

이슈보고서 25-21

ISSUE  
REPORT

# AI와 자산운용: 패러다임의 전환 가능성 고찰

---

권민경 · 남재우

# AI와 자산운용: 패러다임의 전환 가능성 고찰

저자 권민경·남재우\*

인공지능(Artificial Intelligence: AI) 기술, 특히 트랜스포머 아키텍처와 대규모 언어 모델(Large Language Model: LLM)의 비약적 발전은 자산운용 산업의 근본적인 패러다임 전환 가능성을 나타내고 있다. 이러한 변화의 흐름은 크게 두 가지 방향으로 전개되고 있는데, 첫 번째는 LLM의 추론 능력을 활용하여 인간 전문가 팀의 협업적 의사결정 과정을 모방하고 자동화하는 ‘멀티 에이전트 시스템’ 접근법이다. 이 방식은 방대한 비정형 데이터를 종합하고 투자 논거를 제시하는 ‘설명 가능성’을 확보하는 데 중점을 둔다. 두 번째는 트랜스포머 아키텍처를 금융 시계열과 같은 정형 데이터에 직접 적용하여 예측 성능 자체를 극대화하려는 시도이다. 이는 ‘복잡성의 미덕(The Virtue of Complexity)’ 등 새로운 이론적 배경을 바탕으로, 인간의 직관이나 전통적 계량 모델의 한계를 넘어서는 복잡한 비선형 패턴을 발견하는 것을 목표로 한다.

이러한 AI 기반 패러다임 전환기에 자산운용사가 지속 가능한 경쟁력을 확보하기 위한 세 가지 핵심 과제는 다음과 같다. 첫째, AI 모델의 성능을 좌우할 핵심 자산으로서 양질의 데이터를 확보하고 관리하는 체계적인 역량을 구축하는 것이다. 둘째, 개별 고객의 복잡한 재무 목표, 위험 성향, 생애 주기 등을 반영하여 동적으로 운용 전략을 조정하는 ‘초개인화된 투자 솔루션’을 구축하는 것이다. 셋째, 알고리즘의 예측 실패 가능성, 모델의 불투명성, 동일한 인기 모델로의 행태 수렴 등 AI 고유의 새로운 위험 요인을 선제적으로 통제하고 기술에 대한 신뢰를 확보할 수 있는 강력한 내부 거버넌스를 확립해야 한다. 이 세 가지 과제는 AI 기술 도입의 성패를 가늠하는 핵심 기준으로 작용할 것으로 판단된다.

\* 본고의 견해와 주장은 필자 개인의 것이며, 자본시장연구원의 공식적인 견해가 아님을 밝힙니다.

펀드·연금실 연구위원 권민경(mkkwon@kcmi.re.kr) / 펀드·연금실 선임연구위원 남재우(zeuss@kcmi.re.kr)

\*\* 발행: 2025년 12월 1일

## I. 연구 배경

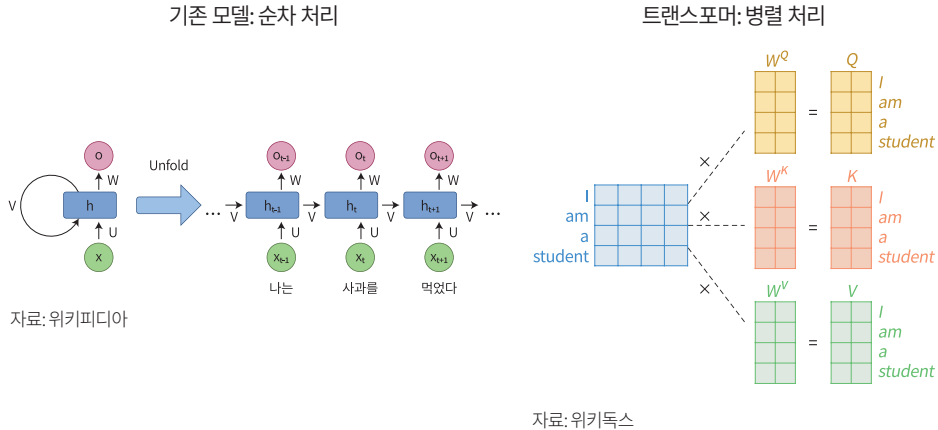
---

2022년 ChatGPT의 등장은 인공지능(Artificial Intelligence: AI) 기술이 특정 전문 영역을 넘어 대중의 일상과 산업 전반에 실질적인 영향을 미칠 수 있음을 보인 분기점으로 평가된다. 이전의 AI 모델이 특정 작업 수행에 국한된 성능을 보인 것과 달리, 대규모 언어 모델(Large Language Model: LLM)은 인간의 언어를 깊이 있게 이해하고 생성하는 능력을 기반으로 범용적인 과제 수행 능력을 선보이며 새로운 가능성을 입증하였다.

이러한 발전의 중심에는 Vaswani et al.(2017)이 제안한 ‘트랜스포머(Transformer)’라는 혁신적인 신경망 아키텍처가 자리하고 있다. 트랜스포머 이전의 자연어 처리 모델들은 주로 순환 신경망(Recurrent Neural Network: RNN)과 같은 순차적(sequential) 구조에 의존하였다. 이 모델들은 문장의 단어를 순서대로 하나씩 처리하는 방식으로 인해, 문장이 길어질 경우 초기 정보가 소실되는 ‘장기 의존성 문제(long-range dependency problem)’를 내재하고 있었다. 또한, 데이터 처리 과정의 순차적 특성은 대규모 병렬 컴퓨팅 자원의 효율적 활용을 저해하여, 모델의 학습 속도와 규모 확장에 근본적인 제약으로 작용하였다.

트랜스포머는 ‘셀프 어텐션(self-attention)’이라는 독창적인 메커니즘을 도입하여 이러한 순차적 구조의 한계를 근본적으로 해결하였다. 셀프 어텐션은 문장 내 모든 단어 간의 관계와 문맥 상 중요도를 동시에 병렬적으로 계산하여, 특정 단어가 문장 내 다른 모든 단어에 어느 정도 ‘집중(attention)’해야 하는지를 학습한다(<그림 1-1> 참고). 예를 들어, “The animal didn’t cross the street because it was too tired.”라는 문장에서 ‘it’의 의미를 파악하기 위해, 모델은 ‘animal’, ‘street’, ‘tired’ 등 문장 전체의 구성 요소를 동시에 고려하여 ‘it’이 ‘animal’을 지칭할 확률이 가장 높음을 추론한다. 이는 순서에 의존하지 않고 단어 간의 복잡한 문맥 관계를 직접 파악하는 방식이다.

&lt;그림 1-1&gt; 트랜스포머에서 모델 규모 확장의 용이성



셀프 어텐션이 가진 핵심적 특징인 병렬 처리 가능성은 현대적 컴퓨팅 하드웨어, 특히 그래픽 처리 장치(Graphics Processing Unit: GPU)의 성능을 극대화하는 기반이 되었다. RNN과 같은 기존 모델들이 데이터를 순차적으로 처리해야 했던 것과 달리, 트랜스포머는 문장 전체 또는 대규모 데이터 배치(batch)를 한 번에 연산할 수 있다. 이러한 아키텍처의 효율성은 모델의 학습 시간을 획기적으로 단축시켰으며, 이전에는 불가능했던 수준의 대규모 모델로 확대할 수 있는 ‘확장성(scalability)’을 확보하는 결정적 계기가 되었다.

트랜스포머 아키텍처가 제공한 모델 규모 확장의 용이성은 LLM의 파라미터(parameter) 수가 폭발적으로 증가하는 결과로 이어졌다. 2017년 Vaswani et al.(2017)의 기본 모델이 6,500만 개의 파라미터를 사용한 데 비해, 2020년 GPT-3는 1,750억 개, 그리고 2023년 공개된 GPT-4는 약 1조 7,600억 개의 파라미터를 보유한 것으로 추정된다. 불과 6년여 만에 약 2.7만 배 증가한 이러한 모델의 양적 팽창은 단순한 성능의 점진적 향상을 넘어, 이전까지 예측하지 못했던 질적인 도약을 야기했으며, Wei et al.(2022)은 이를 ‘창발적 능력(Emergent Abilities)’이라 명명하였다. 창발적 능력이란, 모델의 규모가 특정 임계점을 초과하면서 작은 모델에서는 관찰되지 않던 새로운 능력이 갑자기 발현되는 현상을 의미하며, 이는 모델 성능이 규모에 따라 선형적으로 증가(linear scaling)하는 것이 아니라 특정 지점에서 비선형적인 도약을 보임을 시사한다.

창발적 능력의 대표적인 사례로는 복잡한 ‘추론 능력’과 ‘범용 과제 수행 능력’이 있다. 즉, 소규모 모델은 단순 질의응답 수준에 그치는 반면, 거대 모델은 복잡한 문제 해결을 위해 스스로 중간 사고 과정을 생성하는 ‘추론 능력’을 보인다. 또한, 별도의 학습 없이도 제시된 몇 가지 예시만으로

새로운 과제를 즉각적으로 수행하는 ‘범용 과제 수행 능력’을 발휘한다. 이러한 능력의 발현은 모델의 규모 확장이 단순히 더 많은 정보를 암기하는 과정을 넘어, 데이터에 내재된 근본 원리를 학습하고 이를 유연하게 적용하는 능력을 획득하는 과정임을 시사한다. ChatGPT는 바로 이러한 LLM의 창발적 능력이 구체적인 서비스로 구현되어 그 잠재력을 대중에게 증명한 대표적인 사례이며, AI 기술이 산업 전반의 패러다임을 전환할 수 있다는 인식을 확산시키는 결정적 계기가 되었다.

트랜스포머 아키텍처의 뛰어난 확장성과 범용성은 자연어 처리 영역을 넘어 타 과학 및 산업 분야로 빠르게 확산되는 추세다. 대표적인 사례로, 구글 브레인(Google Brain)의 비전 트랜스포머(Vision Transformer)는 이미지를 여러 개의 작은 패치(patch)로 분할하여 이를 언어의 토큰(token) 시퀀스처럼 처리하는 방식을 채택하였다(Dosovitskiy et al., 2021). 이는 기존의 합성곱 신경망(Convolutional Neural Network: CNN)이 지배적이던 컴퓨터 비전 분야에 새로운 접근법을 제시하며 높은 성능을 달성한 사례다. 또한, 구글 딥마인드(Google DeepMind)의 알파폴드(AlphaFold)는 트랜스포머 구조를 단백질 서열 데이터에 적용하여, 50년 이상 과학계의 난제로 남아있던 단백질 구조 예측 문제를 상당 부분 해결하는 쾌거를 이루었다(Jumper et al., 2021).

이처럼 상이한 도메인(domain)으로의 확장이 가능한 근본적인 이유는, 트랜스포머의 핵심인 ‘셀프 어텐션’ 메커니즘이 본질적으로 ‘범용적인’ 패턴 인식 기능을 갖기 때문이다. 셀프 어텐션은 ‘어떤 요소가 다른 요소에 얼마나 집중해야 하는가’를 학습하는 과정으로, 데이터의 종류에 구애받지 않는다. 즉, 이 메커니즘은 언어 토큰 간의 관계뿐만 아니라, 이미지의 패치 간 공간적 관계나 단백질 서열의 아미노산 간의 구조적 관계와 같이 다양한 데이터 요소 간에 내재된 복잡하고 비직관적인 상호작용을 효과적으로 파악할 수 있다. 이는 트랜스포머가 특정 도메인에 종속되지 않는 범용 아키텍처로서의 잠재력을 보유하고 있음을 증명한다.

트랜스포머 아키텍처의 이러한 범용성은 궁극적으로 과학 연구 패러다임의 근본적인 전환을 예고한다. 기존의 과학 연구가 주로 인간의 직관과 이론에 기반한 ‘가설 검증(hypothesis-driven)’ 방식에 의존했다면, 향후에는 AI가 대규모 데이터 자체에서 유의미한 패턴을 직접 발견하는 ‘데이터 기반 발견(data-driven discovery)’ 방식으로 무게 중심이 이동할 가능성이 높다. 특히 복잡한 문제에서는 유효한 가설을 수립하고 검증하는 과정 자체가 가장 큰 난관인 경우가 많다. 이러한 상황에서 AI는 방대한 데이터에 내재된 통계적 패턴을 자동으로 식별함으로써 전통적인 가설 설정 단계를 우회하거나 보완하며, 연구 개발의 효율성을 극대화하는 핵심 동력으로 작용할 것으로 전망된다(Fudan University et al., 2025).

이러한 패러다임 전환은 자산운용업에도 동일하게 적용된다. 자산운용업의 본질적 가치는 고객의 투자 목표에 부합하는 포트폴리오를 구축하고, 이를 통해 시장 대비 초과수익을 창출하는데 있다. 이러한 가치 창출 과정은 크게 두 가지 핵심 활동으로 구성된다. 첫째, 다양한 자산에 대한 심층적인 분석을 통해 투자 가치를 평가하고, 둘째, 이러한 분석 결과를 바탕으로 위험과 수익의 균형을 고려하여 최적의 포트폴리오를 구성하는 것이다. 전통적으로 이러한 과정은 인간 전문가의 경험과 직관, 그리고 전통적인 계량 모델에 의존해왔다. 그러나 금융 시장의 복잡성이 증대되고 데이터의 양과 다양성이 폭발적으로 증가하는 새로운 환경에서는, 인간의 인지적 한계와 기존 모델의 구조적 제약이 자산운용업의 핵심 가치 창출 능력을 제약하는 요인으로 부상하고 있다.

트랜스포머 기술의 발전은 자산운용업의 핵심 가치 창출 과정에 근본적인 변화를 가져올 잠재력을 지닌다. 특히 트랜스포머 아키텍처가 제공하는 두 가지 핵심 특성은 자산운용업의 본질적 과제와 직접적으로 연결된다. 첫째, LLM의 ‘창발적 능력’은 인간의 언어로 구성된 방대한 비정형 데이터를 이해하고 추론함으로써, 기업 공시, 애널리스트 리포트, 뉴스 기사 등 다양한 정보원을 종합적으로 분석하여 투자 가치를 평가하는 인간 전문가의 복잡한 의사결정 과정을 보조하거나 자동화할 수 있는 가능성을 열었다. 둘째, 트랜스포머 아키텍처의 ‘범용성’과 ‘셀프 어텐션’ 메커니즘은 언어의 영역을 넘어 금융 시계열, 기업 특성 데이터, 그리고 자산 간 상관관계와 같은 정형 데이터에도 직접 적용될 수 있으며, 이를 통해 포트폴리오 구성에 필요한 복잡한 자산 간 상호작용과 동적 패턴을 학습할 수 있다. 이는 인간의 직관이나 전통적인 계량 모델의 한계를 뛰어넘어, 시장 대비 초과수익을 창출할 수 있는 새로운 패턴을 발견할 잠재력을 시사한다.

본고에서는 이러한 패러다임 전환의 가능성을 두 가지 주요 흐름으로 구분하여 심층적으로 고찰한다. 우선 II장에서는 인간 전문가의 지식과 분석 워크플로우를 모방하고 자동화하는 ‘LLM 활용 투자 모델’을 다룬다. 다음으로 III장에서는 금융 데이터 자체에 내재된 패턴을 직접 탐색하는 ‘금융 특화 트랜스포머 모델’을 살펴본다. 이를 바탕으로 IV장에서는 AI 기술의 발전에 대응하여 국내 자산운용업계가 미래 경쟁력 확보를 위해 나아가야 할 방향성을 모색하고자 한다.

## II. LLM을 활용한 투자 모델

앞서 I장에서 논의한 바와 같이, 자산운용업의 핵심 가치는 고객의 투자 목표에 부합하는 포트폴리오를 구축하고 시장 대비 초과수익을 창출하는 데 있다. 이러한 가치 창출 과정의 첫 번째 단계인 ‘투자 가치 평가’는 기업 공시, 애널리스트 리포트, 뉴스 기사, 산업 동향 분석 등 방대한 비정형 데이터를 종합적으로 분석하여 개별 자산의 투자 가치를 판단하는 작업이다. 자산운용 분야에서 AI를 활용하는 첫 번째 주요 흐름은 LLM의 강력한 언어 이해 및 추론 능력을 바탕으로 이러한 투자 가치 평가 과정을 수행하는 인간 금융 전문가 팀의 의사결정을 모방하고 자동화하는 것이다. 이 접근법의 핵심은 AI를 단순한 데이터 처리 도구가 아닌, 분업과 협업을 통해 복잡한 문제를 해결하는 ‘자동화된 전문가 팀’으로 간주하는 데 있다. 또한, 예측의 정확성만큼이나 결정의 근거를 투명하게 제시하는 ‘설명 가능성(explainability)’을 중시하는 특징을 보인다. 이는 자산운용사가 투자 결정의 책임을 지고 고객이나 규제 당국에 투자 근거를 설명해야 하는 업계의 특성과 직접적으로 부합한다.

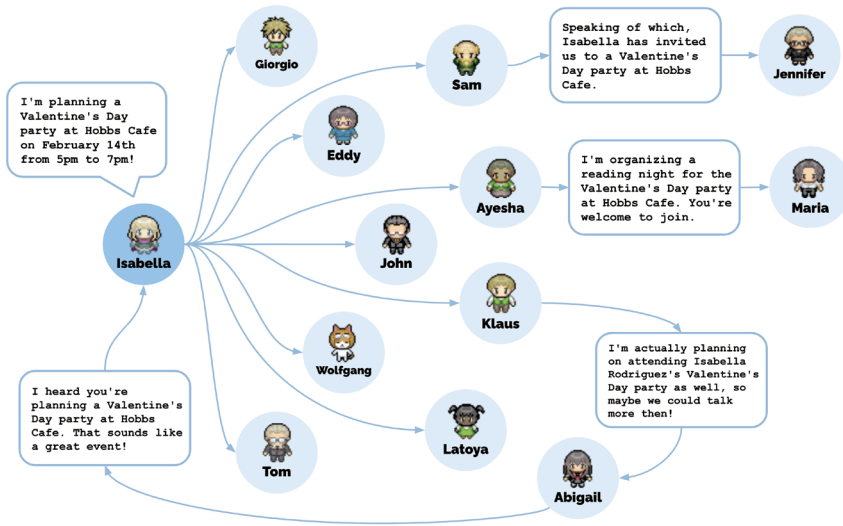
### 1. 멀티 에이전트 시스템

전통적인 투자 결정 과정은 다양한 전문성을 가진 인력들의 협업을 기반으로 한다. 재무제표와 산업 동향을 분석하는 펀더멘털 애널리스트, 주가 차트와 거래량 패턴을 연구하는 기술적 분석가, 그리고 계량 모델을 통해 위험 요인을 평가하는 퀀트 애널리스트 등이 각자의 분석 결과를 공유하고 토론하며 종합적인 투자 결론에 도달한다. LLM을 활용한 멀티 에이전트 시스템(multi-agent system)은 이와 같은 인간 전문가 팀의 협업 워크플로우를 소프트웨어적으로 구현하고 그 효율성을 극대화하려는 시도이다.

이러한 멀티 에이전트 시스템의 구현 가능성은 LLM 에이전트 간의 유의미한 상호작용이 가능함을 보인 실험적 연구들을 통해 뒷받침된다. 대표적으로 Park et al.(2023)의 연구는 가상의 마을 환경에 카페 주인, 화가, 교수 등 각기 다른 역할을 부여받은 LLM 에이전트들을 배치하였다(<그림 II-1> 참고). 그 결과, 에이전트들은 단순한 명령 수행을 넘어 자율적으로 소통하고 협력하며 마치 살아있는 사회처럼 상호작용하는 모습이 관찰되었다. 예를 들어, 한 에이전트가 선거 출마 계획을 알리자 관련 정보가 다른 에이전트들에게 확산되며 토론이 이루어졌고,

또 다른 에이전트가 파티를 기획하자 일부는 자발적으로 준비를 돕고 지인을 초대한 반면, 일부는 개인 사정을 설명하며 불참을 통보하기도 하였다. 이는 LLM 에이전트가 단순 정보 처리 기능을 넘어, 공동의 목표나 사회적 맥락을 이해하고 유기적으로 협업할 잠재력을 가졌음을 시사한다.

<그림 II-1> 멀티 에이전트 시스템의 실험 사례



자료: Park et al.(2023)

자산운용과 같이 고도로 복잡한 영역에서 멀티 에이전트 시스템을 채택하는 것은 단일 LLM이 가진 본질적 한계를 극복하기 위한 전략적 선택이다. 여기에는 두 가지 주요 이점이 있다. 첫째, 전문성의 희석을 방지하고 분석의 깊이를 더할 수 있다. 단일 LLM이 모든 전문가의 역할을 동시에 수행하려 할 때 발생할 수 있는 분석의 피상성을 극복하고, ‘핀더멘털 분석가’, ‘기술적 분석가’ 등 명확히 구분된 역할을 개별 에이전트에 부여하여 전문성을 극대화하는 것이다. 둘째, 방대한 데이터를 동시에 처리할 때 발생하는 기술적 제약을 해결할 수 있다. LLM은 한 번에 처리할 수 있는 정보의 총량, 즉 ‘컨텍스트 윈도우(context window)’의 물리적 한계를 갖는데, 멀티 에이전트 시스템은 방대한 분석 작업을 여러 에이전트에게 병렬적으로 분배함으로써 이 한계를 우회한다.

실제 자산운용 분야의 멀티 에이전트 시스템은 각기 다른 분석 역할을 부여받은 다수의 LLM 에이전트가 독립적으로 정보를 수집, 처리하고 그 결과를 종합하는 형태로 작동한다. 이는



자산운용사가 수행하는 핵심 업무인 ‘투자 가치 평가’ 과정을 구조화하고 자동화하는 데 기여한다. 예를 들어, 특정 기업을 분석하는 과정에서 ① ‘공시 분석 에이전트’는 연간 보고서, 분기 실적 등 정기 공시 자료를 분석하여 재무 건전성과 잠재적 위험 요인을 식별한다. 동시에 ② ‘뉴스 분석 에이전트’는 실시간 뉴스 기사와 소셜 미디어 게시물을 분석하여 시장 심리 및 기업 평판을 평가하며, ③ ‘기술적 분석 에이전트’는 과거 주가 및 거래량 데이터를 기반으로 지지선, 저항선 등 기술적 신호를 포착한다. 이러한 각 에이전트의 분석 결과는 궁극적으로 해당 자산이 포트폴리오에 포함될 가치가 있는지, 그리고 포함된다면 적정 비중은 어느 정도인지를 판단하는 데 활용된다.

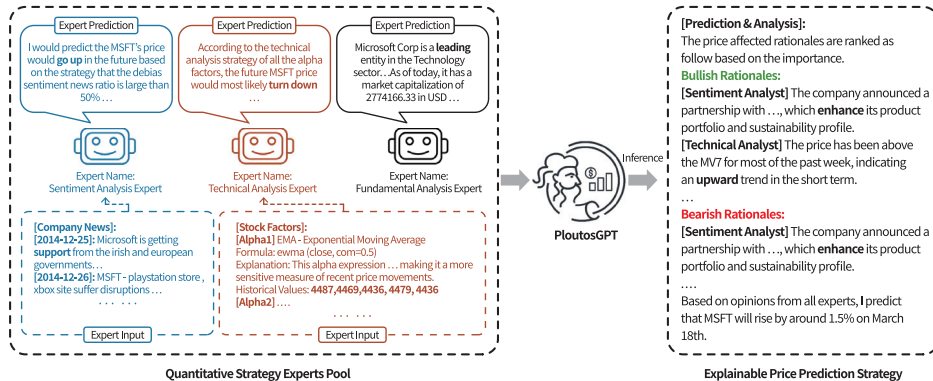
이렇게 각 전문 에이전트가 독립적으로 생성한 분석 결과는 최종적으로 ‘마스터 에이전트 (Master Agent)’로 전달되어 종합된다.<sup>01</sup> 마스터 에이전트는 자산운용사의 투자위원회나 포트폴리오 매니저가 다양한 전문가의 의견을 종합하여 최종 투자 결정을 내리는 과정과 유사하게, 개별 에이전트들로부터 전달받은 상충되거나 보완적인 정보들을 비교하고 이들 간의 논리적 연관성을 파악한다. 이를 바탕으로 최종적인 투자 판단(예: 매수/매도/보유)과 그에 대한 종합적인 근거를 자연어로 생성한다. 이처럼 멀티 에이전트 시스템의 활용은 LLM의 창발적 능력을 구조화된 협업 프레임워크에 적용함으로써, 자산운용업의 핵심 가치 창출 과정인 투자 가치 평가의 효율성과 정확성을 향상시키고, 이를 통해 궁극적으로 초과수익 창출 가능성을 높이려는 시도라고 할 수 있다.

## 2. 적용 사례

멀티 에이전트 시스템의 원리를 실제 투자 모델에 적용한 사례들은 이 접근법의 구체적인 잠재력을 보여준다. 대표적인 예로 Tong et al.(2024)이 제안한 ‘PloutosGPT’를 들 수 있다. 이 프레임워크는 감성 분석, 기술적 분석, 펀더멘털 분석 등 명시적인 역할을 부여받은 개별 에이전트로 구성된다. 각 전문 에이전트는 기업 뉴스, 주가 데이터, 재무 지표와 같이 자신에게 할당된 고유의 데이터를 독립적으로 분석하여 의견을 도출한다. 예를 들어, 동일한 시점에 ‘감성 분석 에이전트’는 긍정적 뉴스를 근거로 매수 의견을 제시하는 반면, ‘기술적 분석 에이전트’는 이동평균선 등 기술적 지표를 바탕으로 주가 하락 가능성을 예측하는 등 상이한 분석을 수행할 수 있다.

01 여기에서는 편의상 ‘마스터 에이전트’로 칭하지만, 모델에 따라 ‘리드 에이전트’, ‘시그널 에이전트’ 등 다양한 명칭으로 불릴 수 있다.

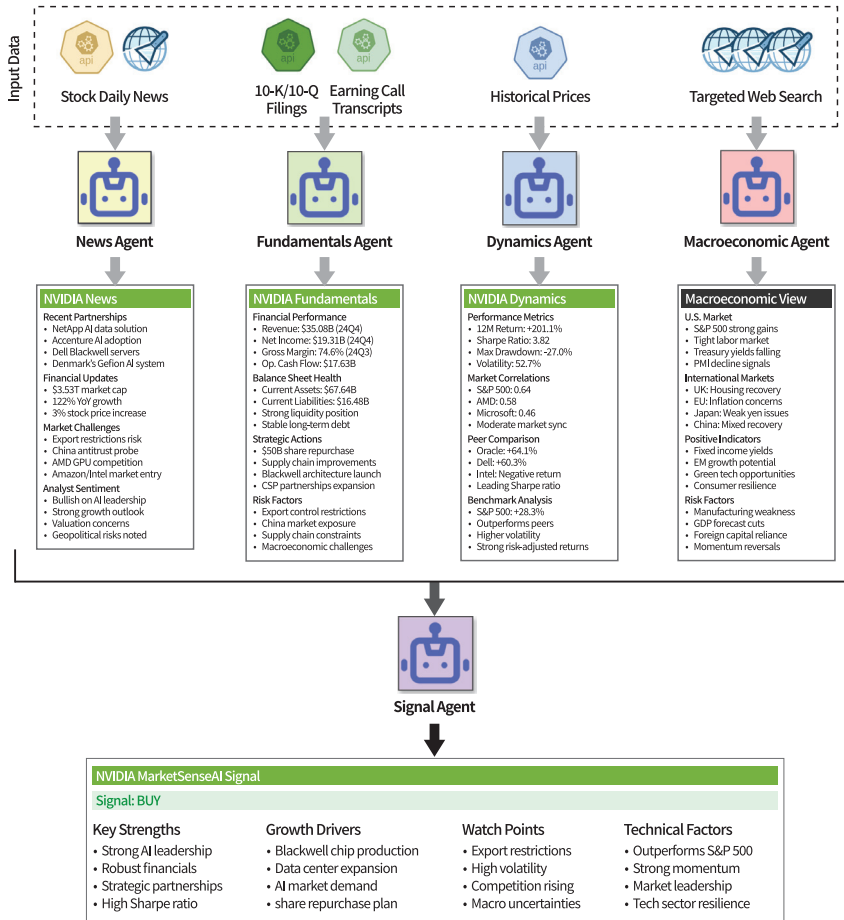
&lt;그림 II-2&gt; PloutosGPT 사례



자료: Tong et al.(2024)

이러한 멀티 에이전트 모델의 핵심은 개별 에이전트의 분석을 기계적으로 종합하는 것을 넘어, 상위 에이전트가 각 의견의 중요도를 판단하고 결합하여 '설명 가능한' 투자 근거를 생성한다는 점이다. PloutosGPT의 경우, 최종 예측치(예: “주가 1.5% 상승 예측”)와 더불어, “긍정적 요인 1: [감성 분석가] A사와의 파트너십 발표”, “긍정적 요인 2: [기술적 분석가] 주가 7일 이동평균선 상회”와 같이 구체적이고 순위가 매겨진 근거를 함께 제시한다(<그림 II-2> 참고). ‘MarketSenseAI’라는 또 다른 모델 역시 유사한 접근법을 취하는데, 이 모델은 뉴스, 펀더멘털, 거시경제 에이전트 등이 수집한 정보를 ‘시그널 에이전트’가 취합하여 최종 신호를 생성한다. <그림 II-3>의 엔비디아 분석 사례에서 보듯이, 단순한 ‘매수’ 신호뿐만 아니라 해당 결정의 배경이 된 핵심 강점(예: AI 리더십), 성장 동력(예: 데이터센터 확장), 주의점(예: 수출 제한) 등을 명확히 정리하여 제공한다.

<그림 II-3> MarketSenseAI 사례



자료: MarketSenseAI 2.0

이처럼 LLM 기반 멀티 에이전트 모델들은 단순한 예측 결과를 넘어 왜 그러한 결론에 도달했는지에 대한 설득력 있는 서사(narrative)를 제공하는 방향으로 발전하고 있다. 이러한 접근법은 자산운용업의 실무 환경에서 특히 중요한 의미를 지닌다. 자산운용사는 투자 결정의 책임을 지고 있으며, 때때로 고객이나 규제 당국에 투자 근거를 설명해야 할 필요가 있다. 또한 포트폴리오 매니저나 투자위원회는 AI가 생성한 투자 판단을 맹목적으로 수용하기보다는, 그 논리적 과정을 비판적으로 검토하고 최종 결정을 내려야 한다. 따라서 ‘블랙박스(black box)’가 아닌 투명한 근거를 제시하는 이러한 모델들은 자산운용업의 업무 특성과 부합하며, AI와 인간 전문가 간의 효과적인 상호작용을 촉진하는 핵심 기제로서 작용할 것으로 판단된다. 이를 통해 자산운용사는

투자 가치 평가의 효율성을 높이면서도, 투자 결정의 신뢰성과 책임성을 유지할 수 있는 가능성을 확보하게 된다.

### 3. 기대효과와 한계

LLM을 활용한 멀티 에이전트 투자 모델은 자산운용업의 핵심 가치 창출 과정인 투자 가치 평가에서 여러 가지 기대효과를 갖는다. 가장 큰 장점은 인간의 물리적, 인지적 한계를 뛰어넘는 정보 처리 능력이다. 인간 애널리스트가 하루에 분석할 수 있는 정보의 양은 제한적이지만, AI 에이전트 시스템은 전 세계에서 발생하는 방대한 양의 정형 및 비정형 데이터를 24시간 실시간으로 분석하고 융합할 수 있다. 특히 연간 보고서, 분기 실적, 일일 뉴스, 실시간 소셜 미디어 데이터와 같이 발표 주기가 상이한 정보들을 유기적으로 결합하여 종합적인 인사이트를 도출하는 능력은 전통적인 계량 분석 모델들과 차별화되는 부분이다. 이러한 정보 처리 능력의 향상은 자산운용사가 더 많은 투자 기회를 발굴하고, 더 신속하게 시장 변화에 대응할 수 있게 함으로써 초과수익 창출 가능성을 높이는 데 기여할 수 있다. 또한, 자연어 기반으로 투자 논리와 근거를 명확하게 제시하는 능력은 모델의 투명성을 높여 투자 결정의 책임성을 강화하고, 규제 당국이나 고객에게 투자 전략을 설명해야 할 때 유용하게 활용될 수 있다. 이는 자산운용업의 신뢰성 확보와 지속 가능한 경쟁력 구축에 중요한 요소이다.

하지만 이러한 장점에도 불구하고, 이 접근법은 자산운용업의 초과수익 창출이라는 핵심 목표 달성에 있어 몇 가지 명확한 내재적 한계를 지닌다. 첫째, 모델의 성능이 근본적으로 인간이 생성한 텍스트 데이터의 질과 범위에 의존한다는 점이다. 이 모델들은 기존의 인간 지식을 학습하고 추론하는 방식으로 작동하기 때문에, 인간이 아직 발견하지 못했거나 언어로 명확하게 정의하지 못한 비직관적인 시장의 이면 관계나 새로운 패턴을 발견하는 데에는 한계가 있다. 즉, 이 모델은 인간 전문가의 능력을 자동화하여 ‘확장’하는 데에는 뛰어나지만, 인간의 지식 체계를 ‘초월’하여 시장에서 새로운 초과수익의 원천을 발견하는 데에는 어려움을 겪는 것으로 평가된다. 이는 자산운용업이 지속적으로 추구해야 하는 경쟁 우위 확보에 제약으로 작용할 수 있다.<sup>02</sup>

02 다만, 현재 LLM의 발전 속도가 매우 빠르므로, 향후 또 다른 창발성이 발현되어 기존 지식의 조합을 통해 새로운 영역으로 확장하거나 초지능이 등장하여 이러한 한계를 극복할 가능성도 배제할 수 없다.

둘째, 이러한 모델들의 장기적인 성과는 아직 충분히 검증되지 않았다는 실증적 한계가 존재한다. 현재까지 제시된 모델들이 특정 기간 동안 우수한 백테스트 성과를 보였음에도, 실제 시장 상황의 변화에도 강건하게(robust) 유지될 수 있는지에 대한 실증적 연구는 여전히 부족한 실정이다(Li et al., 2025).

셋째, LLM 고유의 기술적 문제점이 자산운용업의 위험 관리에 심각한 도전을 제기할 수 있다. 대표적인 것이 ‘환각(hallucination)’ 현상으로, 모델이 사실에 근거하지 않은 내용을 그럴듯하게 생성하는 문제이다. 사실 관계의 정확성이 결정적으로 중요한 투자 분야에서 이러한 오류는 치명적인 손실을 초래할 수 있으며, 이는 자산운용사의 신뢰성과 고객 자산 보호 의무에 직접적인 위협이 된다. 또한, 모델이 학습한 데이터에 내재된 편향이 결과물에 그대로 반영되거나 증폭될 위험도 존재한다. 예를 들어, 특정 시기 시장을 지배했던 과도한 낙관론이나 비관론에 편향되어 학습된 모델은 객관적인 판단을 내리지 못하고 결국 추종 매매의 위험에 크게 노출될 수 있다. 이러한 편향은 자산운용사가 독립적인 투자 판단을 통해 초과수익을 창출하려는 본질적 목표와 상충될 수 있다.

또한, 모델이 제시하는 ‘설명’이 오히려 모델의 실제 한계를 가리는 ‘안전 환상(illusion of safety)’을 초래할 위험을 경계해야 한다. ‘안전 환상’이란, 모델이 그럴듯한 설명을 제공한다는 사실 자체가 사용자로 하여금 해당 모델의 결정 과정을 명확히 이해하고 통제하고 있다고 믿게 만들어, 모델의 실제 오류 가능성이나 내재된 편향을 간과하게 만드는 인지적 편향을 의미한다. 모델이 제시하는 그럴듯한 설명이 실제 성과의 원인인지, 아니면 사후적 합리화에 불과한지에 대한 깊이 있는 검증이 필요하며<sup>03</sup>, 만약 후자에 가까울 경우 자산운용사는 논리적으로는 설득력이 있으나 실제 예측력과 무관한 서사에 의존하는 투자 결정을 내릴 위험에 직면하게 되며, 초과수익 창출이라는 본질적 목표 달성에 실패할 수 있다.

03 예를 들어, Turpin et al.(2023)은 LLM이 제시한 ‘그렇듯해 보이는 설명’과 최종 예측을 이끌어낸 ‘실제 원인’ 간 차이가 있을 수 있다는 사실을 입증하며, 이러한 불성실함(unfaithfulness)은 모델이 정직한 설명을 하도록 훈련되지 않았기 때문이라고 지적하였다.

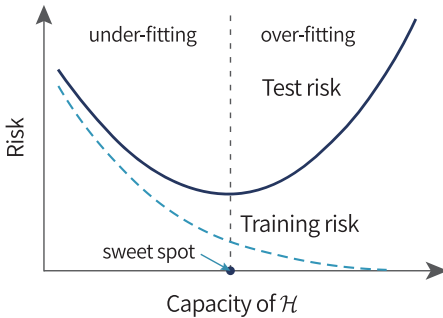
### III. 금융 특화 트랜스포머 모델

앞서 논의한 LLM 에이전트 시스템이 인간의 지식과 추론 과정을 모방하는 데 중점을 두는 반면, 이와는 철학적으로 구분되는 두 번째 AI 접근법이 존재한다. 이는 트랜스포머 아키텍처의 범용성을 활용하여, 이를 금융 시계열이나 기업 특성 데이터와 같은 정형 데이터에 직접 적용하려는 시도이다. 이러한 접근법의 목적은 다음과 같다. 첫째, 자산별 기대수익률의 추정 정확도를 높여 초과수익의 원천을 확장하는 것, 둘째, 시변적(time-varying) 공분산과 교차 자산 상호작용을 더 잘 포착해 포트폴리오 차원의 위험-수익 효율을 개선하는 것이다. 따라서 이 방식은 설명 가능성 확보보다 예측 성능 자체를 극대화하는 데 초점을 두며, 자산 가격 예측에 대한 기존의 통계적 관념에 얽매이지 않고 새로운 방법론을 모색한다는 점에서 전통적인 계량 모델과 차별화된다.

#### 1. 복잡성의 미덕

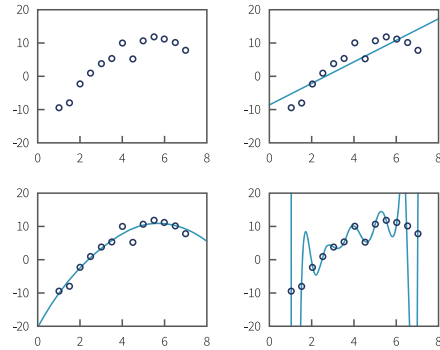
전통적인 계량경제학은 경제 이론에 기반한 소수의 설명 변수를 통해 미래를 예측하려는 시도에 뿌리를 두고 있으며, 모델의 간결성(parsimony)을 중요한 미덕으로 여긴다. 이러한 관점은 통계학의 오랜 원칙인 ‘편향-분산 트레이드오프(bias-variance trade-off)’에 근거한다(Belkin et al., 2019)(<그림 III-1> 참고). 이 원칙에 따르면, 모델이 너무 단순하여(즉, 파라미터가 지나치게 적으면) 데이터에 내재된 근본적인 구조를 포착하지 못해 편향(bias)이 높아지는 ‘과소적합(underfitting)’ 문제가 발생한다(<그림 III-2> 우측 상단 참고). 반대로, 모델이 너무 복잡하여(즉, 파라미터가 지나치게 많으면) 데이터의 본질적 구조가 아닌 노이즈(noise)까지 학습하는 ‘과적합(overfitting)’ 문제가 발생하며, 이로 인해 분산(variance)이 커져 새로운 데이터에 대한 예측력이 저하된다(<그림 III-2> 우측 하단 참고). 따라서 전통적인 계량경제학에서는 이 편향과 분산이 균형을 이루는 ‘최적점(sweet spot)’을 찾는 것(<그림 III-2> 좌측 하단 참고)을 핵심 과제로 인식하였다.

<그림 III-1> 전통적인 편향-분산  
트레이드오프 관계



자료: Belkin et al.(2019)

<그림 III-2> 전통적 관점에서의  
모델 구현 예시

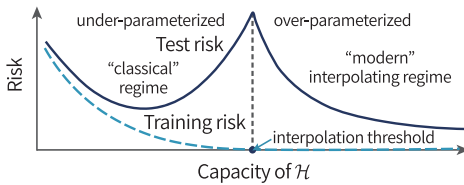


자료: probmods.org

그러나 딥러닝을 위시한 현대 머신러닝의 발전은 이러한 전통적인 통념에 강력한 의문을 제기한다. Belkin et al.(2019)은 모델의 파라미터 수가 데이터의 수를 훨씬 초과하는 특정 임계점을 넘어서면, 오히려 분산이 다시 감소하는 ‘이중 하강(double descent)’ 현상을 발견하였다(<그림 III-3> 참고). 이는 모델의 복잡도와 분산 간의 관계가 전통적인 U자형 곡선이 아닌, W자형 패턴을 보일 수 있음을 의미한다. 즉, 전통적 관점에서 ‘과적합’ 영역으로 간주되던 지점을 지나 파라미터가 극단적으로 많아지는 ‘과잉-파라미터화(over-parameterized)’ 영역에서는, 적절한 ‘최적화 알고리즘’을 활용하여 일반화 성능과 예측력을 오히려 향상시킬 수 있다.<sup>04</sup> 예를 들어, <그림 III-4>의 그래프는 과적합 상태에서 파라미터 수를 더욱 늘렸을 때 예측 모델이 어떻게 변하는지 시각적으로 보여주는 사례이며, 여기에서 파라미터 수가 훨씬 많은 경우(실선으로 표시)가 그렇지 않은 경우(점선으로 표시)보다 분산이 작아 예측에 유리하다는 사실을 확인할 수 있다.

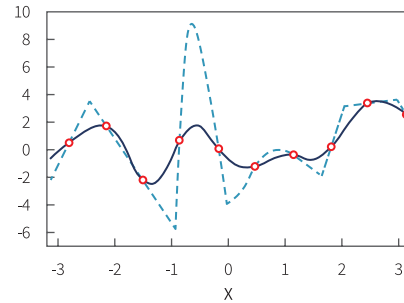
04 파라미터 수가 관측치 수를 초과하는 영역에서는 학습 데이터를 완벽히 설명하는 해(solution)가 무수히 많이 존재할 수 있는데, 최적화 알고리즘은 그중에서도 가장 단순하고 ‘매끄러운(smooth)’ 형태의 해를 선별하는 역할을 한다. 이러한 메커니즘은 모델이 노이즈 대신 일반적인 패턴을 학습하도록 돕는다.

&lt;그림 III-3&gt; 이중 하강 현상



자료: Belkin et al.(2019)

&lt;그림 III-4&gt; 과적합 상태에서의 모델 구현 예시



자료: probmods.org

이러한 패러다임의 전환은 금융 분야의 오랜 논쟁에 중요한 시사점을 제공한다. Goyal & Welch(2008)는 배당수익률, PER, PBR 등 주가 예측에 유용하다고 알려진 전통적인 지표들이 실제로는 시장 지수 예측에 유의미한 기여를 하지 못한다고 결론 내렸다. 이 연구는 금융 분야 최고 학술지인 Review of Financial Studies에 게재되며 오랜 기간 정설로 받아들여졌다. 그러나 Kelly et al.(2024)은 16년 뒤 또 다른 최고 학술지인 Journal of Finance에 게재된 논문에서 정반대의 결론을 제시하였다. 이들은 Goyal & Welch(2008)와 동일한 데이터를 사용하면서도, 의도적으로 파라미터 수를 극단적으로 늘린 복잡한 모델을 적용하여 유의미한 시장 예측력을 확보하는 데 성공하였다.

Kelly et al.(2024)의 발견은, 전통적 관점에서는 ‘과적합’으로 폐기되었을 모델이 오히려 단순한 모델보다 우월한 예측 성과를 낼 수 있음을 증명한 것이다. 이들은 이러한 현상을 ‘복잡성의 미덕(The Virtue of Complexity)’이라는 개념으로 설명하였다. 이 주장의 핵심은 금융 시장이 본질적으로 매우 복잡하고 비선형적인 관계들로 구성되어 있기 때문에, 모델의 복잡성을 의도적으로 높이는 것이 데이터에 내재된 복잡한 패턴을 포착하고 예측력을 높이는 데 필수적이라는 것이다. 이는 모델의 간결성과 해석 가능성을 중시하던 기존 계량경제학의 철학을 근본적으로 뒤집는 접근법이며, 예측 성능을 극대화하기 위해서라면 기꺼이 모델의 복잡성과 비직관성을 감수해야 한다는 새로운 관점을 금융 연구에 제시한 것으로 평가된다.



## 2. 적용 사례

‘복잡성의 미덕’이라는 새로운 철학은 트랜스포머 아키텍처를 금융 정형 데이터 분석에 직접 적용하는 연구들로 구체화되고 있다. 이 접근법은 자연어를 처리할 때 개별 단어를 토큰으로 간주하는 것과 유사하게, 개별 기업의 재무 비율, 주가 변동성, 모멘텀 지표 등 수십, 수백 개의 기업 특성을 토큰으로 취급한다. 이후 트랜스포머의 핵심 메커니즘인 ‘셀프 어텐션’을 이용하여, 이들 특성 간에 존재하는 복잡한 고차원적 상호작용을 학습하는 것을 목표로 한다.

이러한 접근법의 정수는 Kelly et al.(2025)이 제안한 ‘AIPM(Artificial Intelligence Asset Pricing Models)’에서 명확히 드러난다. 이 모델은 A 주식의 미래 수익률을 예측하기 위해 오직 A 주식의 과거 데이터와 특성만을 사용하는 전통적인 ‘자기 자산 예측(own-asset prediction)’의 한계를 근본적으로 해결한다. 이를 위한 AIPM의 핵심 혁신은 트랜스포머 아키텍처의 고유한 장점을 활용한 ‘교차 자산 정보 공유(cross-asset information sharing)’에 있다. 이는 셀프 어텐션 메커니즘이 제공하는 ‘문맥 파악’ 능력을 통해 A 주식 예측 시 B 주식의 모멘텀이나 C 주식의 변동성 같은 교차 자산의 관련성을 동적으로 학습하기에 가능하다. 또한, 이러한 복잡한 상호작용의 연산이 시장 전체의 방대한 자산과 특성에 걸쳐 실제로 가능해진 것은 트랜스포머 아키텍처의 효율적인 ‘병렬 처리’ 능력이 모델의 복잡성을 극단적으로 확장하는 것을 허용했기 때문이다. 결국 AIPM은 이 두 가지 특성(교차 자산 정보 공유 + 병렬 처리)의 결합을 통해 개별 자산을 넘어 시장 전체의 동적인 관계망 속에서 기대수익률을 파악한다. 그 결과, 기존 머신러닝 모델을 압도하는 높은 샤프 비율(Sharperatio)을 달성했으며, 이는 ‘복잡성의 미덕’이 실제 투자 성과 향상으로 이어질 수 있음을 입증한 사례로 평가된다.

Li et al.(2024)의 ‘MASTER(Market-Guided Stock Transformer)’ 모델은 트랜스포머를 활용하여 주가 예측의 또 다른 난제에 접근한다. 주식 시장에서 종목 간 상관관계는 고정되어 있지 않고 시시각각 변하며, 특정 시점에서는 전혀 다른 시간대의 주가 흐름과 연관성을 보이기도 한다. MASTER는 이러한 ‘순간적이고 교차적인(momentary and cross-time)’ 상관관계를 포착하도록 설계되었다. 이 모델은 특정 종목의 현재 주가 흐름을 분석할 때, 다른 종목들의 과거 특정 시점 주가 흐름 중 어떤 부분에 ‘집중’해야 할지를 학습한다. 이를 통해 업종이나 규모와 같은 정적인 관계를 넘어, 시장 상황에 따라 동적으로 형성되는 복잡한 리드-래그(lead-lag) 관계나 군집 행동 패턴을 효과적으로 포착할 수 있다.

이러한 금융 특화 트랜스포머 모델들의 등장은 다른 과학 분야에서도 관찰되는 패러다임 전환과 궤를 같이한다. 예를 들어, 의학 분야에서 Shmatko et al.(2025)은 환자의 건강 기록을 마치 언어처럼 토큰화하여 1,000가지 이상의 질병 발생률을 동시에 예측하는 모델을 개발하였다. 이는 특정 단일 질병(예: 심혈관 질환) 예측에 집중했던 기존 연구들과 달리, 인간이 직접 파악하기 어려웠던 질병 간의 복잡한 상호 의존성을 트랜스포머 아키텍처가 종합적으로 학습했기에 가능한 성과이다. AIPM과 MASTER의 사례는 금융 분야에서도 트랜스포머 모델들이 데이터에 내재된 고차원적이고 동적인 패턴을 직접 학습함으로써, 인간의 사전 가설이나 전통적인 경제 이론의 제약을 뛰어넘는 새로운 예측의 지평을 열고 있음을 보여준다.

### 3. 기대효과와 한계

금융 특화 트랜스포머 모델이 제시하는 가장 큰 잠재력은 예측 패러다임의 근본적인 전환 가능성에 있다. 기존의 쿼트 모델은 인간 전문가가 미래 예측력을 가질 것으로 예상되는 ‘팩터(factor)’(예: 가치, 모멘텀, 퀄리티)를 일일이 찾아내고 그 유효성을 검증하는 ‘가설 검증’ 방식에 가까웠다면, 이 새로운 접근법은 AI가 데이터 속에서 스스로 유의미한 패턴, 즉 새로운 초과수익의 원천을 발견하는 ‘데이터 기반 발견’ 방식에 해당한다. AI가 기업 고유의 효과뿐만 아니라 기업 간의 복잡한 상호작용과 시변적 패턴까지 이해할 수 있게 되면서, 시장 비효율성을 포착하는 능력은 비약적으로 향상될 잠재력을 지닌다. 궁극적으로 이는 금융 시장에서 인간의 직관이나 기존 이론의 범주를 뛰어넘는 완전히 새로운 투자 전략의 탄생 가능성을 시사한다.

그러나 다른 한편으로, 금융 특화 트랜스포머 모델은 심각한 한계를 내포하는데, 가장 큰 문제는 모델 복잡성 증가에 따라 필연적으로 발생하는 ‘설명 불가능성(unexplainability)’이다. 모델 구조가 극도로 복잡해짐에 따라, 모델이 왜 특정 예측이나 투자 신호를 생성했는지 그 이유를 인간이 이해할 수 있는 언어로 설명하기가 거의 불가능해지는 ‘블랙박스’ 문제가 발생한다. 이러한 불투명성은 다음과 같은 현실적인 문제들을 야기한다. 첫째, 투자 운용의 책임과 신뢰 확보가 어려워진다. 예컨대, 고객이나 이사회에 수백만달러 손실의 원인을 단순히 ‘AI 모델의 결정’이라고 설명하기는 어려울 것이다. 둘째, 리스크 관리의 복잡성이 증대된다. 모델이 어떤 가정 하에 작동하는지 명확히 알 수 없으므로, 특정 시장 환경(예: 갑작스러운 유동성 경색, 지정학적 위기)에서 모델이 어떻게 반응할지 예측하고 통제하기가 매우 어렵다. 마지막으로, 규제 준수의 문제가 발생할 수 있다. 때때로 규제 당국은 투자 결정 과정의 투명성과 합리적인 근거 제시를 요구하는

경향이 있는데, 블랙박스 모델은 이러한 요구사항을 충족시키기 어렵다. 따라서 예측 성능의 극대화과 설명 가능성 확보라는 상충되는 목표 사이에서 어떻게 균형을 맞출 것인가는 이 새로운 패러다임이 마주한 가장 중요한 과제로 남아 있다.

## IV. 자산운용업에의 시사점

본고는 AI 기술, 특히 트랜스포머 아키텍처의 발전이 자산운용 산업에 가져올 근본적인 패러다임 전환의 가능성을 다각도로 고찰하였다. 분석 결과, 자산운용 분야의 AI 활용은 크게 두 가지 주요 방향으로 전개되고 있음을 확인하였다. 첫 번째는 LLM 기반 에이전트 시스템을 통해 인간 전문가 팀의 협업적 의사결정 과정을 모방하고 자동화하는 접근법이다. 이는 방대한 비정형 데이터를 종합하고 설명 가능한 투자 논리를 생성하는 데 강점을 지니며, 기존 운용 프로세스를 강화하고 증강하는 역할을 수행한다. 두 번째는 금융 데이터 자체에 내재된 복잡한 패턴을 트랜스포머 모델로 직접 학습하여 예측 성능을 극대화하려는 시도이다. 이는 ‘복잡성의 미덕’이나 ‘이중 하강’ 현상과 같은 새로운 이론적 배경을 바탕으로, 인간의 직관이나 기존 이론의 범주를 뛰어넘는 새로운 초과수익의 원천을 발견할 잠재력을 보여준다. 이 두 접근법은 각각 ‘설명 가능성’과 ‘예측 성능’이라는 상충될 수 있는 가치를 일정 부분 대변하며, 자산운용사에게 중요한 전략적 선택지를 제시한다.

한편으로, 현시점에서 논의되는 AI 모델들이 지닌 명백한 한계점은 이러한 패러다임 전환이 단기적 기술 도입만으로 완성될 수 없음을 시사한다. 현재 제시되는 사례들은 미래 자산운용의 청사진을 제공함에도 불구하고 대부분 초기 실험 단계에 머물러 있으며, 모델의 안정성이나 장기 성과에 대한 실증적 검증은 여전히 부족하다. 더욱이 LLM을 위시한 AI 기술의 발전 속도는 기존의 혁신 주기와 비교할 수 없을 정도로 빨라, 현재의 우위 모델이 수개월 내에 무의미해질 수 있는 높은 기술적 변동성에 노출되어 있다. 따라서 AI 시대 자산운용사의 장기적인 경쟁력은 특정 알고리즘의 우수성이나 단기적인 기술 우위가 아닌, 패러다임 전환의 본질을 꿰뚫는 전사적 차원의 체질 개선 능력에서 판가를 날 것이다. 이는 AI를 중심으로 조직 전체의 워크플로우와 문화를 재설계하려는 장기적 비전과 과감한 실행력으로 발현될 수 있을 것이다.

이러한 요건을 충족하는 기반 위에서, 구체적인 세 가지 핵심 전략 과제를 꼽으면 다음과 같다. 첫째는 미래 가치 창출의 원천이 될 고품질 데이터의 체계적인 확보 및 관리 역량이다. AI 모델의 성능은 학습 데이터의 질과 양에 절대적으로 의존하므로, 양질의 데이터를 적시에 확보하고 이를 AI가 학습 가능한 형태로 정제·활용하여 발견 가능한 패턴의 총량을 늘리는 것이 무엇보다 중요하다. 둘째는 방대한 고객 데이터와 AI 기술을 결합한 초개인화(hyper-personalization)된 맞춤형 투자 솔루션의 구축이다. 이는 단순한 자산 배분 효율성 개선을 넘어, AI 기술로 고객의 복잡다단한 재무 목표와 성향을 실시간으로 반영하는 동적인 솔루션을 제공함으로써 자산운용업의 본질 자체를 재정의할 잠재력을 내포하고 있다. 셋째는 AI 도입에 필연적으로 수반되는 새로운 위험을 식별하고 관리하는 역량 확보다. 알고리즘의 예측 실패 가능성, 모델의 불투명성, 동일한 인기 모델로의 행태 수렴 등 금융 산업의 특수성과 결합된 AI 고유의 위험 요인을 선제적으로 관리하는 능력은 지속 가능한 성장을 위한 필수 조건이다. 이처럼 AI 기술 변화에 대한 깊이 있는 이해를 바탕으로 데이터, 개인화, 그리고 위험 관리라는 핵심 부문에 전략적으로 대응하는 것이야말로 향후 자산운용사의 생존과 성장을 좌우할 결정적 요인이 될 것이다.

## 1. 데이터 확보 및 관리

AI 중심의 인프라 구축에 있어 가장 핵심적인 부분은 양질의 ‘데이터’를 확보하고 활용하는 역량이다. AI 시대의 자산운용사에게 데이터는 더 이상 운용의 보조 수단이 아니라, 그 자체로 가장 중요한 핵심 자산이며 경쟁 우위의 원천이 된다. AI 모델의 성능은 근본적으로 학습 데이터의 질과 양, 그리고 다양성에 의해 결정되기 때문이다. 따라서 미래 경쟁력 확보를 위한 최우선 과제는 리서치, 포트폴리오 구성, 트레이딩, 리밸런싱, 위험 관리 등 각 워크플로우 단계별로 필요한 데이터를 정의하고 이를 효과적으로 통합, 활용하기 위한 체계적인 데이터 전략을 수립하는 것이다. 독점적이거나 고품질의 데이터 자산을 얼마나 확보하고 효과적으로 활용하느냐가 미래 자산운용사의 초과수익 창출 능력을 결정하는 핵심적인 차별화 요인이 될 것이다.

데이터 전략의 첫걸음은 데이터 확보의 범위를 과감하게 확장하는 것이다. 주가와 재무제표 등 전통적인 금융 데이터 수준을 넘어, 잠재적 가치를 지닌 모든 종류의 데이터를 적극적으로 수집하고 축적해야 한다. 여기에는 기업 관련 뉴스 및 공시 자료, 경영진 컨퍼런스 콜 녹취록, 애널리스트 및 외부 연구기관의 분석 보고서, 거래소나 금융정보 제공업체 등 주요 인프라 기관이 제공하는 데이터 등이 포함된다. 또한, 자체적인 리서치나 외부 파트너십을 통해 직접 발굴하는

대안 데이터(alternative data) 역시 중요하다. 뿐만 아니라, 고객의 거래 패턴, 웹사이트 방문 기록, 모바일 앱 사용 데이터와 같이 향후 고객 맞춤형 투자 솔루션 제공의 기반이 될 수 있는 자산운용사 고유의 내부 데이터(proprietary data) 확보 및 활용 전략도 필수적이다. AI 모델은 인간이 보기에는 직접적인 관련성이 없어 보이는 이종 데이터 간의 복잡한 상관관계를 발견할 수 있으므로, 확보된 데이터의 다양성은 곧 발견 가능한 잠재적 패턴의 기하급수적인 확대로 이어질 수 있다.

확보된 데이터를 효과적으로 활용하기 위해서는 데이터 품질 관리 및 거버넌스 체계를 정립하고, 이를 지원하는 기술적 인프라를 구축해야 한다. 데이터에 내재된 오류, 누락, 편향은 AI 모델의 성능을 심각하게 저하시키는 주요 원인이므로, 데이터를 정제하고 결측치를 처리하며 데이터의 출처와 신뢰도를 체계적으로 관리하는 프로세스를 표준화하고 자동화하는 것이 중요하다. 동시에, 방대한 양의 정형 및 비정형 데이터를 효율적으로 저장하고 신속하게 처리할 수 있는 기술 기반이 필수적이다. 이를 위해 원시 데이터를 형태에 구애받지 않고 저장하는 데이터 레이크(data lake)와 분석 목적에 맞게 정제·구조화된 데이터를 저장하는 데이터 웨어하우스(data warehouse) 등을 구축하고, 해당 권한을 가진 직원이 쉽게 데이터에 접근하여 모델을 개발·실험하며 그 결과를 실제 운용 환경에 배포할 수 있는 통합 분석 플랫폼 환경을 제공해야 한다.

## 2. 초개인화된 맞춤형 투자 솔루션 구축

모든 투자자는 동질적이지 않으므로, 앞서 제시된 자산 가격 예측 모델의 효용을 극대화하기 위해서는 자산운용사가 결국 투자자 각각의 고유한 상황에 맞는 투자 솔루션을 개발하고 제공해야 한다. 특히 AI 시대에는 방대한 투자자 데이터를 분석하고 활용하여 초개인화된 맞춤형 투자 솔루션을 낮은 비용으로 제공하는 것이 가능해진다. 이는 단순히 정적인 자산 배분 계획을 제시하는 것을 넘어, 고객의 복잡한 재무 목표, 위험 성향, 생애 주기 등을 실시간으로 반영하여 동적으로 운용 전략을 조정하는 포괄적인 서비스를 의미한다. 이러한 초개인화된 투자 솔루션은 자산운용사의 기존 역할을 넘어서는 궁극적인 경쟁 우위의 원천, 즉 게임 체인저(game changer)가 될 잠재력을 지닌다.

‘투자 솔루션’이라는 용어는 엄밀한 학문적 정의보다는 업계의 실무적 필요에 의해 정착된 측면이 강하다. 전통적인 자산운용 업무가 주로 포트폴리오를 구축하고 운용하는 실행(do) 단계에

집중했다면, 점차 투자 목표 설정 및 계획(plan) 단계와 성과 평가 및 피드백(see) 단계를 포괄하는 전체 자산관리 프로세스로 업무 범위가 확장되면서 ‘투자 솔루션’이라는 용어가 통용되기 시작했다. 본고의 맥락을 고려하여 ‘투자 솔루션’을 정의하자면, ‘고객의 재무 목표 달성을 위해, 자산 배분 · 상품 선택 · 위험 관리 · 성과 평가 등을 통합적으로 설계하여 제공하는 운용 및 자문 서비스 패키지’로 볼 수 있다.

AI 기반 투자 솔루션은 이러한 자산관리의 전 과정, 즉 계획(plan)-실행(do)-평가(see)의 순환 구조에서 인간의 판단을 보조하거나 AI 기술로 대체하려는 시도이다. 과거 전통적인 방식으로는 소수의 고액 자산가에게만 제공 가능했던 맞춤형 서비스를, AI 기술을 통해 다수의 고객에게 낮은 비용으로 초개인화하여 제공할 수 있다. 예를 들어, 마이데이터 등을 통해 확보된 고객의 금융 정보, 비금융 정보, 행동 데이터 등을 AI가 분석하여 개인별 최적 투자 계획을 수립하고(plan), 앞서 논의된 AI 기반 예측 및 최적화 모델을 통해 포트폴리오를 구성하고 운용하며(do), 운용 성과를 AI 기술로 정교하게 분석하고, 그 결과를 다시 계획 단계에 피드백하여 포트폴리오를 동적으로 조정하는(see) 선순환 구조를 구축할 수 있다.

이러한 AI 기반 투자 솔루션의 대표적인 구현 사례는 로보어드바이저(Robo-Advisor: RA)라 할 수 있다. 초기의 로보어드바이저는 알고리즘 개발 지원 등 주로 IT 업무의 생산성 향상에 AI 기술을 제한적으로 활용하였으나, 최근에는 알고리즘을 통한 자산운용 프로세스 전반으로 그 역할이 확대되고 있다. 규칙 기반(rule-based) 알고리즘 운용에 필수적인 정형 · 비정형 데이터 분석을 AI 기술로 처리할 뿐만 아니라, 개별 고객의 목표 수익률과 위험 성향에 부합하는 최적의 운용 알고리즘을 선별하고 추천하는 기능까지 수행한다. 이러한 투자자문 서비스는 자연스럽게 투자 일임 운용으로 확장되어, 고객 계좌 관리 및 최적의 주문 집행까지 포괄하는 맞춤형 투자 솔루션으로 발전하고 있다.

특히 연금 자산운용에서 로보어드바이저가 갖는 잠재력은 OECD(2021)가 제시한 ‘확장가능한 솔루션(scalable solution)’ 개념으로 설명될 수 있다. 이는 디지털 플랫폼 기반의 금융 서비스가 고객 수 확대에 따른 추가 비용(한계비용) 증가폭이 매우 낮아, 광범위한 고객에게 서비스를 효율적으로 제공할 수 있는 확장가능성을 지닌다는 의미이다. OECD(2021)는 “디지털 플랫폼과 자동화된 투자 도구가 낮은 한계비용으로 맞춤형 자문과 자산배분을 제공함으로써, 소매 투자자의 금융 접근성을 개선하고 시장 효율성을 제고할 수 있다”고 평가하였다. 이러한 특성은 소규모 자금을 장기간 안정적으로 운용해야 하는 개인 투자자들에게 합리적인 비용으로 맞춤형 자산관리 서비스를 제공해야 하는 확정기여형 퇴직연금이나 개인형 퇴직연금(Individual Retirement

Pension: IRP) 등 연금 시장의 요구와 정확히 부합한다. 실제로 미국에서 Betterment와 같은 주요 로보어드바이저 전문 기업들은 ‘Betterment at Work’ 등 별도의 기업 연금 플랜 플랫폼을 운영하고 있으며, 피델리티(Fidelity)와 같은 전통적인 대형 연금 사업자 역시 일임 계좌나 하이브리드 로보어드바이저(hybrid RA) 형태로 401k 가입 근로자를 위한 로보어드바이저 상품을 적극적으로 제공하고 있다. 나아가 피델리티, 골드만삭스(Goldman Sachs), 블랙록(BlackRock) 등 주요 금융 그룹들은 IT 기반의 로보어드바이저 전문 기업을 인수하거나 전략적 제휴를 맺는 방식으로 연금 시장에서의 맞춤형 투자 솔루션 제공 역량을 강화하고 있다.

### <참고> 피델리티의 로보어드바이저 서비스 사례

피델리티의 로보어드바이저 운용 사례는 초개인화된 맞춤형 투자 솔루션이 어떻게 구현될 수 있는지를 보여준다. 피델리티는 고용주 제공 퇴직연금 플랜(employer-sponsored plan) 내에서 ‘Fidelity Personalized Planning & Advice at Work’와 같은 일임 계좌 서비스를 제공하는 것으로 알려져 있다. 이 서비스는 401k 플랜에서 허용되는 적격투자상품(Qualified Default Investment Alternative: QDIA) 등을 활용하여, 단순히 투자자의 연령이나 목표 은퇴 시점뿐만 아니라 현재 자산 규모, 연금 외 다른 저축 현황 등 개인의 상세한 재무 정보를 종합적으로 고려하여 초개인화된 포트폴리오를 구축한다. 이를 위해 계좌 가입 초기 단계에서 투자자는 온라인 플랫폼을 통해 자신의 재무 정보와 투자 선호도, 위험 성향 등을 입력하게 된다. 초기 정보 입력이 완료되면, 이후 자산관리의 전 과정은 투자일임 방식으로 자동화되어 진행된다. 자산부채종합관리(Asset Liability Management: ALM) 원칙에 기반한 장기적인 자산 배분 전략 수립, 시장 상황 변화에 따른 주기적인 포트폴리오 리밸런싱 등 핵심적인 의사결정이 알고리즘을 통해 이루어진다.

피델리티가 제공하는 투자 솔루션은 근로 기간 동안의 자산 축적 단계에 국한되지 않고 은퇴 이후 연금 수급(인출) 단계까지 포괄하는 생애주기적 접근법을 취한다. 이는 투자 포트폴리오가 은퇴 시점을 전후하여 단절되지 않고 연속성을 유지하는 것이 중요하며, 특히 세금 관련 자문 및 절세 전략이 중요한 미국의 연금 자산운용 환경에서 인출 단계까지 고려한 맞춤형 설계가 강조되기 때문이다. 이를 위해 피델리티는 자체적인 연구에 기반한 인출 설계 모델인 ‘Managed Cash Flow’ 전략 등을 개발하여, 플랜 참가자가 은퇴 시 안정적인 정기 인출 옵션으로 선택할 수 있도록 제공하고 있다.

### 3. AI 관련 위험 통제

AI 기술의 도입은 자산운용 분야에 새로운 기회를 제공하는 동시에, 기존에는 없었던 새로운 형태의 위험을 야기한다. 따라서 AI 중심의 워크플로우를 성공적으로 구축하기 위해서는 이러한 잠재적 위험을 식별하고 관리하기 위한 체계적인 관리 체계 구축이 필수적으로 수반되어야 한다. AI 도입에 따라 발생할 수 있는 주요 위험으로는 알고리즘 자체의 내재적 한계 및 예측 실패 가능성, 운용 성과 부진 시 책임 소재의 불명확성, 고도화된 사이버 보안 위협, 학습 데이터의 편향성에 따른 차별적 결과 초래 가능성, 그리고 AI 모델 활용 과정에서의 잠재적 이해상충 및 불안 전판매 우려 등을 들 수 있다.

상기한 위험 요인 중 특히 AI 기반 의사결정의 불투명성, 즉 설명가능성 부족 문제는 가장 중요한 관리 과제로 강조된다(OECD, 2021). 본고에서 논의된 트랜스포머 기반 모델들의 복잡성과 창발성이 실제 운용에 적극적으로 구현될수록, 이러한 설명가능성 부족 문제는 더욱 심화될 수 있다. 이 문제는 크게 두 가지 상이한 시나리오로 나누어 살펴볼 수 있다.

첫 번째 시나리오는 금융 특화 트랜스포머 모델과 같이 설명가능성이 극히 제한적인 ‘블랙박스’ 모델의 경우이다. 이러한 모델들은 복잡한 비선형 관계를 포착하여 우수한 예측 성능을 보일 잠재력을 갖추고 있으나, 투자의 성과가 부진할 때 자산운용사가 고객에게 합리적인 설명을 제공할 수 없다는 치명적인 약점을 갖는다. 특히 AI 기반 초개인화 맞춤형 투자 솔루션은 그 효용이 장기적으로 발현되는 특성상 고객이 단기간에 성과를 체감하기 어려울 수 있는데, 손실 발생 시 납득할 만한 설명조차 제공받지 못하는 부정적 경험이 누적된다면, 이는 자산운용사와 고객 간의 신뢰 형성에 근본적인 걸림돌로 작용하여 AI 기반 서비스의 장기적인 이용률을 저해하는 요인이 될 수 있다.

두 번째 시나리오는 LLM 기반 에이전트 시스템과 같이 일정 수준의 설명을 제공함에도 불구하고 발생하는 ‘불완전한 설명가능성’ 문제다. 앞서 ‘안전 환상’ 문제로 지적했듯이, LLM이 결과를 도출하는 과정에서 실제 활용한 논리(혹은 통계적 연관성)와, 결과 도출 후 사후적으로 생성하여 제시하는 설명이 일치하지 않는 ‘불성실한 설명’ 문제가 발생할 가능성이 존재한다. 이는 두 가지 심각한 문제를 야기하는데, 첫째로 투자자가 LLM이 제공하는 그럴듯한 논리를 맹신하게 만들어, 실제로는 검증되지 않았거나 편향에 기반한 모델의 결정을 비판 없이 수용하게 할 위험, 즉 투자자 보호의 실패를 야기할 수 있다. 둘째, 성과가 좋을 때와 나쁠 때의 설명 논리가 일관성을 잃고 사후 합리화에 그친다면, 이는 장기적으로 고객의 신뢰를 무너뜨리고, 더 나아가 금융회사



내부의 위험관리 및 감사 기능의 효과적인 작동마저 저해할 것이다. 모델의 ‘실제 논리’와 ‘제시된 설명’ 간의 불일치는 잠재된 데이터 편향이나 부적절한 모델링 가정을 사전에 식별하고 수정하는 것을 사실상 불가능하게 만들기 때문이다.

설명가능성 문제 외에도, AI 도입이 야기할 수 있는 또 다른 중대한 위험은 시스템 위험의 증폭 가능성이다. 만약 다수의 AI 시스템이 유사한 데이터와 알고리즘(예: 동일한 인기 모델)을 기반으로 시장 상황에 대해 유사한 판단을 내릴 경우, 시장 참여자들이 특정 방향으로 쏠리는 행태 수렴(herding) 현상이 발생할 수 있다. 이러한 동조화는 정상적인 시장 상황에서는 효율성을 높이는 것처럼 보일 수 있으나, 특정 충격이 발생했을 때 매도 혹은 매수 압력을 한 방향으로 집중시켜 시장 변동성을 예기치 않게 증폭시키는 요인으로 작용할 수 있다.

따라서 개별 자산운용사는 AI 모델의 개발, 검증, 배포, 모니터링 및 업데이트에 이르는 전 과정에 걸쳐 발생 가능한 다양한 위험 요인을 식별하고 통제하기 위한 명확한 거버넌스 체계와 내부 통제 프로세스를 수립해야 한다. 또한, 이 모든 과정에서 개인정보 보호 규정 및 데이터 보안 관련 법규를 철저히 준수하는 것은 법적 리스크를 관리하고 고객 및 규제 당국의 신뢰를 유지하는데 필수적인 기반이 될 것이다.

## 참고문헌

- Belkin, M., Hsu, D., Ma, S., Mandal, S., 2019, Reconciling modern machine-learning practice and the classical bias–variance trade-off, *Proceedings of the National Academy of Sciences* 116(32), 15849-15854.
- D’Acunto, F., Prabhala, N., Rossi, A.G., 2019, The promises and pitfalls of robo-advising, *The Review of Financial Studies* 32(5), 1983-2020.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2021, An image is worth 16x16 words: Transformers for image recognition at scale, arXiv preprint.
- Fudan University, Shanghai Academy of AI for Science, Nature Research Intelligence, 2025, *AI for science 2025*.
- Goyal, A., Welch, I., 2008, A comprehensive look at the empirical performance of equity premium prediction, *The Review of Financial Studies* 21(4), 1455-1508.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021, Highly accurate protein structure prediction with AlphaFold, *Nature* 596(7873), 583-589.
- Kelly, B.T., Malamud, S., Zhou, K., 2024, The virtue of complexity in return prediction, *The Journal of Finance* 79(1), 459-503.
- Kelly, B.T., Kuznetsov, B., Malamud, S., Xu, T.A., 2025, Artificial intelligence asset pricing models, NBER Working Paper No. 33351.

- Li, T., Liu, Z., Huang, S., Wang, Z., Chen, E., 2024, MASTER: Market-guided stock transformer for stock price forecasting, *Proceedings of the AAAI Conference on Artificial Intelligence* 38(1), 162-170.
- Li, W., Kim, H., Cucuringu, M., Ma, T., 2025, Can LLM-based financial investing strategies outperform the market in long run?, *arXiv preprint*.
- OECD, 2021, *OECD Business and Finance Outlook 2021: AI in Business and Finance*.
- Park, J.S., O'Brien, J., Cai, C.J., Morris, M.R., Liang, P., Bernstein, M.S., 2023, Generative agents: Interactive simulacra of human behavior, In *Proceedings of the 36th annual acm symposium on user interface software and technology*, 1-22.
- Shmatko, A., Jung, A.W., Gaurav, K., Brunak, S., Mortensen, L.H., Birney, E., Fitzgerald, T., Gerstung, M., 2025, Learning the natural history of human disease with generative transformers, *Nature* 647(8088), 248-256.
- Tong, H., Li, J., Wu, N., Gong, M., Zhang, D., Zhang, Q., 2024, Plutos: Towards interpretable stock movement prediction with financial large language model, *arXiv preprint*.
- Turpin, M., Michael, J., Perez, E., Bowman, S., 2023, Language models don't always say what they think: Unfaithful explanations in chain-of-thought prompting, *Advances in Neural Information Processing Systems* 36, 74952-74965.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017, Attention is all you need, *Advances in neural information processing systems* 30, 5998-6008.
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E.H., 2022, Emergent abilities of large language models, *arXiv preprint*.