**통합 주가 예측 모델 보고서**

**팀 명 : QuantVerse**

**팀 원 : 백준원, 조재상, 조혜진, 최동주**

**목 차**

[1. 프로젝트 개요 4](#_Toc182999597)

[1.1. 프로젝트 목적 4](#_Toc182999598)

[2. 데이터 및 모델 구성 4](#_Toc182999599)

[2.1. 데이터 소스 4](#_Toc182999600)

[2.2. 모델 구성 4](#_Toc182999601)

[2.3.1. 경제상황 예측 모델: 4](#_Toc182999602)

[2.3.2. CAPM 모델을 활용한 주식 가격 평가: 4](#_Toc182999603)

[2.3.3. 산업군 분석 및 개별 기업 분석. 4](#_Toc182999604)

[2.3.4. 기술적 분석 기반 평가: 5](#_Toc182999605)

[3. 분석 과정 5](#_Toc182999606)

[3.1. 경제상황 예측 5](#_Toc182999607)

[3.2. 산업군 분석 5](#_Toc182999608)

[3.3. 개별 주식 분석 5](#_Toc182999609)

[3.4. 기술적 분석 5](#_Toc182999610)

[4. 모델링 5](#_Toc182999611)

[4.1. 경제 예측 모델링 5](#_Toc182999612)

[4.2. 산업별 모델링 6](#_Toc182999613)

[4.3. 주식 가격 평가 모델링 6](#_Toc182999614)

[4.4. 기술적 분석 모델링 6](#_Toc182999615)

[5. 결과 및 논의 6](#_Toc182999616)

[5.1. 경기 예측 결과 6](#_Toc182999617)

[5.2. 산업별 분류 결과 6](#_Toc182999618)

[5.3. 개별 주식 가격 평가 6](#_Toc182999619)

[5.4. 기술적 분석 결과 7](#_Toc182999620)

[6. 결론 및 향후 과제 7](#_Toc182999621)

# 1. 프로젝트 개요

## 1.1. 프로젝트 목적

본 프로젝트는 다양한 기법을 활용하여 경제상황을 예측하고, 이를 바탕으로 통합적인 주가 예측 모델을 개발하는 것을 목표로 합니다. 구체적으로 주식 자체의 자료만을 사용하여 분석하는 기술적 분석과 개별 기업의 재무제표를 분석하는 기본적 분석, 개별 주식의 수익률과 위험에 기반한 자산가격결정모형(CAPM) 및 기타 거시경제 지표를 활용한 주식 가격 예측 모델을 만들 예정입니다.

본 프로젝트에 앞서 기존에 진행된 다양한 주가 예측 모델에 대해 소개하고 그 한계를 확인하였습니다. 먼저 ARIMA모델을 활용한 주가 예측 방법론이 있습니다. 시계열 데이터를 분석하고 미래를 예측하는데 사용되는 통계적 기법으로서, 자기회귀, 차분, 이동평균을 결합한 모델입니다. 해당 모델을 활용하여 주가의 등락을 예측한 결과 가장 높은 값으로 60%를 보고하는 논문이 있었습니다. 하지만 선형적인 결과값을 보여주기 때문에 정확한 가격을 제시하기 어렵고, 일반적으로 상승장이 오래 지속되었던 과거 데이터를 감안하면 지금과 같은 하락 추세를 예측하기 어렵다는 단점이 있습니다.

다음으로 LSTM 및 GRU를 활용한 주가 예측 방법론이 있습니다. 이러한 딥러닝 모델은 입력 데이터의 특징 패턴을 학습하여 기존 통계 및 머신러닝 기법보다 주가 예측 성능이 우수하다고 평가됩니다. 이에 따라 주가 예측 모델이 수익이 나게끔 동작했으나 동 기간 전체 시장의 수익대비 미미 하였습니다. 또한 동 모델을 활용하였을 때 데이터 수의 제한으로 모델의 복잡성으로 인한 과적합 문제가 발생하고 ARIMA와 마찬가지고 상승장의 데이터를 학습하였을 때 지금과 같은 하락 추세 예측이 어렵다는 단점이 있습니다.

이 외에도 SVM, HMM 등의 모델을 활용한 주가예측 기법이 있었습니다. 위의 모델을 활용한 논문 및 프로젝트에서는 모델이 어떤 데이터를 학습하는지와 외부의 경제 충격을 어떤 방식으로 모델에 적용했는지에 따라 예측 정확도가 달라졌지만, 대부분의 프로젝트에서 한 개 내지 다섯 개의 종목 주가 예측을 하여 일반화 시키기 어렵다는 한계점이 존재하였습니다. 이에 본 프로젝트에서는 KOSPI 200 (시가총액 기준 상위 200개 기업)의 기업을 선정하여 INPUT 값으로 KOSPI200에 해당하는 종목명을 넣으면 한 달 후의 주가를 예측하는 조금 더 일반화 된 모델을 더 높은 확률로 예측하고자 합니다. 이에 아래의 다섯가지 기준을 본 모델에 적용하여 예측 모델을 구현하고자 합니다.

* 경기선행지수
* CAPM 기반 개별 종목의 민감도
* 특정 산업의 민감도
* 재무제표 지표(수익성, 안정성, 성장성 등)
* 기술적 분석

# 2. 데이터 및 모델 구성

## 2.1. 데이터 소스

* **경기선행지수**: 통계청에서 제공되는 종합적인 거시경제 지표를 활용하였습니다.
* **개별 주식 데이터**: 한국거래소(KRX) 통계자료로 개별 종목의 일일 시가, 종가, 저가, 고가, 거래량 등의 자료 및 기타 지수 정보, 국채 정보를 활용하였습니다.
* **기업별 재무제표** : Dart 전자공시 통계자료로 개별 기업의 분기, 반기, 사업 보고서를 활용하여 재무분석 자료를 활용하였습니다.
* **추가 지표**: 월별/연도별 주요 경제 지표와 기존 데이터 소스에서 가져온 자료를 통해 만든 보조지표(기술적 분석 지표)입니다.

## 2.2. 모델 구성

### 2.3.1. 경제상황 예측 모델:

* + 경기선행지수를 기반으로 향후 6개월 경제 동향 예측.

### 2.3.2. CAPM 모델을 활용한 주식 가격 평가:

* + CAPM은 포트폴리오의 자산가격을 결정하는 이론적 모형으로 해당 이론의 베타(B)는 개별 종목이 전체 시장의 수익률에 대비해 얼마나 민감한지를 나타냅니다. 이에 따라 전체 시장의 흐름에 대비해 개별 종목이 어느 정도 상승, 하락하는지 알려주는 지표로써 활용됩니다.
  + 경제상황과 Beta 값을 활용한 이론적 주가 계산.

### 2.3.3. 산업군 분석 및 개별 기업 분석.

* + 산업별 분류 및 각 산업별 주가 민감도 계산
  + 개별 기업 재무제표 분석에 기반한 주가 예측

### 2.3.4. 기술적 분석 기반 평가:

* + 주가 변동성 분석.
  + 저평가 및 고평가 여부 판단.

# 3. 분석 과정

## 3.1. 경제상황 예측

* 경기선행지수 변동성을 분석하여 짧은 기간 내 경제 상황을 예측.
* 최근 1년간 데이터를 기반으로 경기 변화 패턴 탐지.

## 3.2. 산업군 분석

* 각 산업군별 주가에 대한 민감도를 산출 및 분석
* 산업별 가중치 설정으로 경기 방어 산업군, 경기 민감 산업군 분석

## 3.3. 개별 주식 분석

* 각 주식의 민감도(Beta)를 산출하여 시장 대비 위험성 분석.
* 개별 기업 재무제표를 활용한 기본적 분석 진행..

## 3.4. 기술적 분석

* 월별/연도별 주요 기술적 지표를 분석하여 주가 흐름 예측.
* 매수/매도 시점을 도출하기 위한 추가 분석 진행.

# 4. 모델링

## 4.1. 경제 예측 모델링

* **목적**: 경기선행지수를 활용하여 단기 경제 동향 예측.
* **방법**: 통계적 회귀분석 및 시계열 예측 모델 적용.
* **결과**: 예측 모델의 성능 지표(SMAPE, RMSE)와 실제 경기 데이터 비교.

## 4.2. 산업별 모델링

* **목적**: 산업별 주식 가격 민감도를 활용해 경기방어주, 경기민감주 분류.
* **방법**: 분류 모델 적용.
* **결과**: 실제 주가 데이터와 모델의 산업별 예측 값.

## 4.3. 주식 가격 평가 모델링

* **목적**: CAPM을 활용한 개별 주식 가격 평가.
* **방법**: 주식의 Beta 값을 계산하고 시장 기대수익률 적용.
* **결과**: 이론적 주가와 실제 주가 비교 분석.

## 4.4. 기술적 분석 모델링

* **목적**: 주가 변동성을 분석하여 매수/매도 판단.
* **방법**: 이동평균선, RSI 등 기술적 지표 활용.
* **결과**: 주요 시그널의 신뢰도 및 매수/매도 판단 정확도 평가.

# 5. 결과 및 논의

## 5.1. 경기 예측 결과

* 경기선행지수로부터 추출된 경제상황 변동성 결과 요약.
* 경제 예측 모델의 정확도 검증.

## 5.2. 산업별 분류 결과

* 각 산업의 민감도 분석을 통해 실제 산업의 흐름 예측 결과
* 실제 주가와 차이를 분석 후 모델 개선 가능성 논의.

## 5.3. 개별 주식 가격 평가

* Beta 값과 이론적 주가 비교.
* 실제 주가와의 차이를 분석하여 모델 개선 가능성 논의.

## 5.4. 기술적 분석 결과

* 주요 기술적 지표의 변화와 향후 주가 흐름 예측 결과.
* 저평가 및 고평가 판단에 대한 신뢰성 검토.

# 6. 결론 및 향후 과제

* 통합 주가 예측 모델의 성능 평가.
* 모델 개선을 위한 추가 데이터 및 알고리즘 제안.
* 실제 투자 시 적용 가능성을 높이기 위한 실행 계획 마련.