**□ 모듈프로젝트 산출물**

클라우드 AI 융합 전문가 양성 과정

|  |
| --- |
| 모듈 프로젝트 2   * **보조지표에 따른 주식투자 수익률 분석 –** * 1조 : 허현준(조장), 신종유, 최숙희, 민병석 |

**2021년 8월 30일**

**클라우드 AI 융합 전문가**

**이름 : 최숙희**

|  |
| --- |
| **목 차**  Ⅰ. 투자 종목 선정  1.1. 종목 선정 기준  1.2. 종목 선정을 위한 외부 데이터 테이블화  1.3. SQL Query문  Ⅱ. 보조 지표를 이용한 주식투자  2.1. Bollinger Band 기반 주식투자  2.2. RSI 기반 주식투자  2.3. MACD 기반 주식투자  2.4. 3가지 지표 종합 적용  Ⅲ. 결론  Ⅳ. 소감문 |

**Ⅰ. 투자 종목 선정  
1.1. 종목 선정 기준**

**- 재무제표분석을 통한 경영이 안정된 2개 기업 선정**

**- 중첩 조건 4개, 단독 조건 2개 사용**

* **중첩 조건 : 부채비율, 유동비율, ROE, PER**
* **단독 조건 : ‘당기 순이익’ 상위 종목 중 ‘2018년 총 거래량’ 순으로 검토**

**1.2. 종목 선정을 위한 외부 데이터 테이블화**

|  |
| --- |
| 전자공시 시스템 ‘DART’(https://opendart.fss.or.kr/disclosureinfo/fnltt/dwld/main.do)에서 2019년 ‘재무상태표’, ‘손익계산서’, ‘현금흐름표’ 데이터를 다운받아서 각각 ‘jaemo’, ‘sonic’, ‘hungum’ 테이블로 AWS의 MySQL에 저장.    Python을 이용해서 유가 증권시장의 주식정보를 스크래핑 할 수 있는 ‘pykrx’패키지를 이용해서 ‘종목코드’, ‘시가총액’등의 정보를 데이터프레임으로 변환했다.          AWS의 MySQL에 ‘ciga’테이블을 추가하기 위해 다음과 같은 방식을 사용했다.            executemany()를 이용해서 %s에 대입할 값은 리스트나 튜플을 담은 리스트여야 하므로, pykrx패키지를 이용해 불러온 데이터프레임을 딕셔너리의 형태로 변환한 후에, KEY값이 ‘data’인 부분을 리스트로 받아서 매개변수에 대입한다.    AWS의 MySQL에 테이블이 저장 되었으므로 코랩 등을 이용해서 데이터분석을 할 수 있었다.  최종적인 테이블관계도는 다음과 같다.  **C:\ai\module_project2\table\테이블 관계도.png** |

**1.3. SQL Query문**

* **삼성전자(강의자료에서 분석함)를 제외한 2가지 종목 선정**
* **최종적으로 ‘비덴트’와 ‘바른손이앤에이’를 투자 종목으로 선정**

|  |
| --- |
|  |
|  |

**Ⅱ. 보조 지표를 이용한 주식투자**

* **사회적 이슈와 같은 외부 요소는 배제하고 오직 기술적 지표인 ‘보조 지표’에 따른 적절한 매매 시점 파악과 그에 따른 수익률 분석**
* **지표 별 수익률과의 관계를 분석**

**2.1. Bollinger Band 기반 주식투자**

|  |  |  |
| --- | --- | --- |
| **분석 자료 데이터 프레임화** | | |
|  | | |
| **데이터 프레임 주요 코드** | | |
| **1.**     * **코랩으로 AWS에 저장된 MySQL 정보를 부름** * **불러온 정보를 df로 저장** * **해당 자료는 비덴트 자료. STK\_CD를 바른손으로 교체시, df 결과 출력 가능** | **2.**     * **df 종가(C\_PRICE)의 20일간의 이동평균, 상위밴드, 하위밴드, PercentB를 데이터 프레임으로 표시** * **PercentB가 1일 때의 값은 UB, 0일 때의 값은 LB. 종가 UB를 초과할 때는 이상고가, LB 미만일 때는 이상저가 의미** * **이상저가에 매수, 이상고가에 매도** * **PercentB가 이틀 연속으로 이상저가, 이상고가인 날짜 :**  1. **비덴트 : 02-08(매수), 05-16(매도)** 2. **바른손이앤에이 : 03-26(매수), 04-05(매도), 07-24(매수), 09-06(매도)** | |
| **시각화 결과 – 비덴트** | | |
|  | | |
| **시각화 결과 – 바른손이앤에이** | | |
|  | | |
| **시각화 주요 코드** | | |
| 1.       * **데이터 프레임을 활용해 종가 이동 평균을 선 그래프로 시각화** | | 2.       * **데이터 프레임을 활용해PercentB를 막대 그래프로 시각화** |
| 3.     * **PercentB 그래프를 선 그래프로 시각화** * **그래프를 확대하여 상세 정보를 확인 가능** * **그래프에 마우스가 접근 시, 상세 정보 팝업** * **팝업 정보 활용해, 데이터 프레임으로 확인한 이상저가일(매수일), 이상고가일(매도일) 비교 가능** | | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **수익금 / 수익률** | | | | |
| **비덴트** | **매수일** | **매도일** | **수익금** | **수익률** |
| **1차** | **2019-02-08** | **2019-05-16** | **125,393,080 원** | **25.394 %** |
| **바른손이앤에이** | **매수일** | **매도일** | **수익금** | **수익률** |
| **1차**  **2차** | **2019-03-26**  **2019-07-24** | **2019-04-05**  **2019-09-06** | **113,027,310 원**  **108,171,960 원** | **13.028 %**  **-4.294 %** |

**2.2. RSI 기반 주식투자**

|  |
| --- |
| **시각화 결과 – 비덴트** |
|  |
| **시각화 결과 – 바른손이앤에이** |
|  |
| **시각화 주요 코드** |
| RSI, RSI signal 그리고 매매의 기준선(30%, 70% 지점)을 선 그래프로 표시 후에 매매 적합  시점과 실제 매매 일을 marker로 표시했다. |
| **데이터 분석 주요 코드** |
| RSI는 주가의 상승/하강 각각의 변화량에 기반한 지표이므로 주가(종가)를 나타내는 ‘C\_PRC’ 컬럼을 정수형으로 변환했다.        전일 대비 종가의 차이를 나타낸 컬럼 ‘DIFF’의 값을 상승/하락 분으로 구분하는 UD 함수를 생성 및 적용 후, 각각의 절댓값을 컬럼 'U', 'D'에 대입했다.    rolling()과 mean()을 이용해서 ‘U’, ‘D’컬럼에 저장된 값들의 14일의 이동 평균을 구했다.    RSI와 RSI signal(RSI 6일의 이동 평균)을 구한 후, 데이터프레임에 컬럼으로 추가했다.        - 과매도, 과매수 구간의 시작점을 나타낼 30%지점과 70%지점을 표시하기 위한 컬럼을 추가했다.  - 주식매수 적합 시점을 파악하기 위해서, RSI와 RSI signal 값이 모두 0.3미만인 시점이  생긴다면 과매도 상황으로 판단하고, 그 다음날에 예산(1억 원)으로 가능한 만큼 주식을  매수했다.  - RSI, RSI signal 값이 모두 0.3미만인 행에는 1, 이 외에 행에는 0을 대입하는 함수 ‘TimeToBuy’를 이용해서, ‘time\_to\_buy’컬럼을 추가했다.      - 연속적으로 RSI값이 0.3미만으로 떨어지는 경우가 생기더라도, 3일이 넘도록 지속되지는 않는 경향이 있으므로, RSI와 RSI signal 값이 0.3미만으로 떨어진 최초의 날 이후로 며칠을 기다리기 보다는 최초로 떨어진 그 다음날에 바로 매수하는 것이 적절하다고 판단했다.  - ‘time\_to\_buy’ 컬럼에서 RSI와 RSI signal 값이 연속적으로 0.3 미만인 구간에서는 모두 1일 것이므로, 0.3미만이 되는 최초의 날들을 파악하기 위해서 ‘first\_crashed’ 컬럼을 추가했다.  - ‘time\_to\_buy’ 컬럼에 diff()를 사용한다면, 최초로 0.3미만이 되는 행의 값은 1이 되고, 이외에는 0 또는 -1이 되기 때문에, np.where() 사용해서 ‘first\_crashed’ 컬럼의 값이 1인 행의 인덱스를 찾았다.  - 최초로 RSI기준을 만족한 다음 날의 행 인덱스 알아보면, 2/11, 5/3, 7/25, 7/31, 12/6, 12/26 6개의 시점이 나오므로, 매도 적합일을 고려해서 2/11과 7/25일에 각각 예산범위에서 일괄적으로 매수했다.    매도일을 결정할 때에는 RSI 와 RSI signal이 모두 0.7을 초과하면 과매수 상황으로 판단하고, 매도 적합시점으로 선택했다. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **수익금 / 수익률** | | | | |
| **비덴트** | **매수일** | **매도일** | **수익금** | **수익률** |
| **1차**  **2차** | **2019-02-11**  **2019-07-25** | **2019-06-28**  **2019-09-23** | **154,781,650 원**  **209,019,730 원** | **54.78 %**  **35.04 %** |
| **바른손이앤에이** | **매수일** | **매도일** | **수익금** | **수익률** |
| **1차**  **2차** | **2019-06-24**  **2019-10-11** | **2019-09-19**  **2019-12-27** | **81,662,875 원**  **90,052,825 원** | **-18.33 %**  **10.27 %** |

**2.3. MACD 기반 주식투자**

|  |
| --- |
| **시각화 결과 – 비덴트** |
|  |
| **시각화 결과 – 바른손이앤에이** |
|  |
| **시각화 주요 코드** |
| - ‘masu\_df’ 와 ‘mado\_df’는 매수/매도 날짜 정보를 가지고 있는 데이터프레임  - 매수 일은 MACD 선이 아래에서 Signal 선을 치고 올라갈 때 매수구간이라 판단  - 매도 일은 ‘MACD선 > Signal 선’ 조건을 만족하고 “macdhist” 칼럼 값이 고점일 때 매도하지만, 고점을 알 수 없음으로 떨어진 다음날 매도  - MACD 지표 추가 매수 일은 ‘MACD 선 < Signal 선’ 조건을 만족하고 이격이 가장 클 때 매수하지만, 다른 보조지표 참고 후 매수 일로 설정 |
| **데이터 분석 주요 코드** |
| - MACD는 MACD선과 Signal선으로 구성, 따라서 ‘MACD’/’Signal’ 컬럼 추가  - MACD선은 단기/장기 이동 평균의 차이로 구함, 단기는 12일 장기는 26일의 이동평균이기 때문에 데이터프레임에 ‘ema\_short’/’’ema\_long’ 컬럼 추가 후 두 컬럼의 차이의 값으로 ‘MACD’ 컬럼 추가  - Signal선은 MACD의 9일 지수 이동 평균을 사용하여 컬럼 값 추가  - MACD선과 Signal선의 두 이동평균선의 차이가 가장 큰 시점을 찾아 추세 변화의 신호로 삼기 위해서, MACD선과 Signal선의 값을 빼서 ‘macdhist’ 컬럼 추가 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **수익금 / 수익률** | | | | |
| **비덴트** | **매수일** | **매도일** | **수익금** | **수익률** |
| **1차**  **2차**  **3차**  **4차**  **5차**  **6차**  **7차** | **2019-02-22**  **2019-04-02**  **2019-05-15**  **2019-06-24**  **2019-08-13**  **2019-09-03**  **2019-10-18** | **2019-02-28**  **2019-04-09**  **2019-05-17**  **2019-07-01**  **2019-08-21**  **2019-09-25**  **2019-11-08** | **92,537,313원**  **84,328,358원**  **77,320,461원**  **96,221,019‬원**  **91,036,177원**  **90,819,939원**  **113,759,965원** | **-7.46 %**  **-15.67 %**  **-22.68 %**  **-3.78 %**  **-8.96 %**  **-9.18 %**  **13.76%** |
| **바른손이앤에이** | **매수일** | **매도일** | **수익금** | **수익률** |
| **1차**  **2차**  **3차** | **2019-04-05**  **2019-05-22**  **2019-09-02** | **2019-04-22**  **2019-06-04**  **2019-09-24** | **116,510,903원**  **155,250,779원**  **167,035,491원** | **16.51 %**  **55.25 %**  **67.03%** |

**2.4. 3가지 지표 종합 적용**

* **분석한 3가지 보조 지표에서 중첩하는 매매 시기에 주식 매수 및 매도**
* **단일 지표 고려 시, Bollinger Band의 수익률이 가장 높음**
* **지표 종합 시, 수익률이 높아질 확률 상승**
* **단기적으로 수익률을 올리는 관점에서 보조 지표의 활용은 유용**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **결과 – 비덴트** | | | | |
| Bollinger Band와 RSI를 기준으로 매수하고, Bollinger Band와 MACD를 기준으로 매도한 일자와 종가를 2019년 일별 종가 그래프에 표시했다. | | | | |
| **수익률 결과** | **매수** | **매도** | **결산** | **수익률** |
|  | **2019-02-11** | **2019-05-17** | **115,128,420 원** | **15.130 %** |
| **중첩 지표** | **(B, R)** | **(B, M)** |  |  |
| **결과 – 바른손이앤에이** | | | | |
| Bollinger Band와 MACD를 기준으로 매수하고, Bollinger Band와 RSI를 기준으로 매도한 일자와 종가를 2019년 일별 종가 그래프에 표시했다. | | | | |
| **수익률 결과** | **매수** | **매도** | **결산** | **수익률** |
|  | **2019-04-01** | **2019-05-30** | **214,878,420 원** | **114.878 %** |
| **중첩 지표** | **(B, M)** | **(B, R)** |  |  |
| **주요 코드** | | | | |
| Plotyly의 graph\_objects를 이용해서 2019년 각 종목의 종가를 선 그래프로 표시하고, 2개 이상의 지표 기준을 통과한 시점을 marker로 표시했다. | | | | |

**Ⅲ. 결론**

* **2가지 지표 이상의 기준을 충족하는 시점에서 매수/매도했을 경우에, 단일 지표를 기준으로 했을 때에 비해서 수익률이 상승했다.**
* **3가지 보조 지표 중에서 항상 수익을 내는 지표는 없었지만, MACD의 수익률이 가장 편차가 컸다.**
* **보조 지표를 통해서 매매 신호를 확인했더라도 주가변동에 기인하는 외부적 요인이 많으므로 지표를 맹신할 수는 없다.**
* **Bollinger Band와 RSI 지표의 경우, 주가의 하락세를 매수 신호로 삼고 상승세를 매도 신호로 삼는다. 이는 종가의 이동 평균을 기반으로 지표 값이 계산되기 때문으로 보인다. 그러나 해당 지표들은 단기간의 이동 평균을 바탕으로 하기 때문에 상대적인 종가의 흐름을 나타낼 뿐 절대적인 종가를 나타내지는 않는다.**
* **MACD 지표의 경우, 주가가 횡보할 때 MACD선과 Signal선의 교차가 자주 일어나는 거짓 매매 신호가 빈번하게 확인됐다. 따라서 MACD 지표를 참고할 때는 RSI나 Bollinger Band등의 추세를 나타내는 지표를 함께 참고하는 것이 적절하다고 판단된다.**

**Ⅳ. 소감문**

|  |  |
| --- | --- |
| 성명 | 프로젝트 후 소감 |
| 허현준 | 주식 데이터를 분석 후 분석된 데이터에 근거해 주식을 매매해야 한다는 2번째 모듈 프로젝트 주제를 보고 주식에 대한 기본지식이 없어서 당황했고 따라서 데이터 분석의 방향과 분석에 근거한 인사이트 도출이 어떻게 이루어져야 하는지에 대한 시간 투자도 프로젝트 진행에 있어 상당 부분을 차지했다.  하지만 훌륭한 팀원들과 많은 의견을 나누니 프로젝트의 방향이 빨리 나왔고 무사히 마무리할 수 있었다.  처음 해보는 차후에 일어날 일을 과거 데이터분석을 통해 예측하는 데이터분석은 좋은 경험이었고 아쉬운 부분이 없다고는 말할 수 없다.  하지만 다른 조의 인상 깊었던 분석 방법들 중 하나인 주식시장에서 매매에 관한 가설을 데이터분석을 통한 결과 도출과 같은 다양한 접근법을 보고 배웠다는 것에 더 많은 가치를 둔다. |
| 최숙희 | 두번째 모듈 프로젝트는 첫 번째 프로젝트의 경험을 바탕으로, 비대면 상황이지만 팀원들과의 의견 공유와 분업이 잘 이루어졌다. 줌의 화면 공유 기능을 활용하여 서로 고민하는 부분이나 진행 상황을 공유하는 것도 어색하지 않아졌다.  데이터분석 과정에서 다양한 보조 지표를 활용하기 위해서 외부 데이터를 추가적으로 사용하게 되었는데, 코랩 등의 개발 환경에서 데이터분석을 하기 위해서는 외부데이터를 추가적으로 활용할 때마다 AWS의 MySQL에 테이블로 만드는 작업이 필수적이었다. 이를 위해서 execute()와 executemany()의 활용법에 대해 자세히 알아보게 되었다. 특히 executemany()의 경우, 많은 양의 데이터를 테이블에 추가해야하는 경우에, 데이터가 대입될 부분을 %s로 대체하면, %s부분들에 다중 입력이 가능하다는 점에서 아주 유용했다. 복잡한 SQL Query문을 작성하는 것 이외에 어떤 상황에서 활용할 수 있을지는 떠오르지 않지만 알아두면 좋을 개념을 익히게 된 것 같다.  필요한 테이블들을 구축한 후에도, 여러가지 조건들을 하나의 SQL Query문에 담기 위해서 JOIN을 이용하는 부분에서도 많은 고민을 하게 됐다. 수업시간에 받은 데이터와 외부 데이터의 종목 코드가 온전히 일치하지 않아서 수업시간에 자세히 배우지 않은 OUTER JOIN을 이용해야 하는 경우를 걱정했으나 다행히 수업시간에 배운 INNER JOIN만을 사용해도 조회할 수 있었다. 여러가지 조건을 중첩적용해서 데이터를 조회하는 SQL Query문을 고민하면서 SELECT, FROM, WHERE, HAVING, GROUP BY 그리고 ORDER BY 부분까지 각 쓰임새를 고민해 볼 수 있었다.  프로젝트 이전까지 주식과 보조 지표에 대한 개념 이해가 없던 탓에 시각화 결과물을 보면서 수익률과의 연관성을 찾아보는 것이 어려웠지만 수업시간 이외에도 팀원들과 함께 의논하며 결론을 찾아보는 노력을 한 점이 뿌듯하다. |
| 신종유 | 요즘 자산 증식의 일환으로 주식 투자가 많은 관심을 받고 있다. 나는 프로젝트 이전에는 주식에 문외한이었다. 하지만, 이번 기회를 통해 기초적인 주식 지식과 투자 방법에 대해 공부해봤고, 그로 인해 경제적인 시야가 넓어졌다고 생각하기에, 의미 있는 경험을 하였다고 생각한다. 또한, 이번 프로젝트로 파이썬을 통한 데이터 시각화를 다시 한번 연습해 볼 수 있어 도움이 되었다고 생각한다.  다소 아쉬운 점은 주가 변동 지표를 종합적으로 검토해 투자를 하기 어려웠다는 것이다. 주가 변동 지표의 매수/매도 결정일이 겹치는 시기가 없었다. 따라서 3가지 지표 중 2가지 지표만 고려해서 투자해야 했다. 또한 매수/매도일 각각에 적용하는 종합 지표가 서로 달랐다. 그래서 종합 지표를 검토한 이점을 찾기 쉽지 않았다. 다만 종합 검토시, 손실을 보지 않았다는 점에서 의미를 찾을 수 있을 것 같다. |
| 민병석 | 요즘 가장 핫한 재테크 방법인 주식을 분석한다는 것에 상당히 흥미로웠습니다.  주식을 조금 해본 주식 투자 초보자로서 `과연 주식의 오르내림을 예측할 수 있을까?`이라는 의문이 있었지만, 보조지표들을 검색해본 결과 많은 사람이 공식을 세웠고 분석 방법을 이미 만들어놨습니다.  그 보조지표들 중에 3가지를 참고하여 모의 투자를 해봤는데, 수익률이 나쁘지 않게 나왔습니다.  이 프로젝트로 주식 종목을 선정하는 과정에서 굉장히 유용하다는 생각이 들었습니다. 재정적으로는 안정적이고 저평가된 주식을 SQL을 통해 찾아냈고, 그 이후에 거래량이 많은 순서대로 나열했습니다.  이미 2019년의 자료이기 때문에 그 리스트에 나온 회사들은 현시점인 2021년에 보았을 때 그대로 투자했다면 모두 대박이 났을 것입니다. 보조지표들을 이용해서 단기적인 투자를 하는 것은 수익률이 날 확률이 높지만 좋은 수익률 기대하기는 힘들다는 생각이 들었고, 리스크가 있는 작업이라고 생각합니다. 왜냐하면 주식은 보조지표들 말고도 외부적인 뉴스나 트랜드를 따라가기 때문에 그 요소들은 수치로 환산하기도 어렵고, 워낙 변수가 많기 때문입니다.  예를 들어 아무도 코로나라는 전염병이 생길 거라 예상하지 못했습니다. 종합적으로 종목 선정하는 것은 아주 좋으나 단기 투자를 위해 분석하는 것은 다소 위험성이 있다고 판단되었습니다.  인간은 신이 아니기 때문에 가장 효율적인 투자는 좋은 종목을 찾아 저점에 매수한 뒤 더 이상 발전할 가능성이 없다고 판단 될 때 매도 하는 것이 가장 좋다고 생각합니다. |