[인사말]

시작하기에 앞서, 이렇게 본선에서 발표할 기회를 주셔서 감사하다는 말씀을 먼저 드리고 싶습니다. 저희 팀은 경북대학교 금융공학 동아리 DART에서 만나 퀀트투자에 높은 관심을 가지고 스터디를 진행하고 있습니다.

저는 감성 분석이 결합된 MVO전략, 심층 분석과 적용 가능성에 대해 발표하기로 한 최원준입니다. 발표시작 하겠습니다.

(다음장) [자산배분의 중요성 -> MVO의 단점]

금융시장은 다양한 요인에 의해 큰 **변동성**을 가지고 있으며 또한 일정수준 이상의 **불확실성을** 내포하고 있습니다. 이러한 불확실성 속에서 투자자들은 항상 어떻게 하면 안전하게 **초과수익률을** 올릴까? 고민하며 **최적의** **포트폴리오를** 구성하고자 노력합니다. 이러한 관점에서 자산배분 전략이란 **오랫동안** **연구되던** 주제입니다. 그중 마르코위츠의 Mean Variance Optimization 전략 이하, MVO 전략은 현대 자산배분 전략의 핵심으로 알려져 있습니다. **MVO**는 주어진 목표 수익률 하에서 위험을 최소화하는 전략으로 주어진 가정하에서 이론적으로 완벽한 자산배분 전략 중 하나입니다. 그러나, MVO가 성립하기 위해 필요한 많은 가정들은 **이것이 실제 투자에** 적용하기 힘들게 만듭니다. MVO가 가진 **문제점**들 중 저희가 주목한 점은 위험과 수익률 예측이 단순해 정확하지 않다는 것이고, 특정 자산에 대부분의 자산이 쏠려버리는 코너해문제, 시장의 동적 상황을 즉각적으로 반영하기 힘들다는 점이었습니다.

(다음장) [감성분석의 특징 -> 자산배분에 적용하는 기존 연구]

저희는 이러한 문제점들을 해결하고자 감성분석을 금융시장에 적용하고자 합니다. 감성분석은 회전율이 높고 잡음이 많은 대신 수익률 예측이 우수하다는 장점을 가지고 있습니다. 또한, 시장이 급변할 경우 시장의 동적상황을 빠르게 반영해 주곤 합니다. 이러한 장점을 활용하여 감성분석을 통한 자산배분 전략을 구현해 보고자 합니다. 감성분석을 활용한 자산배분 모델은 AI의 발전 이후 꾸준히 연구되는 주제였습니다. 예를 들어 국내에서는 2023년 김동재씨의 “**BERT를 활용한 뉴스 기사 감성분석과 블랙-리터만 모형을 결합한 자산 배분 전략 제안**” 같은 논문이 대표적으로 존재합니다.

(다음장) [기존 연구의 데이터 측면 문제점 => 데이터 취합 방식]

그런데, 저희가 생각했을 때 해당 전략에는 아쉬운 점이 존재했습니다. 금융 시장은 금리, CPI등 다양한 전문용어들이 존재합니다. 따라서 해당 용어들을 학습한 금융 특화 자연어 처리 모델인 Financial BERT가 나오기까지 이르렀습니다. **비슷한 방식으로** 철강, 건설, 헬스케어 섹터 혹은 **개별 주식들의 뉴스에서도** 전문용어들이 발생합니다. 이 **또한 특화 모델이** 필요함에도 불구하고 기본 Fin-BERT 모델을 사용하는 것은 적합하지 않다고 생각했습니다. 또한 각 전문섹터에 특화된 모델일지라 하더라도 이것이 **주가에 미치는 영향을 파악하는** 것은 **금융 도메인 내에서 감성분석을** 수행하는 것 보다 정확도가 떨어질 것이라 생각하였습니다. 따라서 저희는 금융 특화 모델이 존재하는 만큼 금융 데이터에서만 FinBERT 모델을 사용하는 방법을 고안했으며, 이를 기존 감성분석을 활용한 블랙리터만 자산배분 전략과 **비교해 보고자** 하였습니다. 금융데이터는 모두 크롤링을 통해서 수집하였습니다. 해당 데이터는 매일경제의 코스피 검색데이터, 한국경제의 코스피 태그데이터, 블롬버그에서 가져온 해외 뉴스 데이터를 사용하였습니다. 총 8년간 37,205개의 뉴스기사를 수집하였습니다. 각 뉴스 데이터의 예시는 다음과 같습니다. (짜잔 5초간 보여줌)

(다음장) [sentiment index 방법론]

KR Fin-BERT 모델의 성능을 **육안으로 관찰해본 결과** 실제 주가 상승에 영향을 줄만한 요인들은 긍정문으로, 주가 하락에 영향을 줄만한 요인들은 부정문으로, 그 외 문장들은 중립 문으로 판별한 예시를 확인해 볼 수 있었습니다. 위 데이터랑 KR-Fin-BERT 모델을 사용하여 각 뉴스별 sentimental score를 **다음 식과 같이** 정규화 하였습니다. 먼저 해당 뉴스를 문장별로 분할 한뒤 KR-FINBERT 모델을 통해 긍정문과 부정문 그리고 중립 문의 개수를 판별하였습니다. 그 후 해당 뉴스에서 긍정문 개수 총합 - 부정문 개수 총합 / 긍정과 부정문 개수의 총합으로 해당 뉴스의 sentiment score를 계산하였습니다. 그 후 월간 min max scaling을 한달마다 반복적으로 시행해 월간 sentiment index**를 최종적으로 계산하였습니다**.

(다음장) [그래프 잘나옴]

Sentimental index 그 자체는 잡읍이 **많아 사비츠키 골레이필터를** 사용하여 평탄화 시킨 뒤 그래프를 그려보았습니다. 육안으로 관찰해 보았을 때 코스피 주가에 선행하는 모습을 관찰할 수 있었습니다. 실제로 월간으로 데이터를 shift하면서 상관관계를 찍어본 결과 **주가와 상관관계가** 0.6까지 증가하여 매우 높은 상관관계를 보였습니다. 인과관계 분석을 진행하기 위해 각 데이터의 정상성을 확인해 보았습니다. ADF 검정 결과 코스피 종가와 감성 시계열 모두 95%유의 수준에서 정상성을 보임을 확인하였습니다. granger 인과검정 결과. 벤치마크로써 설정한 M2, BTC, sent-index, 경기선행지수를 사용한 결과보다 저희 sentimental-index가 인과관계로써 가장 유의한 지표임을 확인할 수 있었습니다.

(다음장) [sentiment index와 코스피 주가와의 관계]

이러한 데이터들을 토대로 저희는 **감성지수를 자산배분에 사용하기로** 결정하였습니다. 해당 데이터의 상관관계와 인과관계를 분석해본 결과 감성분석 결과가 실제 주가에 미치는 영향에는 일정 **시차가 존재함을** 확인할 수 있었습니다. 감성 지수가 각 시차별 미치는 영향을 **충격반응함수를 통해 계량**적으로 확인해보았습니다. 분석 결과, 시차 0에서 충격이 가장 컸음을 관찰할 수 있었으나 감성지수를 뽑아내는 것은 후행 적 지표이므로 시차1로부터의 충격을 관찰해 보았습니다. 관찰 결과 충격 **량이 각 시차별 75 88 88임**을 확인하였습니다. **각 시차별 미치는 영향이 유사하다는** 가정하에 감성지수가 주가에 미치는 영향을 월마다 0.3,0.35,0.35 정도로 가정하여 가중치를 설정하였습니다. 해당 가중치를 활용하여 **무위험 자산에 대한 투자 비중을** 조정하였다**. 즉,** 이전 1개월, 2개월, 3개월 전의 감성지수에 각각 30%, 35%, 35%의 가중치를 부여하여 종합 감성지수를 계산하였습니다. 이 종합 감성지수를 기반으로, 무위험 자산에 배분할 비중을 다음과 같이 설정하였습니다.

(다음장) [백테스트 결과]

실험은 백테스트를 통해 진행했습니다. Sentimental MVO에 사용될 **위험 자산군은** KRX 철강, 건설, 에너지 등 다양한 섹터에 존재하는 ETF를 토대로 구성하였습니다**. 무위험 자산** 수익률의 경우 연 3.5%로 설정하였습니다. **평가기**준은 샤프비율, MDD, 소르티노 비율, 정보비율, 수익률 등을 사용하여 평가하였습니다. 2018년부터 2024년까지 기본적인 MVO 전략을 활용했을 경우와 감성지수를 결합한 MVO 전략을 사용했을 경우를 비교하였습니다. **왼쪽 그래프를** 보시면 알 수 있듯이 하락장에서 **무위험 자산의 비중이 굉장히 높게 잡혀** 수익률을 효과적으로 방어할 수 있었으며, 상승장에서는 **위험자산의 비중이 높아** 수익률을 많이 확보한 것을 볼 수 있었습니다. 평가기준을 통해 기본 MVO전략과 비교해본 결과에서도 모든 지표에서 성능이 우수함을 확인할 수 있었습니다. 이러한 수치는 각 KRX index와 비교해 보았을 때도 우수한 성능을 보였습니다. 자세한 내용은 위 도표를 통해 확인 가능 합니다.

(다음다음장) [기존 연구와 비교]

또한 각 섹터별 감성지수를 사용하여 감성지수를 활용한 **블랙리터만 모델을 약식으로 구현해** 보았습니다. 해당 모델과 비교 하였을때도 상대/적으로 우수한 성과를 보였습니다. **블랙리터만의 파라미터 값은** 널리 사용되는값 혹은 MVO 전략에서 사용한 값과 동일하게 실험하였다. 무 위험 자산 수익률의 경우 3.5%로 가정하였으며 위험 회피계수는 2.5로 설정하였습니다. 시장 포트폴리오 가중치는 ETF를 사용한 관계로 모든 자산에 대해 동일 가중치를 사용하였으며 Ω이 값은 0.05로, τ의 값은 0.02로 설정하였습니다.

(다음장) [연구2 – 데이터 변경]

또한 금융 특화 데이터들 중 저희는 여러 데이터들을 병합하여 사용하였습니다. **각 데이터는** 코스피에 대한 직접적인 데이터, 코스피내 주식과 관련 있는 데이터, 해외 뉴스 데이터라는 **각기 다른 특징을** 가지고 있습니다. 어느 데이터가 자산배분에 활용가능성이 높은 지 파악하기 위해 각 기간 동안 데이터들을 변경해보면서 백테스트를 추가로 수행해보았습니다. 그 결과, 모든 데이터가 기본 MVO와 비교하였을 때, 자산배분에 사용하기 유의함을 알 수 있었습니다. 또한 세부 지표를 통해 판단해 본 결과 코스피와 직접적인 관련이 있는 한국경제의 코스피 태그 데이터가 자산배분에 활용하기 가장 적절하다는 점을 파악할 수 있었습니다.

(다음장) [시사점]

이러한 저희의 연구는 다음과 같은 **시사점**을 도출해 내었습니다. 먼저, MVO 수익률 예측이 단순하여 자산 배분 전략에 실증적으로 사용하기 부정확하다는 점을 감성점수로 인해 보완하여 활용할 수 있음을 수익률, 샤프비율, MDD 등 다양한 평가 지표를 통해 보여주었습니다. 또한 감성분석 데이터를 통해 특정 섹터나 주식에 치우치기보다 금융 도메인과 직접적 관련이 있는 데이터를 사용하는 것이 더 좋은 성과를 보여줌을 평가지표들을 통해 시사하였습니다. 마지막으로, 블랙리터만에 감성지수를 활용하는 것 보다 코스피 데이터를 활용하여 무 위험 자산과 위험 자산 사이의 투자 비중을 조절하는 것이 자산배분에서 더 효과적임을 보여주었습니다.

(다음장) [인사말]

이상으로 감성 분석이 결합된 MVO 전략 : 심층 분석과 적용 가능성에 대한 발표를 모두 마치겠습니다. 긴 시간 동안 집중하여 들어주셔서 **감사**드립니다. (꾸벅)