 증권 정보 서버(30)는 통상의 국내외 및 해외 증시 변동 정보 및 과거 주가 변동내용을 갖고 있으며, 특정 기간에 상승, 하락한 종목과, 특정 기업을 종목별, 테마별 그룹화한 데이터를 가지고 있다.
- [0028] 또한, 상장사 정보는 KRX(한국거래소)에서 참고할 수 있고, 국내?외 및 해외 증시 변동 정보는 코스피, 코스닥, 다우, 홍콩, 나스닥 등을 참조하여 수집할 수 있다.
- [0029] 운영서버(40)는 본 발명의 전반적인 동작을 제어 및 운영하는 서버로서, 온라인 네트워크 상에서 주식과 관련된 빅데이터를 수집하고, 수집된 빅데이터의 목표 단어를 감성 분석하며, 상기 감성 분석된 데이터를 종류별 또는 테마별로 분류 축척하여 이를 객관적 수치 데이터로 변환하고, 변환된 데이터를 분석하여 미래의 주가 변동을 예측하는 일련의 제어 과정을 수행한다.
- [0030] 최근, 디지털 경제의 확산으로 우리 주변에는 규모를 가늠할 수 없을 정도로 많은 정보와 데이터가 생산되는 '빅데이터(big data)' 환경이 도래하고 있다. 빅데이터란 과거 아날로그 환경에서 생성되던 데이터에 비하면 그 규모가 방대하고, 생성 주기도 짧고, 형태도 수치 데이터뿐 아니라 문자와 영상 데이터를 포함하는 대규모 데이터를 말한다.
- [0031] PC와 인터넷, 모바일 기기 이용이 생활화되면서 사람들이 도처에 남긴 발자국(데이터)은 기하급수적으로 증가하고 있다. 쇼핑의 예를 들어 보자. 데이터의 관점에서 보면 과거에는 상점에서 물건을 살 때만 데이터가 기록되었다. 반면 인터넷쇼핑몰의 경우에는 구매를 하지 않더라도 방문자가 돌아다니는 기록이 자동적으로 데이터로 저장된다. 어떤 상품에 관심이 있는지, 얼마 동안 쇼핑몰에 머물렀는지를 알 수 있다. 쇼핑뿐 아니라 은행, 증권과 같은 금융거래, 교육과 학습, 여가활동, 자료검색과 이메일 등 하루 대부분의 시간을 PC와 인터넷에 할애한다. 사람과 기계, 기계와 기계가 서로 정보를 주고받는 사물지능통신(M2M, Machine to Machine)의 확산도 디지털 정보가 폭발적으로 증가하게 되는 이유다.
- [0032] 사용자가 직접 제작하는 UCC를 비롯한 동영상 콘텐츠, 휴대전화와 SNS(Social Network Service)에서 생성되는 문자 등은 데이터의 증가 속도뿐 아니라, 형태와 질에서도 기존과 다른 양상을 보이고 있다. 특히 블로그나 SNS에서 유통되는 텍스트 정보는 내용을 통해 글을 쓴 사람의 성향뿐 아니라, 소통하는 상대방의 연결 관계까지도 분석이 가능하다. 게다가 사진이나 동영상 콘텐츠를 PC를 통해 이용하는 것은 이미 일반화되었고 방송 프로그램도 TV수상기를 통하지 않고 PC나 스마트폰으로 보는 세상이다.
- [0033] 예를 들면, 트위터(tweeter)에서만 하루 평균 1억 5500만 건이 생겨나고 유튜브(YouTube)의 하루 평균 동영상 재생건수는 40억 회에 이른다. 글로벌 데이터 규모는 2012년에 2.7제타바이트(zettabyte), 2015년에는 7.9제타바이트로 증가할 것으로 예측하고 있다(IDC, 2011). 1제타바이트는 1000엑사바이트(exabyte)이고, 1엑사바이트는 미 의회도서관 인쇄물의 10만 배에 해당하는 정보량이다(Lynman, P., & Varian, H., 2003).
- [0034] 본 발명은 상기와 같은 빅데이터를 이용하여, 특히 SNS를 통해 주식에 관련된 생각과 의견을 분석하고 이를 통하여 주가를 예측할 수 있도록 한다.
- [0035] 이를 위하여, 먼저 본 발명은 도 1에 도시된 바와 같이, 온라인 네트워크 상에서 주식 또는 주식 상장사 등과 관련된 빅데이터를 수집하는 수집단계를 수행한다(S100).
- [0036] 빅데이터의 수집은 본 발명의 운영서버(40)에서 수행될 수 있다.
- [0037] 이때, 상기 빅데이터는 SNS 데이터(10)를 포함할 수 있으며, 상기 SNS 데이터(10)는 온라인 네트워크, 특히 인터넷 상에서 수집할 수 있는 트위터, 블로그 및 페이스북 데이터중 적어도 어느 하나를 포함하는 것이 바람직하다.
- [0038] 이후, 상기 수집된 빅데이터의 목표 단어를 감성 분석하는 감성 분석단계를 수행한다(S110).
- [0039] 여기서, 상기 감성 분석단계(S110)는, 상기 수집된 빅데이터에서 목표 단어를 일련의 체계에 따라 분석하여 긍정적 문장과 부정적 문장을 분류하는 단계를 포함한다.
- [0040] 즉, 본 발명에서 말하는 감성 분석은 문장내에서 목표가 되는 단어에 대해 긍정적인지 부정적인지 구분하는 언어처리 기법으로서, SNS를 통해 많은 정보가 생성되고, 신뢰성 있는 주가 관련 글이 축적됨에 따라 이를 분석하

여 주가 예측 실행하기 위한 것이다.

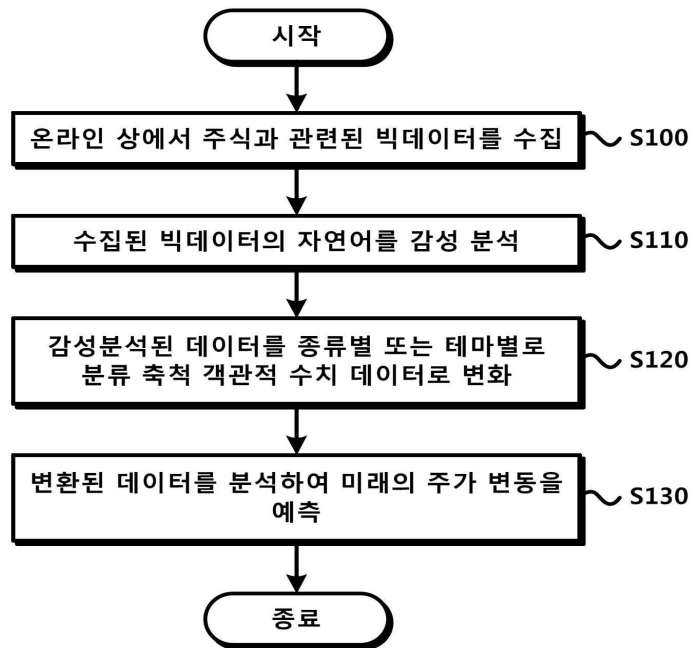
- [0041] 예로서, 트위터 상의 주식 관련 글을 수집하여 감성 분석을 실행하고, 감성 분석된 자료를 바탕으로 주가 변동 예측할 수 있게 되는 것이다.
- [0042] 이후, 상기 감성 분석된 데이터를 종류별 또는 테마별로 분류 축적하고, 이를 객관적 수치 데이터로 변환하는 데이터변환단계를 수행한다(S120).
- [0043] 여기서, 증권 정보 서버(30)는 통상의 국내외 및 해외 증시 변동 정보 및 과거 주가 변동내용과, 특정 기간에 상승, 하락한 종목과, 특정 기업을 종목별, 테마별 그룹화한 데이터를 제공하고, 운영서버(40)는 상기 데이터를 제공받아 상기 감성 분석된 데이터를 종류별 또는 테마별로 분류 축적하여 이를 객관적 수치 데이터로 변환하고, 변환된 데이터를 분석하여 미래의 주가 변동을 예측하는 일련의 제어 과정을 수행할 수 있다.
- [0044] 이후, 상기 변환된 데이터를 분석하여 미래의 주가 변동을 예측하는 주가변동예측단계를 수행한다(S130).
- [0045] 상기 주가변동예측단계(S130)에서는, 국내외 및 해외 증시 변동 정보 및 과거 주가 변동내용을 수집하고, 같은 기간에 상승, 하락한 종목을 그룹화하고, 기업을 종목별, 테마별 그룹화한 데이터를 축적하여 주가변동을 예측하는 것이 더욱 바람직하다.
- [0046] 따라서, 본 발명에 의하면, 주식을 예측함에 있어 감성 분석과 한국 및 해외 증시의 변동률을 측정하며, 종목과 테마별 그룹에 따라 예측 모델의 데이터로 삼아 사용자에게 주가 상승이 예측되는 종목을 추천할 수 있다.
- [0047] 본 명세서에 기재된 본 발명의 실시예와 도면에 도시된 구성은 본 발명의 가장 바람직한 실시예에 관한 것이고, 발명의 기술적 사상을 모두 포괄하는 것은 아니므로, 출원시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다. 따라서 본 발명은 상술한 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자라면 누구든지 다양한 변형실시가 가능한 것은 물론이고, 그와 같은 변경은 청구범위 기재의 권리범위 내에 있게 된다.

부호의 설명

- [0048] 10 : SNS 데이터
- 20 : 인터넷
- 30 : 증권 정보 서버
- 40 : 운영 서버

도면

도면1



도면2

