

2024 NH투자증권 빅데이터 경진대회 (예선)

팀명					
팀원	성명	생년월일	학교	학과	연락처
주제명					

분석 보고서

1. **분석 목적 및 목표**

이 분석의 목적은 NH투자증권 데이터와 미국 주식/ETF 시세 정보를 활용하여 투자자들에게 유용한 인사이트를 도출하는 것입니다. 주요 목표는 다음과 같습니다:

- 산업 내 경쟁자 간 상관관계를 도출하여 시장 내의 영향력을 평가합니다.
- 고객의 공포와 관심도를 파악하여 시장 참여도를 분석합니다.
- 주식의 유동성 및 투자 성과를 단기 및 장기 관점에서 평가합니다.
- 회사 및 투자 전문가의 투자 활동을 분석하여 산업별로 어떤 종목에 투자하는지 파악합니다.

2. **제공된 데이터**

- NH MTS (모바일 트레이딩 시스템) 데이터: 고객 활동, 매수/매도 정보, 회사 보유 현황을 포함합니다.
- 미국 주식 및 ETF 시세 정보: 개별 주식 및 ETF에 대한 세부 구성 데이터를 포함합니다.

주요 데이터 필드:

- CUS_CGR_LLF_CD (고객 구성 대분류 코드) 및 CUS_CGR_MLF_CD (고객 구성 중분류 코드): 투자자 프로파일 구분.
- RGS_TGT_CNT (관심 종목 등록 수) 및 VW_TGT_CNT (종목 조회 수): 사용자 관심도.
- LSS_IVO_RT (손실 투자자 비율) 및 PFT_IVO_RT (수익 투자자 비율): 투자 성과 지표.
- IFW_ACT_CNT (신규 매수 계좌 수) 및 OFW_ACT_CNT (전량 매도 계좌 수): 매수/매도 트렌드 파악.

3. **분석 방법론**

1. **투자자 행동 분석**:

- 평균 보유 기간(TCO_AVG_HLD_TE_DD_CNT)을 분석하여 투자자들이 장기 혹은 단기 전략을 취하는지 파악합니다.
- 고객 세그먼트 분석: CUS_CGR_LLF_CD와 CUS_CGR_MLF_CD를 활용하여 투자자의 실력, 나이, 자산 규모에 따른 행동 패턴을 분석합니다.

2. **유동성 및 변동성 분석**:

- VW_TGT_CNT(종목 조회 수)와 RGS_TGT_CNT(관심 종목 등록 수)를 분석하여 유동성을 평가합니다.
- MM1_TOT_PFT_RT, MM3_TOT_PFT_RT, YR1_TOT_PFT_RT 데이터를 활용하여 단기(1개월), 중기(3개월), 장기(1년) 성과를 분석합니다.

3. **시장 심리 분석**:

- LSS_IVO_RT(손실 투자자 비율) 및 PFT_IVO_RT(수익 투자자 비율)를 통해 시장 심리를 분석하고, 두 비율을 통해 공포와 낙관주의 트렌드를 파악합니다.
- IFW_ACT_CNT(신규 매수 계좌)와 OFW_ACT_CNT(전량 매도 계좌)를 분석하여 투자자들이 시장에 진입하는지, 혹은 이탈하는지를 파악합니다.

4. **전문가 투자 추적**:

- 고급 투자자(CUS_CGR_LLF_CD = "1")의 행동을 추적하고, 이들의 투자 활동이 시장 성과에 미치는 영향을 분석합니다.

4. **서비스 아이디어**

1. **산업 상관관계 및 경쟁자 추적 서비스**:

- 동일 밸류 체인 내에서 경쟁자 간의 상관관계를 추적하는 서비스를 제공합니다. 주식 성과 및 투자자 활동 데이터를 활용하여 시장 점유율과 경쟁자 반응을 파악합니다.

2. **고객 심리 대시보드**:

- 손실/수익 비율과 계좌 유입/유출 데이터를 기반으로 시장의 공포와 관심도를 시각화하는 대시보드를 제공합니다.

3. **유동성 및 투자 유형 평가**:

- 종목의 유동성과 장단기 투자 적합성을 실시간으로 평가하는 서비스를 개발합니다. VW_TGT_CNT 및 RGS_TGT_CNT와 거래량 데이터를 결합하여 제공합니다.

4. **전문가 투자 추적 서비스**:

- 특정 산업 내에서 전문 투자자들의 행동을 추적하는 기능을 제공하여, 일반 투자자들이 전문가 전략을 분석하고 따를 수 있게 합니다.

5. **결과 해석**

- 상관관계 분석: 경쟁자 간의 상관관계를 도출하여 경쟁 또는 협력 관계를 평가합니다.
- 시장 심리 인사이트: 손실/수익 비율을 통해 시장의 흐름을 예측하고, 향후 시장의 상승 혹은 하락 가능성을 분석합니다.
- 유동성 트렌드: 관심도와 거래량 증가를 반영하여 투자에 적합한 종목을 도출합니다.

서비스 기획 아이디어 및 발전 방향

3. 향후 서비스 발전 방향 및 아이디어 앞으로의 분석 발전 방향은 최신 기술과 연구 논문을 구현하는 것에 중점을 둡니다. 특히, 두 가지 연구 논문을 활용하여 분석 역량을 강화하고자 합니다.
4. 서비스 아이디어
5. 산업 상관관계 및 경쟁자 추적 서비스:
 - 동일 밸류 체인 내에서 경쟁자 간의 상관관계를 추적하는 서비스를 제공합니다. 주식 성과 및 투자자 활동 데이터를 활용하여 시장 점유율과 경쟁자 반응을 파악합니다.
5. 고객 심리 대시보드:
 - 손실/수익 비율과 계좌 유입/유출 데이터를 기반으로 시장의 공포와 관심도를 시각화하는 대시보드를 제공합니다.
6. 유동성 및 투자 유형 평가:
 - 종목의 유동성과 장단기 투자 적합성을 실시간으로 평가하는 서비스를 개발합니다. VW_TGT_CNT 및 RGS_TGT_CNT와 거래량 데이터를 결합하여 제공합니다.
7. 전문가 투자 추적 서비스:
 - 특정 산업 내에서 전문 투자자들의 행동을 추적하는 기능을 제공하여, 일반 투자자들이 전문가 전략을 분석하고 따를 수 있게 합니다.
8. 결과 해석
 - 상관관계 분석: 경쟁자 간의 상관관계를 도출하여 경쟁 또는 협력 관계를 평가합니다.
 - 시장 심리 인사이트: 손실/수익 비율을 통해 시장의 흐름을 예측하고, 향후 시장의 상승 혹은 하락 가능성을 분석합니다.
 - 유동성 트렌드: 관심도와 거래량 증가를 반영하여 투자에 적합한 종목을 도출합니다.

향후 서비스 발전 방향 및 아이디어 앞으로의 분석 발전 방향은 최신 기술과 연구 논문을 구현하는 것에 중점을 둡니다. 특히, 두 가지 연구 논문을 활용하여 분석 역량을 강화하고자 합니다.

1. Scaling Core Earnings Measurement with Large Language Models:
 - 대형 언어 모델(LLM)을 사용하여 미래의 핵심 수익과 순이익을 평가하는 기술을 도입할 계획입니다. 이를 통해 Value Chain과 Supply Chain 데이터를 기반으로 기업의 수익성을 보다 정확하게 분석할 수 있습니다.
2. Text2SQL is Not Enough: Unifying AI and Databases with TAG:
 - 데이터베이스 쿼리와 LLM 기반 추론을 결합하는 방식을 제시하는 논문을 적용하여, 데이터 분석의 효율성을 극대화할 수 있습니다. 밸류체인 및 순이익 평가에 있어 보다 정교한 분석과 NH 데이터에 관한 고급 쿼리 및 시각화 자동화가 가능해집니다.

이와 같은 기술적 발전을 통해 서비스 아이디어에 다음과 같은 기능을 추가할 수 있습니다:

1. 산업 상관관계 및 경쟁자 추적 서비스:
 - 대형 언어 모델을 활용하여 산업 내 경쟁자 간 상관관계를 보다 심층적으로 분석하고, 밸류체인 및 서플라이 체인 데이터를 기반으로 시장 점유율과 경쟁자 반응을 추적하는 서비스를 제공합니다.
2. 고급 AI 기반 고객 심리 대시보드:
 - AI 기반 쿼리 시스템을 통해 고객 심리를 실시간으로 분석하여, 손실/수익 비율과 계좌 유입/유출 데이터를 자동화된 방식으로 시각화하는 기능을 제공합니다. 이를 통해 시장의 공포와 관심도를 더 정확하게 파악할 수 있습니다.
3. AI-driven 유동성 및 투자 유형 평가 서비스:
 - AI 기반 데이터 추론을 활용하여 종목의 유동성과 장단기 투자 적합성을 실시간으로 분석하고, VW_TGT_CNT 및 RGS_TGT_CNT 데이터를 결합하여 보다 정확한 유동성 평가와 예측을 제공합니다.
4. 전문가 투자 추적 서비스의 고도화:
 - 전문가 투자에 대한 심층 분석을 통해 전문 투자자들의 전략을 추적하고, 그들의 행동이 시장에 미치는 영향을 예측하는 기능을 제공합니다. AI 기반 데이터 분석을 통해 투자자들이 보다 스마트한 의사결정을 할 수 있도록 돕습니다.

이러한 발전 방향을 통해 서비스는 기존 분석의 한계를 극복하고, 더 나아가 AI와 데이터베이스를 결합한 고도화된 분석 서비스를 제공할 수 있을 것입니다.