**통합 주가 예측 모델 보고서**

**팀 명 : QuantVerse**

**팀 원 : 백준원, 조재상, 조혜진, 최동주**

**목 차**

[1. 프로젝트 개요 2](#_Toc182926998)

[1.1. 프로젝트 목적 2](#_Toc182926999)

[2. 데이터 및 모델 구성 2](#_Toc182927000)

[2.1. 데이터 소스 2](#_Toc182927001)

[2.2. 모델 구성 2](#_Toc182927002)

[경제상황 예측 모델: 2](#_Toc182927003)

[CAPM 모델을 활용한 주식 가격 평가: 2](#_Toc182927004)

[기술적 분석 기반 평가: 2](#_Toc182927005)

[3. 분석 과정 3](#_Toc182927006)

[3.1. 경제상황 예측 3](#_Toc182927007)

[3.2. 개별 주식 분석 3](#_Toc182927008)

[3.3. 기술적 분석 3](#_Toc182927009)

[4. 모델링 3](#_Toc182927010)

[4.1. 경제 예측 모델링 3](#_Toc182927011)

[4.2. 주식 가격 평가 모델링 3](#_Toc182927012)

[4.3. 기술적 분석 모델링 3](#_Toc182927013)

[5. 결과 및 논의 4](#_Toc182927014)

[5.1. 경기 예측 결과 4](#_Toc182927015)

[5.2. 개별 주식 가격 평가 4](#_Toc182927016)

[5.3. 기술적 분석 결과 4](#_Toc182927017)

[6. 결론 및 향후 과제 4](#_Toc182927018)

# 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)
* 기술적분석

# 2. 데이터 및 모델 구성

## 2.1. 데이터 소스

* **경기선행지수**: 향후 경제 활동의 전반적인 방향을 예측하기 위한 지표로, 통계청에서 제공하는 경기종합지수 데이터를 활용함.
* **개별 주식 데이터**: 한국거래소(KRX) 통계자료를 통해 특정 종목별로 데이터 수집이 가능함. 기간을 정할 수 있어 시장 동향과 특정 주식의 과거 흐름을 분석할 수 있도록 구성함.
* **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

### 2.2.1. 경제상황 예측 모델:

* + 경기선행지수를 기반으로 향후 6개월 경제 동향 예측.

### 2.2.2. 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.
  + 각 주식의 Beta 값을 산출하기 위해 시장 지수(KOSPI)와 개별 주식 간의 회귀 분석을 진행합니다.
  + 경제이론을 바탕으로 Beta값과 CAPM 방정식 등을 활용하여 각 종목의 이론적 기대 수익률 및 적정 주가를 계산합니다.

### 2.3.3. 산업군 분석 및 개별 기업 분석.

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

### 2.2.4. 기술적 분석 기반 평가:

* + 기술적 분석은 금융 시장을 분석하고 예측하는 기법 가운데 하나로 주로 시세 동향 그래프를 이용해 분석합니다. 이동평균선, RSI, MACD등의 많은 보조지표와 도구를 사용하여주가의 변동성 분석합니다.
  + 보조 지표로 구한 기대 수익률과 기술적 분석을 통해 저평가 및 고평가 여부도 판단합니다.

# 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. 개별 주식 분석

* 각 주식의 민감도(Beta)를 산출하여 시장 대비 위험성 분석.
* 주식의 기대수익률을 바탕으로 이론적 주가 산출.

## 3.3. 기술적 분석

* **목적:** 주가의 흐름을 예측하거나 매수/매도 시점을 판단.
* **방법:** 다양한 지표를 조사 및 분석하고 기대 효과 분석
  1. 이동평균선 (Moving Average)
     + 특정 기간의 평균 주가를 선으로 나타내 주가의 추세를 시각화
     + 일정 기간 동안의 종가의 평균을 계산하여 구함. (예시: 5일 이동평균선은 경우 최근 5일 간의 종가의 평균임)
     + 장기 및 단기 추세 분석에 유용하며, 골든크로스나 데드크로스를 통해 매수/매도 신호를 제공
  2. RSI (Relative Strength Index)
     + 주가의 과매수/과매도 상태를 판단하기 위해 0~100 사이의 값으로 표시되는 모멘텀 지표
     + 매도/매수 타이밍을 잡는데 활용
     + 일정 기간 동안의 상승폭과 하락폭을 이용하여 계산하며 보통 14일을 기준으로 상승폭의 평균과 하락폭의 평균을 이용해 현재 주가의 상대적인 강도를 나타냄.
     + 70 이상일 경우 과매수 구간, 30 이하일 경우 과매도 구간
  3. 볼린저 밴드 (Bollinger Bands)
     + 주가의 변동성을 시각화한 밴드로, 이동평균선을 중심으로 상단과 하단에 위치
     + 밴드의 확장과 축소를 통해 변동성을 파악 및 밴드 경계에서의 반등을 매매 전략으로 활용 가능
     + 이동평균선과 표준편차를 이용해 상단 밴드와 하단 밴드를 계산하며, 밴드 경계에서의 반등을 매매 전략으로 활용할 수 있음.
  4. PSAR (Parabolic Stop And Reverse)
     + 주가가 추세 전환 시점을 나타내는 점으로 표시
     + 추세의 방향성과 추세 종료 지점을 예측해 매수/매도 타이밍을 결정
  5. ATR (Average True Range)
     + 주가의 변동성을 측정하는 지표로, 평균 진폭을 나타냄.
     + 변동성이 큰 시점을 파악하여 리스크 관리에 활용
     + 특정 기간 동안의 진폭(고가와 저가의 차이, 혹은 전일 종가와의 차이 등)을 이용해 평균 변동성을 계산함.
  6. MACD (Moving Average Convergence Divergence)
     + 이동평균선의 차이를 활용해 추세의 강도와 방향성을 나타냄.
     + MACD선과 시그널 선의 교차를 통해 매매 신호 포착 가능.
     + 단기 이동평균선과 장기 이동평균선의 차이를 구하여 계산함.
  7. 오실레이터 (Oscillator)
     + 일정 범위 내에서 움직이며, 주가의 과매수/과매도 상태를 나타냄.
     + 주가의 전환 시점을 예측하여 단기 매매 전략에 활용
  8. 하한선/상한선
     + 주가의 일정 범위 내 상한과 하한을 나타내는 선
     + 돌파 여부에 따라 매매/매도 신호를 제공
  9. 저항선/지지선 (Resistance/Support Levels)
     + 주가가 상승하거나 하락을 멈추는 경계선
     + 주가가 저항선을 돌파하거나 지지선을 이탈할 때 중요한 매매 신호를 제공
  10. %K, %D (Stochastic Oscillator)
      + 주가의 현재 위치를 일정 기간 동안의 고점과 저점 사이에서 비율로 나타내는 지표
      + 둘의 교차를 통해 매매 타이밍을 파악 가능
      + %K는 현재 주가가 일정 기간 동안 고점과 저점 사이에서 어디에 위치하는지를 나타내며, %D는 %K의 3일 이동평균선임.
  11. CCI (Commodity Channel Index)
      + 주가의 추세와 평균 가격 사이의 차이를 측정
      + 주가의 과매수/과매도 상태를 파악해 매매 기회를 제공
  12. ADX (Average Directional Index)
      + 추세의 강도를 나타내는 지표, 방향성과는 무관함
      + 강한 추세와 약한 추세를 구분해 추세 기반 전략 수립에 도움을 줌
      + PDI(Positive Directional Indicator)와 NDI(Negative Directional Indicator)를 이용해 추세의 강도를 측정하며, 100에 가까울수록 강한 추세, 0에 가까울수록 약한 추세임.
  13. 이격도 (Price Oscillator)
      + 주가와 이동평균성 간의 괴리율을 나타내는 지표
      + 주가가 평균에서 얼마나 떨어져 있는지 파악하여 과매수/과매도 상태를 분석
      + 주가가 이동평균선에서 얼마나 떨어져 있는지를 백분율로 나타낸 값.
  14. Envelope
      + 이동평균선을 기준으로 일정 비율만큼 상하로 이동한 선으로 구성된 지표
      + 이동평균선에서 일정 비율 만큼 상하로 이동한 상한선과 하한선으로 구성되며, 주가의 추세와 반전 가능성을 파악해 매매 전략을 도출
  15. OBV (On-Balance Volume)
      + 거래량과 주가의 움직임을 결합한 지표로, 거래량의 흐름을 추적
      + 거래량 변화를 통해 주가의 상승/하락 가능성을 예측
      + 주가가 상승할 때 거래량을 더하고, 주가가 하락할 때 거래량을 빼는 방식으로 계산됨.
  16. Klinger (Klinger Oscillator)
      + 거래량 흐름과 주가의 추세를 결합한 지표
      + 장기적인 자금 흐름을 파악해 추세 변화 신호를 감지
  17. PDI (Positive Directional Indicator)
      + 일정 기간 동안의 상승폭을 기반으로 계산된 지표, 상승 추세의 강도를 나타냄
      + 일정 기간 동안의 고가와 저가를 비교하여 상승폭을 측정하고, 강한 상승 흐름일 때 매수 신호를 제공
  18. NDI (Negative Directional Indicator)
      + 일정 기간 동안의 하락폭을 기반으로 계산된 지표, 하락 추세의 강도를 나타냄
      + 일정 기간 동안의 고가와 저가를 비교하여 하락폭을 측정하고, 강한 하락 흐름일 때 매도 신호를 제공

# 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. 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.

## 4.3. 기술적 분석 모델링

* **목적**: 주가 변동성을 분석하여 매수/매도 판단.
  1. 주식 시장에서 주가의 변동성 분석을 통해 향후 주가의 상승 또는 하락을 예측하며, 더 나아가 상세한 주가까지 예측하는 것이 이 모델의 주요 목표임.
  2. 이를 통해 투자자들이 주가 흐름을 예상하고, 그에 맞는 매매 전략을 수립할 수 있도록 지원함.
  3. 본 모델은 20 영업일 뒤의 주가 예측을 기준으로 하며, 포괄적인 예측을 위해 최소 하루 뒤부터 최대 1달까지 여러 모델로 확장하여 나타냄.
  4. 모델 구성 중 앞에 소개드리는 지표(1~5)를 기본 지표로 삼음.
* **방법**: 이동평균선, RSI, PSAR, ATR 등 기술적 지표 활용.
  1. 이동평균선
     + 주가의 5일, 20일, 50일 이동평균선(MA5, MA20, MA50)을 분석해 각 이동평균선의 차이를 통한 단기 상승, 중간 상승, 강한 상승 등의 추세를 점수화 하여 상승 비율과 하락 비율을 계산함.
     + MA5가 MA20보다 높은 상태에서, MA50보다도 높으면 강한 상승 추세로 판단하고 상승 비율에 0.12를 추가함. 그렇지 않다면 중간 상승 추세로 판단하고 상승 비율에 0.09를 추가함.
     + MA5가 MA20보다 낮은 상태에서, MA50보다도 낮으면 강한 하락 추세로 판단하고 하락 비율에 0.12를 추가함. 그렇지 않다면 중간 하락 추세로 판단하고 하락 비율에 0.09를 추가함.
     + MA5가 MA20보다 높다면 단기 상승 추세로 판단하여 상승 비율에 0.07을 추가함. 그 반대라면 단기 하락 추세로 판단하여 하락 비율에 0.07을 추가함.
  2. RSI (Relative Strength Index)
     + 과매수 및 과매도 상태를 판단하여, 매수/매도 추세를 확인함.
     + RSI가 70 이상일 경우 과매수로 판단하여 하락 비율을 0.06 증가, 30 이하일 경우 과매도로 판단하여 상승 비율을 0.06 증가시킴
  3. 볼린저 밴드
     + 주가의 20일 이동평균을 기준으로 위 아래로 표준편차를 더하고 뺀 밴드임.
     + 상단 밴드를 돌파할 경우 과매수 상태로 판단하여 하락 비율을 0.04 증가.
     + 하다 밴드를 돌파할 경우 과매도 상태로 판단하여 상승 비율을 0.04 증가.
  4. PSAR (Parabolic Stop And Reverse)
     + 주가와의 위치 관계를 통해 매수 혹은 매도 신호를 확인하는 지표임.
     + SAR 값이 현재 주가 아래에 위치할 경우 상승 추세로 판단하여 매수 신호를 나타내어 상승 비율을 0.05 증가시킴.
     + SAR 값이 현재 주가 위에 위치할 경우 하락 추세로 판단하여 매도 신호를 나타내어 하락 비율을 0.05 증가시킴
  5. ATR (Average True Range)
     + 특정 기간 동안의 주가 변동 폭을 평균한 값으로 주가의 변동성을 측정하는 지표임.
     + 변동성이 증가할 경우 하락 비율을 0.03 증가시키고, 감소할 경우 상승 비율을 0.03 증가시킴.
  6. MACD (Moving Average Convergence Divergence)
     + 단기(12일)와 장기(26일) 지수이동평균선의 차이를 기반으로 생성된 MACD 선과 시그널 선(MA9)의 교차를 분석함.
     + MACD선이 시그널 선을 상향 돌파하면 매수 신호로 판단하여 상승 비율을 0.05 증가시킴. 그 반대라면 매도 신호로 보고 하락 비율을 0.05 증가시킴.
     + 그 외 MACD값이 0보다 크면 상승 추세로 판단하여 상승 비율을 0.03 증가시킴. 그 반대의 경우 하락 비율을 0.03 증가시킴.
  7. 오실레이터 (Oscillator)
     + 주가가 과매수 또는 과매도 상태에 있는지를 판단하여 추세 방향성을 확인함.
     + 오실레이터 값이 +80 이상일 경우 과매수로 판단하여 하락 비율을 0.04 증가시킴. 반대로 -80 이하일 경우 과매도 상태로 판단하여 상승 비율을 0.04 증가시킴
  8. 하한선/상한성 & 저항선/지지선
     + 주가가 특정 범위 내에서 움직이는지를 확인하기 위해 설정한 선들을 활용함.
     + 상한선을 돌파한 경우에는 과매수로 판단하여 하락 비율을, 하한선 아래로 내려갈 경우 과매도로 판단하여 상승 비율을, 각각 0.03씩 증가시킴
     + 저항선 위로 돌파할 경우 강한 상승 추세로 판단하여 상승 비율을, 지지선 아래로 이탈할 경우 강한 하락 추세로 판단하여 하락 비율을, 각각 0.05씩 증가시킴
  9. %K, %D (Stochastic Oscillator)
     + 주가의 현재 위치를 일정 기간 동안의 고점과 저점 사이에서 측정하여 과매수/과매도 상태를 확인함.
     + %K 값이 80 이상이면 과매수로 판단하여 하락비율을, 20 이하이면 과매도 상태로 판단하여 상승 비율을 각각 0.02 증가시킴.
     + %K 선이 %D 선을 상향 돌파하면 상승 추세로 판단하여 상승 비율을, %K 선이 %D 선을 하향 돌파하면 하락 추세로 판단하여 하락 비율을, 각각 0.04 증가시킴
  10. CCI (Commodity Channel Index)
      + 주가가 이동평균선 대비 얼마나 떨어져 있는지를 기준으로 과매수/과매도 상태를 판단함
      + CCI가 100 이상이면 과매수 상태로 판단하여 하락 비율을, -100 이하이면 과매도 상태로 판단하여 상승 비율을, 각각 0.03 증가시킴.
  11. ADX (Average Directional Index)
      + 추세의 강도를 확인하기 위한 지표로, 값이 높을수록 강한 추세임.
      + ADX 값은 25를 기점으로 잡았으며, 값을 넘어갈 경우 강한 추세로 판단하여 상황에 맞게 상승 비율 또는 하락 비율을 0.05 증가시킴.
  12. 이격도 (Price Oscillator)
      + 주가와 이동평균선 사이의 괴리율을 기반으로 과매수/과매도 상태를 판단함.
      + 이격도가 105% 이상이면 과매수 상태로 판단하여 하락 비율을, 95% 이하이면 과매도 상태로 판단하여 상승 비율을, 각각 0.03 증가시킴.
  13. OBV (On-Balance Volume)
      + 거래량 흐름을 분석하여 자금 유입 또는 유출을 확인함.
      + OBV가 증가 또는 감소함에 따라 추세를 판단하여 상승/하락 비율을 0.03 증가시킴.
  14. Klinger Oscillator
      + 장기적인 자금 흐름을 분석하여 매수 및 매도 신호를 확인함.
      + Klinger Oscillator가 양수에서 음수로 전환되면 매도 신호로 판단하여 하락 비율을 0.05 증가시킴. 그 반대의 경우는 매수 신호로 보고 상승 비율을 0.05 증가시킴.

* **결과**: 주요 시그널의 신뢰도 및 매수/매도 판단 정확도 평가.

|  |  |
| --- | --- |
|  |  |
|  |  |

* + - 각 데이터는 5월 2일자의 krx open api 데이터를 바탕으로 20 영업일 뒤인 6월 3일의 주가를 예측한 결과임.
    - SK 하이닉스, 현대차, NAVER의 경우 상승/하락을 맞췄으나 LG 에너지 솔루션의 경우 반대로 예측함.
    - 거래량이 높은 8 개 종목들을 가지고 2024년의 특정한 날을 기점으로 20 영업일 뒤의 주가를 예측했을 때의 정확도는 아래와 같음.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 종목 | 삼성전자 | SK하이닉스 | 현대차 | LG에너지솔루션 | 삼성바오로직스 | NAVER | KB금융 | 기아 |
| 정확도 | 43.97% | 56.54% | 45.54% | 47.12% | 63.87% | 48.16% | 39.79% | 51.3% |

# 5. 결과 및 논의

## 5.1. 경기 예측 결과

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

## 5.2. 개별 주식 가격 평가

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

## 5.3. 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.