

[Pixie: AI-Agent Personalized Investment Service]

Seonkwon Kim, and Byeongwoong Hwang, Hyunjoon Kim
School of Data Science, Hanyang University

Research Motivations

In recent years, major banks and securities firms have mainly used AI for task.

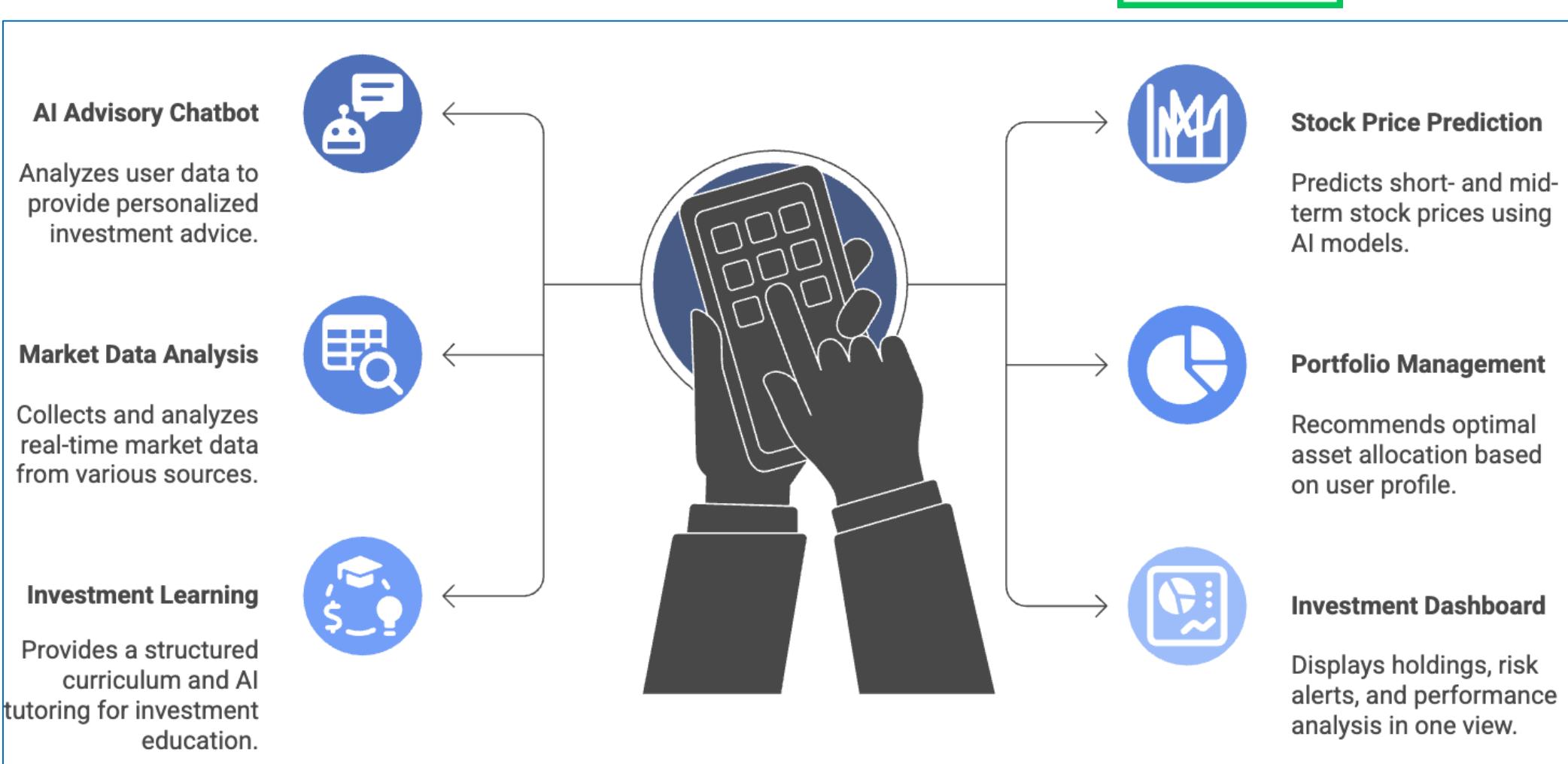
However, they still lack true personalization and remain unreliable due to limited quantitative reasoning and LLM hallucinations.

We propose a personalized investment service powered by AI agents to provide more adaptive and reliable guidance.

User Study

Our Services

(Link) 



AI-Agent Module

Personalized AI



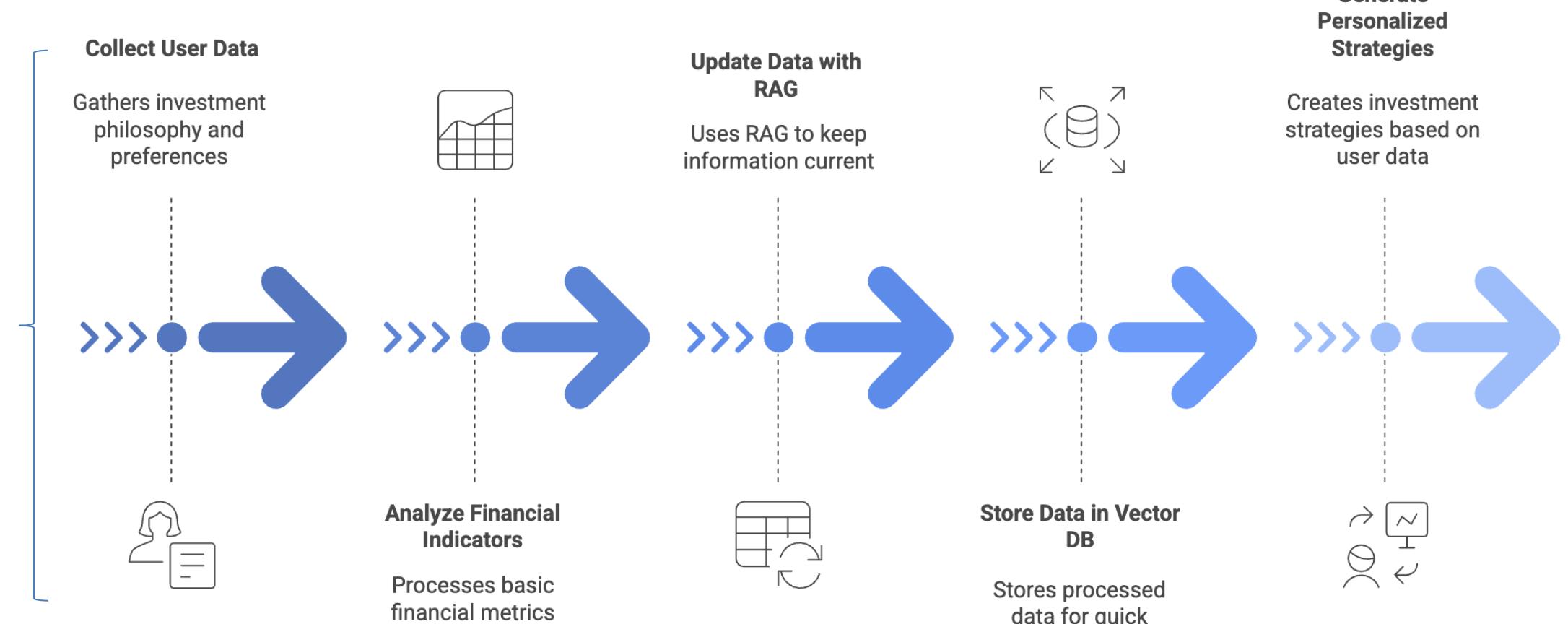
Investment Profile Survey + Chat History

Automated Scoring Using LLM

Final Personality Embedding Output & 6 Investment Personality Indicators

The embedding is directly inserted into the LLM prompt, enabling highly personalized responses and investment insights.

Financial AI



Response Integration

AI-A (Initial Analysis)

AI-A interprets the user's intent and establishes the initial investment strategy.



Final Integration Agent

The final integration agent combines all analysis results to generate the final investment recommendation.

AI-B (Data Analysis)

AI-B analyzes real-time financial and market data.

Models and Technology Stack

LLM & Embedding Model

Model	OpenAI GPT-4 / GPT-4o-mini	Naver CLOVA Studio (HCX-DASH-001)	Ko-SBERT	FAISS	ARIMA	LSTM
Model Description	Transformer-based, 1.75T+ parameters	Korean language-specialized LLM	Korean version of Sentence-BERT	High-performance vector search library	Combines AR, I, and MA components	RNN designed to capture long-term dependencies
Use Case	Personalized investment suggestions, query interpretation, real-time financial data analysis	Backup LLM, Korean market sentiment analysis	User profile embedding, news embedding, financial document embedding	News vector search, user profile matching, stock similarity search	Short-term stock price forecasting	Long-term stock trend forecasting
Key Parameters	temperature, max_tokens, top_p, frequency_penalty, presence_penalty	temperature, maxTokens, topP, repeatPenalty, includeAiFilters	Optimized for Korean language, multitask learning	Supports indexing millions of vectors, GPU acceleration	Combines AR, I, and MA components	Uses memory cell structure to retain important information

Tech Stack

