



# 파이썬을 이용한 금융상품 가격 예측 서비스 구현

키워드

PyQt

Python

Scikit-learn

Tensorflow

## ✓ 분석 개요

### 분석 배경

- 은행 예금으로는 목돈 마련이나 노후 대비가 어려워지면서 일반인들의 주식, 채권, 펀드 투자가 급증하고 있음

### 분석 목적

- 주식 가격과 경제 뉴스 기사를 수집/분석하여 투자자들의 합리적인 의사결정을 지원

### 분석 목표

- 주식 가격 데이터와 뉴스 기사를 분석하여 최적의 매매 타이밍을 추천하는 자동화된 분석 시스템 구축
- 사용자가 분석 결과를 직관적으로 파악할 수 있는 GUI 애플리케이션 개발

### 분석 내용

- 주식 가격 예측
  - 5개의 시계열 모델(DNN, CNN, RNN, LSTM, ARIMA)을 구현하고 각 모델의 예측 성능을 RMSE를 기준으로 비교 분석
  - 예측 성능이 가장 뛰어난 최적의 데이터 크기를 도출하여 예측 정확도를 향상
  - 주식 보조 지표 활용과 하이퍼 파라미터 튜닝을 통한 예측 정확도 향상
  - 5일간의 주가를 예측하고, 예측 결과를 기반으로 매수/매도 타이밍 추천
- 주식 시장 이슈 분석
  - 웹 크롤링을 통해 사용자가 선택한 종목의 관련 뉴스 기사를 수집
  - 기사에서 키워드를 추출하여 워드클라우드로 시각화

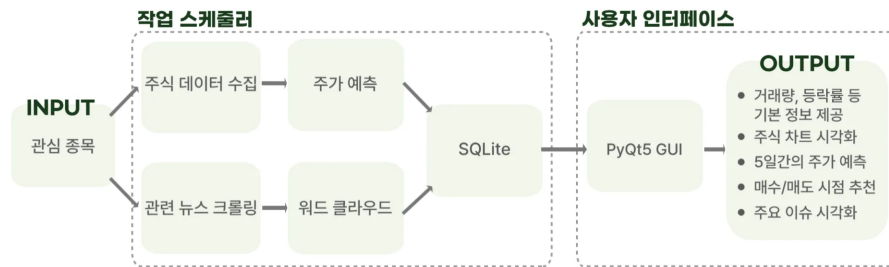
### 분석 결과

- 자동화된 데이터 수집 및 분석 시스템 구현
  - 작업 스케줄러를 통해 매일 저녁 자동으로 데이터를 수집하고 분석을 수행
  - 분석 결과를 DB에 저장하여 사용자 요청 시 분석 결과를 즉시 제공
- 애플리케이션 구현을 통한 정보 제공
  - PyQt5 기반의 GUI를 통해 분석 결과를 직관적으로 제공
  - 워드클라우드를 통한 주요 이슈 시각화로 투자 의사결정 지원
  - 매수/매도 타이밍 추천 기능으로 실질적인 투자 가이드 제공

## ✓ 분석 프로세스

### 1. 프로젝트 구성도

## 프로젝트 구성도



## 2. 데이터 수집 및 저장

- 키움증권 Open API를 활용하여 관심 종목의 주가 데이터를 수집 후 sqlite에 저장
  - python 32bit 환경 사용
- 한경, YTN, 중앙일보 등 주요 언론사의 관심 종목 관련 뉴스 기사를 웹 크롤링 후 sqlite에 저장
  - Python의 BeautifulSoup 패키지 사용

## 3. 분석 내용

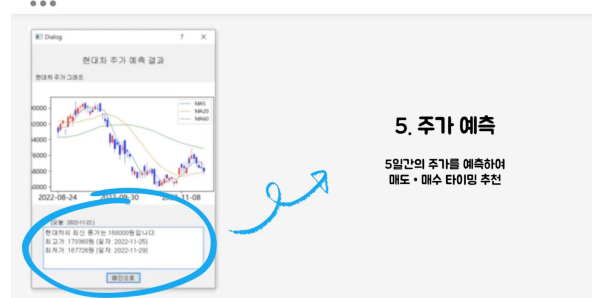
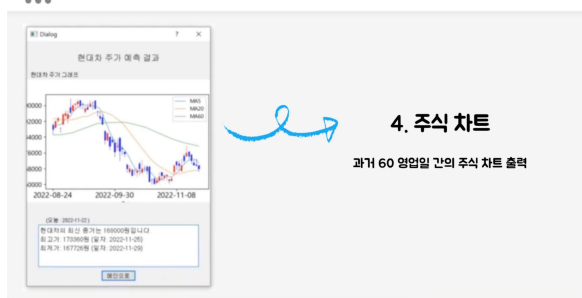
- 금융 상품 가격 예측
  - 다양한 시계열 모델(DNN, CNN, RNN, LSTM, ARIMA)을 구현하고 RMSE 기준으로 예측 성능을 비교하여 최적의 모델을 선택
  - 예측 정확도를 높이기 위해 최적의 데이터 크기를 도출
  - 기준금리, 통화량, 환율 등의 보조 지표를 활용하여 예측의 정확도를 향상
  - 하이퍼파라미터 튜닝을 통해 예측 정확도를 개선
- 분석을 통한 정보 제공
  - 관심 종목들의 5일간의 주가를 예측하고 최적의 매수,매도 시점을 추천
  - 관심 종목들의 거래량, 가격 상승률, 시가총액 정보를 제공
- 관련 금융 상품 키워드 제공
  - 크롤링한 기사에서 키워드를 추출하여 워드클라우드를 시각화
- 자동화
  - 거래량이 열렸던 날 저녁에 자동으로 예측 모델을 미리 실행해 DB에 저장해두고 필요 시 DB에 접속해서 분석 결과를 사용자에게 바로 제공할 수 있도록 함.

## 4. 분석 결과

### 애플리케이션 구현

- Python pyqt5 사용





## 5. 문제점과 해결 방안

1

- 문제점: 금융상품 예측 모델 생성 및 예측 결과 도출까지 10분 이상이 소요되어 실시간 예측과 결과 제공에 어려움이 있었다.
- 해결방안: 거래장이 종료된 날 저녁에 미리 예측 모델을 실행하고 결과를 DB에 저장한 후, 사용자 요청 시 저장된 예측 결과를 즉시 제공하는 방식으로 개선하였다.

2

- 문제점: 주가 예측 성능이 알고리즘과 데이터 크기에 따라 일관성이 부족했다.
- 해결방안: 주가 예측 시 다양한 알고리즘을 적용하고, 2개월·6개월·1년 등 여러 기간의 데이터로 실험을 수행하여 예측 정확도가 가장 높은 알고리즘과 최적의 데이터 기간을 도출하여 예측에 사용했다.