

< 주식시장 기상청: 오늘의 날씨는? - 개발기획서 >

<빅데이터 부문>

팀 한탕

고려대학교 통계학과 강현규

고려대학교 통계학과 이충은

고려대학교 통계학과 장희중

- 목차 -

1. 초록
2. 개요
3. 서비스 및 디자인 설명
 - 3.1. 오늘의 날씨
 - 3.1.1. 서비스 설명
 - 3.1.2. 사용 데이터
 - 3.1.3 메인 알고리즘 및 Flow chart
 - 3.1.4 디자인 및 기대효과
 - 3.2. 일기 예보
 - 3.2.1. 서비스 설명
 - 3.2.2. 사용 데이터
 - 3.2.3. 메인 알고리즘 및 Flow chart
 - 3.2.4. 모델 설명
 - 3.2.5. 디자인 및 기대효과
 - 3.3. 뉴스 해시태그
 - 3.3.1. 서비스 설명
 - 3.3.2. 사용 데이터
 - 3.3.3. 메인 알고리즘 및 Flow chart
 - 3.3.4. 모델 설명
 - 3.3.5. 디자인 및 기대효과
 - 3.4. MY 날씨
4. 추가 활용방안
 - 4.1. 이상기후 탐지 경보
 - 4.2. 개인별 예보시스템
5. 참고문헌

1. 초록

모바일거래시스템(Mobile trading system, MTS)의 영역에서 성공적인 금융 투자를 위해서는 데이터 기반 의사 결정이 필수적입니다. 이에, 본 연구는 투자자들에게 실시간 주식 시장에 대한 직관적이고 이해하기 쉬운 정보를 제공하기 위해 "주식시장 기상청"이라는 새로운 서비스를 제시합니다.

여기에는 실시간으로 변화하는 주식시장 데이터를 현재 날씨로 변환하는 "오늘의 날씨", 내일의 날씨를 예측하는 "주식 일기예보", 시장과 관련된 현재의 뉴스 트렌드를 해시태그로 보여주는 "뉴스 해시태그", 개인이 보유한 주식 종목 별 날씨를 보여주는 "MY 날씨" 서비스가 포함됩니다. 서비스를 활용함으로써 "주식시장 기상청"은 투자자들이 방대한 양의 금융 정보를 직관적으로 받아들이고, 제공받은 인사이트를 통해 합리적인 의사결정을 하도록 지원합니다.

본 보고서는 "주식시장 기상청" 서비스의 개념과 구현 방법을 설명하며, 향후 미래에 증권 앱 "M.STOCK"에서의 활용 가능성을 다룹니다.

2. 개요

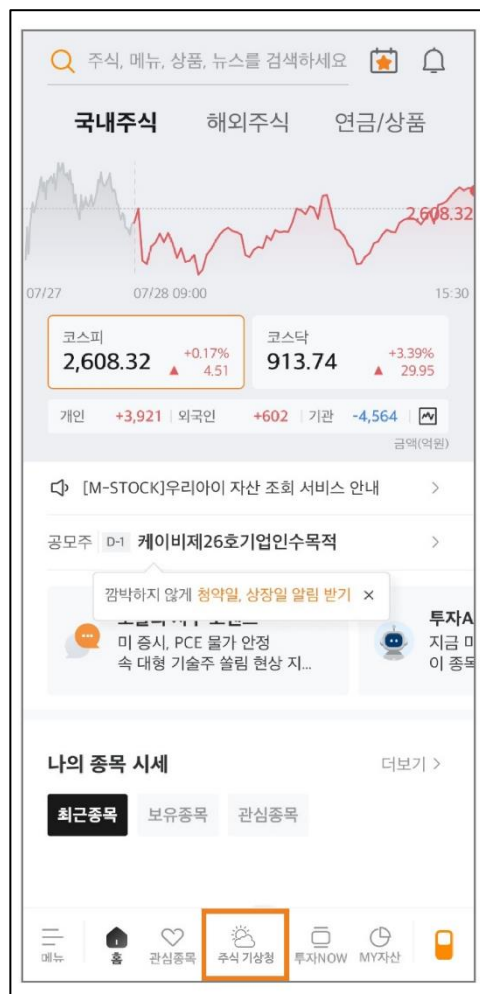
한국소비자원, 증권 앱 서비스 소비자 만족도 및 이용실태 조사(2021)에 의하면, 최근 주식시장에 대한 관심이 높아지며 모바일거래시스템(Mobile trading system, MTS)을 활용해 주식을 거래하는 소비자들의 비율이 크게 증가했습니다.

대부분의 증권 회사들은 투자자들을 자사의 고객으로 유치하기 위해 다양한 방식으로 앱을 발전시키고 있습니다. 기존 사용자들은 회원가입 및 계좌개설의 복잡성으로 인해 초기에 결정한 증권사를 잘 바꾸지 않는 특성이 있습니다. 그렇기 때문에 신규 고객을 확보하는 것이 매우 중요하며, 이를 위해 증권 앱은 신규 투자자들에게 적합한 서비스를 제공할 수 있어야 합니다.

금융 및 투자의 세계에서 사용 가능한 정보의 양은 종종 압도적으로 느껴질 수 있으며, 특히 이러한 데이터를 완전히 분석할 수 있는 도구나 전문 지식이 부족한 신규 개인 투자자들에게는 더욱 그렇습니다. 이들이 꾸준히 증권 앱을 활용해 주식을 거래하고, 투자에 흥미를 갖도록 하기 위해서는 시장지수, 거래량 등과 같이 어려운 용어와 숫자로만 이루어진 화면이 아닌, 보다 직관적으로 시장에 대한 정보를 줄 수 있는 UI/UX가 필요합니다.

또한 소비자24에서 2021년 증권 앱 사용자 1800명을 대상으로 한 설문조사 결과, 실시간 기업 정보 및 뉴스 정보를 제공하는 서비스의 만족도가 가장 높은 것으로 나타났습니다. 실시간으로 많은 정보들이 업데이트 되는 현대 사회에서 어떤 일들이 있었는지 하나하나 찾아보는 것은 투자자 입장에서 많은 시간과 비용이 필요한 작업이기에, 이는 당연한 결과입니다.

이러한 상황들을 종합적으로 고려했을 때, 실시간 주식시장의 정보들을 명확하고 직관적으로 제공하는 서비스를 신규 고객에게 제공한다면 매우 만족도가 높을 것으로 판단했습니다. 날씨는 모두에게 친숙하며 나타내는 바가 분명합니다. 따라서 "기상청"이라는 큰 틀 내에서 "일기예보" ("오늘의 날씨" 및 "My 날씨")를 활용해 현재 주식시장과 본인이 투자한 주식의 주가 상황을, "일기예보"를 통해 내일 날씨의 예측을, "뉴스 해시태그"를 활용해 실시간 뉴스 트렌드를 제시하는 서비스를 고안했습니다.



왼쪽에 제시된 그림과 같이 “M.STOCK” 탭에 새로운 서비스인 “주식 기상청”을 추가하고, 들어가면 소비자에게 오른쪽 그림과 같은 정보들을 제공하는 방식입니다.

3. 서비스 및 디자인 설명

3.1. 오늘의 날씨

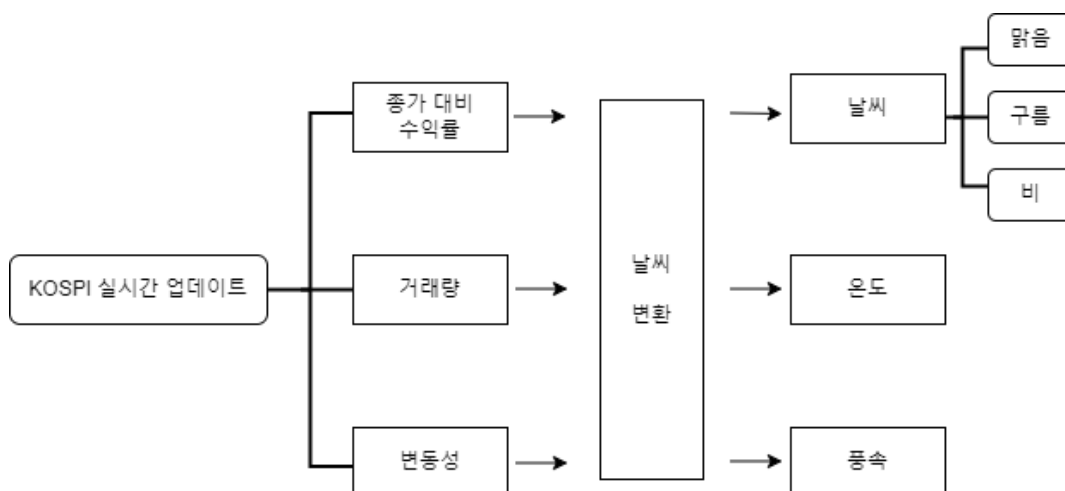
3.1.1. 서비스 설명

지수의 수익률, 거래량, 변동성은 주식 시장의 흐름을 나타내는 매우 중요한 지표입니다. 하지만 경험이 많지 않은 개인 투자자들의 경우 위와 같은 지표가 무엇을 의미하는지 파악하기 어렵습니다. “오늘의 날씨”는 위 지표들을 날씨 형태로 변환함으로써, 투자자들이 지표들의 관계와 현재 시장의 상황을 직관적으로 이해할 수 있도록 하는 것을 목표로 합니다.

3.1.2. 사용 데이터

- 실시간 KOSPI 지수, 거래량 및 변동성 지수(VKOSPI): 네이버 금융, Investing.com 크롤링
- 지표들의 과거 데이터: Python “yfinance”패키지

3.1.3. 메인 알고리즘 및 Flow chart



1) 데이터셋 업데이트: 크롤링을 통해 실시간 KOSPI지수, 거래량, 변동성 지수 업데이트

2) 날씨로 변환

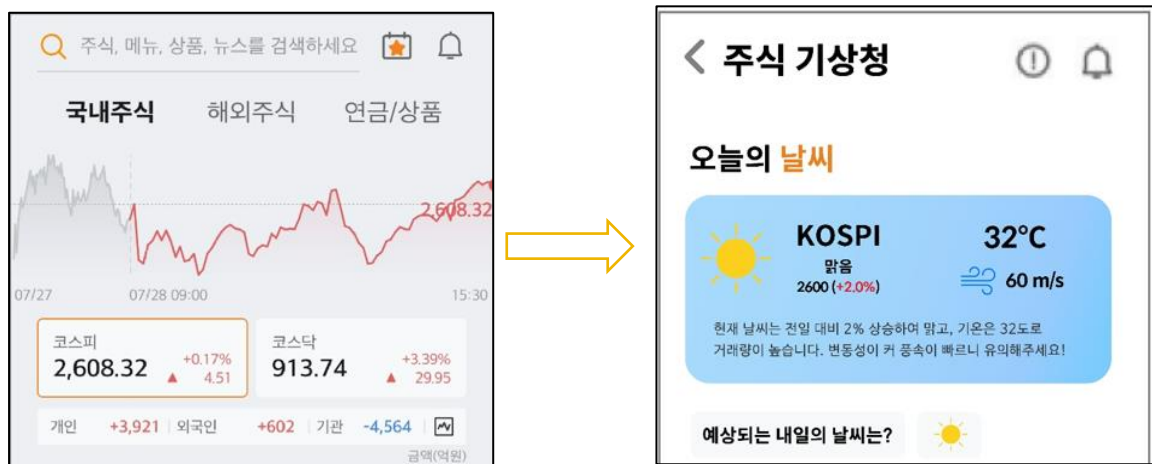
2-1) KOSPI지수: 종가대비 수익률을 계산하여 $\pm 0.5\%$ 를 기준으로 맑음, 구름, 비로 분류

2-2) 거래량: 과거 1년 간 거래량의 최소값과 최대값을 이용하여 현재의 거래량을 0~40 사이의 온도로 스케일링 하여 표현

2-3) 변동성: 현재의 변동성을 풍속으로 표현

3) 결과 종합: 날씨로 변환된 3가지 지표를 시각화하고, 간단한 문장으로 요약

3.1.4. 디자인 및 기대효과



일반적으로 주식 시장과 같은 복잡한 시스템에서는 수많은 지표와 수치들이 존재하며, 이들 간의 상대적인 크기나 중요성을 파악하는 것은 전문가가 아니라면 어려운 일입니다.

이러한 지표들을 날씨라는 모두에게 익숙한 개념으로 변환하면, 사람들은 직관적으로 시장 상태를 파악할 수 있게 됩니다. 예를 들어, 대부분의 사람이 맑은 날씨를 긍정적으로 인식하므로 KOSPI지수의 등락을 날씨로 표현한다면, 날씨가 "맑음"일 때 지수가 상승했음을 쉽게 받아들일 수 있습니다. 또한 높은 온도는 뜨거운 이미지와 연결되므로, 거래량을 온도로 표현하면 거래 활동의 활발함을 직관적으로 이해할 수

있습니다.

이런 방식은 상대적인 수치를 표현하는 데도 유리합니다. 즉, 어떤 하루가 다른 날보다 더 덥거나 바람이 더 세게 부는 것처럼, 어떤 날의 거래량이나 변동성이 평균에 비해 어느 정도인지를 쉽게 파악할 수 있습니다.

따라서, 이런 방식은 주식 시장의 복잡성을 단순화하고, 시장 상태를 빠르게 파악할 수 있게 해주며, 상대적인 지표의 중요성을 이해하는 데 도움을 줍니다. “주식 시장 기상청”은 오른쪽 이미지에 제시된 “오늘의 날씨” 서비스를 제공하여 날씨로 변환된 시장 지표를 한 눈에 알아볼 수 있도록 합니다. 나아가 날씨정보들을 요약하여 간결한 문장을 제시함으로써 투자자들에게 시장 상태를 효과적으로 전달합니다.

3.2. 일기 예보

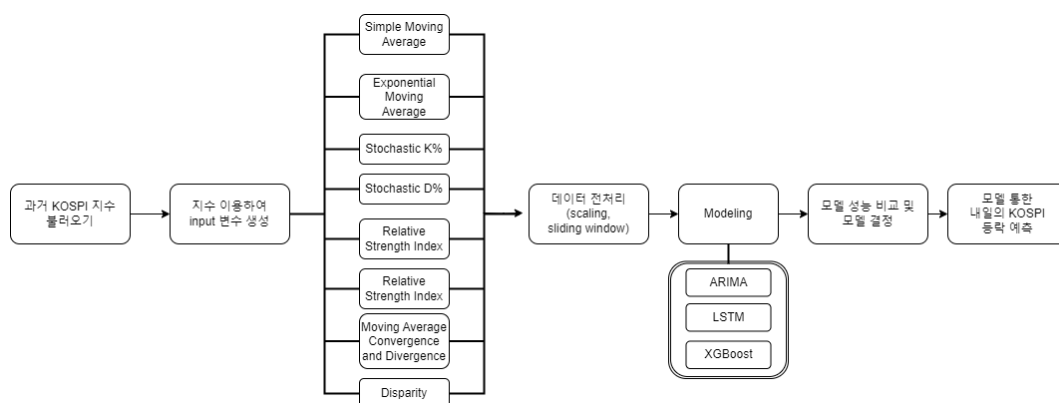
3.2.1. 서비스 설명

날씨 형태로 오늘의 상황을 보여주는 것에서 더 나아가 일기예보 형식으로 내일의 KOSPI지수 등락을 예측하는 서비스입니다. 예측 모델은 과거 KOSPI지수와 관련된 기술적 지표들을 이용한 “ARIMA”, “LSTM”, “XGBoost”와 같은 방법들로 구성되며, 이를 통해 내일의 날씨를 예측하여 투자자들에게 제시합니다.

3.2.2. 사용 데이터

- 과거 KOSPI지수 데이터: Python “yfinance”

3.2.3. 메인 알고리즘 및 Flow chart



- 1) 데이터셋 구축: Python “yfianance” 패키지 활용한 KOSPI 종가 데이터 구축
- 2) 기술적 지표 변수 생성
- 3) 데이터 전처리: 스케일링, 윈도우 크기에 맞게 데이터 재구조화
- 4) 모델링 및 성능평가: ARIMA, LSTM, XGBoost 모델링 후 각각의 성능 비교
- 5) 모델 결정 및 예측값 산출: 성능이 가장 좋은 모델 선정 및 이를 통한 KOSPI 등락 예측

3.2.4. 모델 설명

본 모델은 논문 “XGBoost 모형을 활용한 코스피 200 주가지수 등락 예측에 관한 연구”(하대우, 김영민, 안재준, 2019)를 참고하였습니다.

내일의 KOSPI지수 등락을 예측하는 모델을 구축하고자 과거 KOSPI지수를 그대로 사용하는 것이 아닌, KOSPI지수의 움직임을 잘 나타낼 수 있는 기술적 지표들을 활용하였습니다. 다음은 사용한 기술적 지표들입니다.

- **SMA(simple moving average)**: 특정 기간 동안 주가가 가지고 있는 방향성을 수치화
- **EMA(Exponential moving average)**: 최근 시점에 더 높은 가중치를 부여하는 이동평균 방법
- **Stochastic지표**: 일정 기간 동안 종가를 바탕으로 주가의 상승/하락 강도를 나타내는 단기적 지표
- **RSI(Relative Strength Index)**: 현재의 주가 추세 강도를 백분율로 표현하여 상대적 추세를 측정할 수 있는 지표
- **MACD(Moving Average Convergence and Divergence)**: 단기이동평균 값과 장기이동평균 값과의 차이를 이용하여 주가의 추세를 파악할 수 있는 지표
- **Disparity**: 주가와 이동평균 값과의 차이를 측정하는 지표

위 지표들을 입력변수로 하고, KOSPI의 일별 종가를 출력변수로 하여 모델을 학습시킬 예정입니다. 시계열이라는 데이터의 특성을 고려하여 과거 데이터를 누적하여 다음 시점의 값을 예측하는 “Sliding window” 방법론을 적용합니다. 이 방법론에서

과거 누적된 데이터를 윈도우라고 하는데 예를 들어 윈도우 사이즈를 5라고 설정하면 과거 5일의 데이터를 가지고 다음날의 값을 예측하는 방식입니다.

사용할 모델로는 시계열데이터를 다루는 전통적인 통계 모형인 "ARIMA", Neural Network 모형인 "LSTM", 그리고 다양한 분야에서 일반적으로 좋은 성능을 보인다고 알려진 "XGBoost"가 있습니다. 이들을 활용해 지수가 상승할지 하강할지를 분류하는 모델링 과정을 거치고, 오분류율이 가장 작은 모델을 최종 모델로 선정합니다. 이후 최종 모델의 예측 결과가 상승이면 맑음, 하강이면 비로 내일의 날씨를 예보합니다.

3.2.5. 디자인 및 기대효과

오늘의 날씨



다음과 같이 오늘의 날씨 하단에 내일의 날씨가 표시됩니다. 이는 투자자들에게 시장의 움직임을 파악하는 데 도움을 주는 예측정보를 제공합니다. 이는 의사결정 시 고려할 수 있는 여러 정보들 중 하나로써 기능할 수 있을 것으로 기대됩니다.

3.3 뉴스 해시태그

3.3.1. 서비스 설명

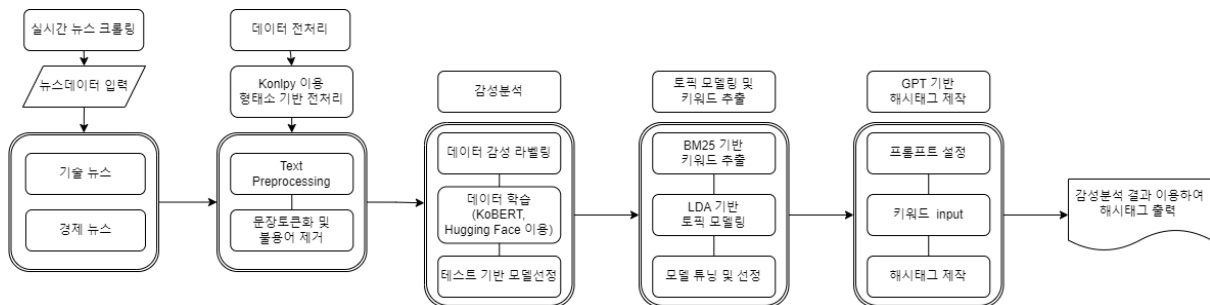
"주식시장 기상청"은 단순히 주식시장의 현황을 표시하는 것뿐만 아니라, 사용자들이 서비스의 메인 화면을 통해 현재 시장의 트렌드까지 함께 알 수 있도록 뉴스 데이터를 이용한 "뉴스 해시태그" 서비스를 함께 제공합니다. 이 서비스는 금융 뉴스에서 오늘의 키워드를 찾아 간결하고 쉽게 이해할 수 있는 해시태그 형식으로 표시해 줍니다. 이를 위해 "자연어 처리(NLP)", "감성분석", "LDA 토픽 모델링", "랭킹 알고리즘 BM25를 활용한 키워드 추출"을 진행하고, 이 결과를 "ChatGPT API"를 이용해 해시태그로 변환하는 과정을 거칩니다.

3.3.2. 사용 데이터

네이버 뉴스의 기술, 경제 섹터에서 업데이트 시점을 기준으로 최근 200개의 기사를 크롤링하여 만든 csv파일을 사용할 예정입니다.

3.3.3. 메인 알고리즘 및 Flow chart

많은 선행연구(Shao, 2022; Liu, 2020) 등에서는 LDA를 활용하여 뉴스의 토픽을 정하고, Kashoob et al(2015)는 LDA에 BM25를 혼합한 LDA+BM25를 사용하여 토픽 내 단어의 랭킹을 정하고 있음을 참고하여 모델링을 진행하였습니다.



- 1) 데이터셋 구축: 크롤링을 통해 위에서 언급한 csv파일 생성
- 2) 전처리: Python “Konlpy” 패키지를 이용하여 결측치 제거, 토큰화, 불용어 제거, 형태소 기반 추출, 패딩 등과 같은 전처리 진행
- 3) 감성 분석
 - 3-1) 훈련 데이터 구축: 과거 금융 뉴스에 대한 감성 라벨링
 - 3-2) 감성 모델 구축: Kobert/Hugging face 등의 모델을 금융 데이터에 적합한 모델로 학습
 - 3-3) 모델 선정 및 분석: 학습된 모델들의 성능을 비교하여 가장 좋은 모델 선정 후 감성분석 진행
- 4) LDA 토픽 모델링: (2)의 데이터에서 6개의 토픽 추출
- 5) BM25를 활용한 단어 별 중요도 Rank: 토픽 별 포함된 단어의 빈도를 기반으로 함
- 6) 프롬프트 구축: (5)의 결과를 입력했을 때 적절한 해시태그를 반환하도록 프롬프트 설정

- 7) API 이용 해시태그 제작: Chat GPT API 프롬프트를 입력하여 토픽 별 해시태그를 제작
- 8) 해시태그 별 감성분석: (3), (4)의 결과를 활용하여 토픽의 감성을 분류하고 이를 해시태그에 시각화
- 9) 일별 2회 업데이트 진행

3.3.4. 모델 설명

“뉴스 해시태그” 서비스의 최종 목표는 경제 및 기술 뉴스의 트렌드를 사용자들에게 간결하게 전달하는 것입니다.

이를 위해 가장 먼저 할 것은 “감성분석”입니다. “감성분석”이란 텍스트 데이터 내의 긍정/부정적인 감성을 판별하는 기법입니다. 주식 시장의 흐름은 다양한 요인에 의해 영향을 받지만, 그 중 하나가 바로 뉴스와 같은 대중 매체의 정보입니다. 특히, 경제 및 기업에 대한 뉴스는 해당 기업의 주식 가격에 큰 영향을 줄 수 있으며, 긍정적인지 부정적인지에 따라 시장의 반응이 달라질 수 있습니다. 경제 및 기술 뉴스만의 감성을 분석하는 모델이 따로 존재하지 않기 때문에 직접 모델을 훈련시키는 과정을 거쳐야 합니다. 같은 분야 과거 뉴스 데이터의 감성이 긍정적인지 부정인지 직접 라벨링 한 후 이를 훈련데이터로 Kobert/Hugging face 등의 모델을 학습시킵니다. 학습된 모델을 경제 및 기술 뉴스의 감성을 분석하는데 특화된 모델이 되고, 이를 활용해 최근 뉴스의 감성을 긍정, 부정, 중립으로 분류합니다.

그 다음 과정은 “LDA(잠재 디리클레 할당)”입니다. 분석을 통해 투자자들에게 제공될 약 200개의 뉴스들이 제공하는 다양한 정보를 몇 개의 주제로 요약하는 과정이 이에 해당합니다. “LDA”는 관찰 집합을 관찰되지 않은 그룹으로 설명할 수 있도록 하는 생성 통계 모델입니다. 이 모델은 텍스트 데이터가 여러 토픽의 혼합으로 생성되고, 이러한 토픽이 확률 분포에 따라 텍스트 내 단어를 생성한다고 가정하므로, 이러한 단어들의 분포를 확인하고 클러스터링 함으로써 공통적인 주제를 갖는 단어들의 클러스터인 토픽을 생성할 수 있습니다. 예를 들어, LDA결과 다음과 같은 토픽을 얻을 수 있습니다. “토픽_1: 시장, 은행, 금리, 상승, 감독 ...”

이후 각 토픽을 가장 잘 대표하는 단어를 선정하여 해시태그 형태로 만드는 과정이 필요합니다. 이때 “BM25” 기법을 활용합니다. “BM25”는 일반적으로 검색 엔진에서 특정 검색어와의 관련성에 따라 문서의 일치성을 판단하는 순위 알고리즘입니다.

문서 내용에 검색어 출현 빈도가 높을수록, 다른 문서에는 검색어가 출현하지 않을수록, 문서 내용이 적을수록 더 큰 점수를 부여합니다. "BM25" 알고리즘의 결과 점수를 활용하여 각 토픽 별로 이를 대표하는 단어들을 식별할 수 있습니다. 가령, 위에서 언급한 "토픽_1"의 bm25점수가 {금리 100, 은행 98, 상승 77, 감독 65, 시장 44 ...} 라면 토픽 1의 키워드는 금리, 은행, 상승 정도로 예상할 수 있습니다.

추출한 키워드를 해시태그 형태로 변환하기 위해 "Chat GPT API"를 이용합니다. 항상 적절한 해시태그를 출력하도록 키워드 입력 프롬프트를 구축하고, 이를 "Chat GPT"에 입력합니다. 앞선 예시는 "#은행금리상승" 이라는 해시태그를 출력할 것입니다.

마지막으로 해당 주제를 구성하는 기사들의 감성분석 결과를 바탕으로 해시태그의 색을 결정합니다. 긍정이면 붉은색, 부정이면 파란색, 중립이면 검은색을 사용합니다.

정리하면, "LDA"를 통해 뉴스 데이터를 임의의 주제로 분류하고, "BM25"를 통해 임의의 주제를 나타내는 키워드를 추출한 뒤, "Chat GPT API"를 이용해 이를 해시태그 형태로 변환하고 "감성분석"을 이용해 해시태그의 색을 결정하는 모델을 구상했습니다.

3.3.5. 디자인 및 기대효과



“뉴스 해시태그” 서비스는 오른쪽 이미지와 같이 투자자들에게 제공될 예정입니다. 기존 “M.STOCK”앱에서는 실시간 뉴스와 그 중 중요하다고 판단되는 뉴스를 선정해서 제공하는 방식의 서비스를 제공했습니다. 이와 달리 “뉴스 해시태그”는 전반적인 뉴스의 트렌드를 키워드와 색의 조합을 이용해 투자자에게 전달하는 방식을 선택했습니다. 실시간으로 많은 양의 뉴스들이 업데이트 되는 디지털 사회에서 개인투자자가 다양한 주제의 뉴스들을 하나하나 파악하는 것은 많은 시간과 비용을 요구하는 일이기에 해시태그 형태의 요약적 제시는 충분히 좋은 호응을 얻을 것으로 기대됩니다.

나아가 투자자들이 “뉴스 해시태그”를 통해 전반적인 시장의 흐름을 살피고 관심 있는 키워드를 터치했을 때, 해당 토픽에 대한 미래에셋증권의 자세한 AI뉴스가 제시되도록 한다면, 기존 서비스와 밀접하게 연동되어 시너지 효과를 기대할 수 있을 것 같습니다.

3.4. MY 날씨



“오늘의 날씨”(3.1)에서는 코스피 지수에 대한 여러 지표들을 날씨 형태로 변환하는 서비스를 제공했다면, “MY 날씨”는 투자자 개인이 보유하고 있는 각 종목들에 대한 지표들을 날씨로 변환하여 소비자들에게 제공합니다. KOSPI지수가 아닌 개별주식 데이터를 크롤링한다는 것 이외에는 사용된 알고리즘과 서비스의 구조가 “오늘의 날씨”와 동일합니다. 대상이 개별 주식이므로 아래 그림과 같이 주가 상황에 대한 간단한 코멘트가 추가적으로 가능하다는 점이 차이점입니다.

테슬라의 **주가**는 상승중이지만 **거래량**은 하락중입니다.
상승 추세가 멈출 수 있으니 각별히 유의하시길 바랍니다.

4. 추가 활용 방안

“주식시장 기상청” 서비스의 가장 큰 장점이자 차별화 전략은 복잡하고 어려운 시장 지표 대신 친숙하고 직관적인 날씨라는 매개체를 이용하여 투자자들에게 정보를 전달한다는 점입니다. 이를 “M.STOCK”에서 기존에 제공하고 있었던 서비스와 연동한다면, 좋은 효과를 얻을 수 있을 것으로 기대됩니다.

4.1. 이상기후 탐지 경보

“M.STOCK”이 제공했던 서비스 “Dr.big의 시그널”에서는 주식에 중요한 영향을 미치는 주요 네 가지 시그널(공시, 가격변화, 거래량변화, 보조지표)이 발생하면 투자자에게 알림을 줍니다. 날씨를 활용하여 이러한 시그널을 표현하면 위에서 언급한 바와 같이 직관적으로 정보를 전달할 수 있습니다. 예를 들어, 가격의 급격한 상승을 “Dr.big의 시그널” 서비스가 탐지했다면

" *맑은 날입니다! *: [주식명] 주식이 눈부시게 빛나고 있어요! 현재 가격은 [현재가격]로, [증가율]% 상승하였습니다. 오늘은 주식시장에서 기분 좋은 날입니다! "

와 같은 알림을 개인에게 보낼 수 있습니다.

4.2. 개인별 예보시스템

기상 예보에서 기상 캐스터는 오늘의 날씨뿐만 아니라 날씨에 따른 행동요령을 함께 알려줍니다. 가령 비가 오는 날에는 우산을 챙겨라, 눈이 오는 날에는 도로가 미끄러우니 주의하라는 멘트가 기상 예보에 함께 포함됩니다. 이러한 점에서 착안해 “주식시장 기상청”이 제공하는 “MY날씨”와 “M.STOCK”에서 제공하는 “Dr.big의 투자진단” 서비스를 결합하여 개인별 투자 특성을 고려한 직관적이고 효과적인 조언을 제공하면 좋을 것이라 생각했습니다.



왼쪽 이미지는 개인의 투자패턴을 5가지 기준으로 나타내는 기존의 서비스입니다. 이를 기반으로 가령, 분산투자를 잘하지 않는 고객에 대해서는 다음과 같은 조언을 해줄 수 있습니다.

“투자비중이 높은 주식이 전일대비 -8%로 폭우주의보가 발효되었습니다. 기온은 33도로 거래량 높습니다. 주식 가격 움직임에 각별히 주의해주세요!”

“MY주식”의 날씨 정보를 활용하여 현재 보유 주식의 상황을 직관적으로 전달하므로 초기 투자자들도 무리 없이 이해할 수 있으며, 본인의 성향에 맞는 예보이므로 만족도도 높을 것으로 기대됩니다.

5. 참고문헌

하대우, 김영민, 안재준(2019).XGBoost 모델을 활용한 코스피 200 주가지수 등락 예측에 관한 연구. 한국데이터정보과학회지, 30(3), 655-669.

<http://dx.doi.org/10.7465/jkdi.2019.30.3.655>

한국소비자원(2021.06.29) 보도자료 '증권앱 소비자 만족도, NH 투자증권 나무 가장 높아'

<https://www.kca.go.kr/home/sub.do?menukey=4002&mode=view&no=1003155749>

Chanyuan Liu (2020). Analysis of Relationship Between Hot News and Stock Market—Based on LDA Model and Event Study. J. Phys.: Conf. Ser. 1616 012048.

<https://doi.org/10.1088/1742-6596/1616/1/012048>

Said Kashoob, James Caverlee, and Krishna Kamath (2010). Community-based ranking of the social web. In Proceedings of the 21st ACM conference on Hypertext and hypermedia (HT '10). Association for Computing Machinery, New York, NY, USA, 141–150.

<https://doi.org/10.1145/1810617.1810642>

Shao, D., Li, C., Huang, C. et al (2022). A news classification applied with new text representation based on the improved LDA. Multimed Tools Appl 81, 21521–21545.

<https://doi.org/10.1007/s11042-022-12713-6>