International Journal of Advanced Multidisciplinary Research and Educational Development
Volume 2, Issue 2 | March – April 2026 | www.ijamred.com
ISSN: 3107-6513

AI Data Analysis Agent Using Large Language Models

Krithika G, V. Yogashri
Department of Computer Science, Rathinam College of Arts and Science (Autonomous), Coimbatore, Tamil Nadu, India
kalaikrithika14@gmail.com , yogashri.cs@rathinam.in

Abstract— *This paper proposes a novel AI-powered Data Analysis Agent built upon Large Language Models (LLMs) to transform how users interact with complex datasets. The agent integrates natural language understanding with automated data processing pipelines, enabling non-technical users to query, explore, and derive insights from structured and unstructured data through conversational interfaces. By leveraging LLMs such as GPT-4 and Claude, the system interprets user intent, formulates analytical plans, and executes data operations including aggregation, visualization, anomaly detection, and trend forecasting. The architecture employs a multi-tool orchestration framework where the LLM acts as a central reasoning engine, delegating tasks to specialized analysis modules. Evaluations conducted on real-world datasets demonstrate that the agent achieves high accuracy in responding to analytical queries while significantly reducing the technical barrier for data exploration. This research contributes a scalable, extensible framework for building LLM-driven analytical assistants applicable across domains such as healthcare, finance, agriculture, and education.*

Keywords— *Large Language Models, Data Analysis Agent, Natural Language Interface, Tool Augmentation, LLM Orchestration, Conversational AI, Automated Data Processing, GPT-4, Claude.*

I. INTRODUCTION

The exponential growth of digital data across industries has created both an opportunity and a challenge. Organizations generate vast volumes of structured data in databases and spreadsheets as well as unstructured data in reports, logs, and communications. Extracting actionable insights from this data traditionally requires specialized expertise in data science, SQL, or statistical programming, creating a significant accessibility gap for domain experts who lack technical proficiency.

Recent advances in Large Language Models (LLMs) have demonstrated remarkable capabilities in natural language understanding, reasoning, and code generation. These capabilities position LLMs as ideal candidates for building intelligent analytical agents that bridge the gap between domain knowledge and data science expertise. An LLM-based data analysis agent can interpret ambiguous queries expressed in plain language, devise analytical strategies, execute computational steps, and communicate results in a narrative, human-readable format.

This paper presents a comprehensive design and implementation of an AI Data Analysis Agent powered by LLMs. The proposed system enables users to describe their analytical goals in natural language and receive accurate, well-explained results without requiring programming skills. The agent is designed to handle multi-step analytical workflows, maintain conversational context across interactions, and produce reproducible analyses through structured tool use.

II. EXISTING SYSTEM VS. PROPOSED SYSTEM

A. Existing System Analysis

Conventional data analysis workflows rely on a combination of manual scripting, dashboard tools, and fixed query interfaces. Business intelligence platforms such as Tableau and Power BI enable visualization but require users to understand the tool's query logic. SQL-based systems demand structured queries that precisely mirror the underlying schema. These approaches present several limitations:

- **Inflexibility:** Fixed templates cannot adapt to novel or composite analytical questions.
- **High Technical Barrier:** Users must learn domain-specific query languages or programming environments.
- **Limited Context:** Individual queries are stateless, preventing progressive analytical conversations.
- **Unstructured Data Exclusion:** Traditional tools predominantly handle structured tabular data, ignoring textual or document-based information.

B. Proposed System Framework

The proposed AI Data Analysis Agent addresses these limitations through a unified conversational interface powered by LLMs. The system accepts free-form natural language instructions and autonomously determines the appropriate sequence of data operations. Key advantages include:

- **Natural Language Interface:** Users express queries in plain English without knowledge of SQL or programming syntax.
- **Multi-Step Reasoning:** The agent decomposes complex requests into manageable sub-tasks and executes them sequentially.

ISSN: 3107-6513

- **Context Retention:** Conversational memory allows iterative refinement of analyses across multiple dialogue turns.
- **Multimodal Data Support:** The agent handles both structured (CSV, databases) and unstructured (PDFs, reports) data sources.

Table I. Comparison of Existing vs. Proposed System

Feature	Traditional	LLM Agent
Adaptability	Static rules	Dynamic & context-aware
Data Handling	Structured only	Structured & unstructured
User Interface	Fixed query	Natural language
Explainability	Limited	High (narrative)
Scalability	Requires re-engineering	Plug-in extension

III. METHODOLOGY

A. System Architecture

The AI Data Analysis Agent is structured around a layered architecture comprising four primary modules: the Natural Language Processing (NLP) Interface, the Orchestration Engine, the Tool Execution Layer, and the Response Generation Module. The NLP Interface receives user input and passes it to the LLM, which identifies the analytical intent and generates a tool-use plan. The Orchestration Engine coordinates the sequencing and execution of this plan.

B. Data Preprocessing and Feature Engineering

Raw data ingested by the agent undergoes an automated preprocessing pipeline. This pipeline performs missing value imputation using statistical strategies (mean, median, or mode-based), detects and flags outliers using the Interquartile Range (IQR) method, and applies normalization to numeric features. Categorical variables are encoded using label encoding or one-hot encoding depending on cardinality. The preprocessing decisions are made dynamically by the LLM based on the data characteristics it observes.

C. LLM Orchestration and Tool Integration

The core of the system is an LLM configured with a set of tool definitions representing available analytical capabilities. When a user submits a query, the LLM generates a structured tool-call plan in JSON format specifying which tools to invoke, with what parameters, and in what sequence. Supported tools include:

- **DataLoaderTool:** Reads CSV, Excel, JSON, and relational database sources.
- **StatAnalysisTool:** Computes descriptive statistics, correlations, and distributions.
- **TrendForecastTool:** Applies time-series models including ARIMA and Prophet.
- **AnomalyDetectorTool:** Identifies outliers using Isolation Forest and z-score methods.
- **VisualizationTool:** Generates charts and graphs using Matplotlib and Plotly.
- **ReportWriterTool:** Compiles insights into formatted textual or PDF reports.

IV. SYSTEM IMPLEMENTATION & DESIGN

A. Frontend and Backend Development

The frontend provides a conversational web interface developed using React.js with a chat-style layout. Users can upload datasets, type queries, and receive responses that may include textual explanations, rendered charts, and downloadable reports. The backend is implemented in Python using the FastAPI framework. It manages user sessions, routes messages to the LLM via the Anthropic or OpenAI API, handles tool execution in isolated sandboxes, and maintains a conversation history buffer to support multi-turn interactions.

B. System Specification

- **Hardware:** Intel Core i7 or equivalent processor, 16 GB RAM minimum, 512 GB SSD for local dataset storage and model caching.
- **Software:** Python 3.11+, Node.js 20+, FastAPI, React.js, LangChain, OpenAI/Anthropic SDK, Pandas, NumPy, Matplotlib, Plotly, scikit-learn, PostgreSQL.
- **LLM Backend:** Anthropic Claude or OpenAI GPT-4 via API, with function-calling/tool-use capabilities enabled.
- **Deployment:** Containerized using Docker, orchestrated with Kubernetes for scalable multi-user deployments.

V. RESULT AND DISCUSSION

The proposed AI Data Analysis Agent was evaluated on a diverse set of real-world datasets including financial transaction records, healthcare patient logs, and agricultural yield databases. Testing covered three dimensions: analytical accuracy, response naturalness, and task completion rate. The agent was compared against a baseline system using conventional SQL queries and Python scripts written by analysts.

Across all evaluation categories, the LLM-based agent outperformed the baseline, particularly on complex multi-step queries and tasks requiring narrative explanations. The agent successfully handled ambiguous queries by requesting clarifications when needed and recovering gracefully from malformed inputs. Response generation latency averaged 3.2 seconds for standard queries and 8.7 seconds for complex multi-tool workflows.

ISSN: 3107-6513

Table II. Performance Results of AI Data Analysis Agent

Task	LLM Agent	Baseline
Data Summarization	94.2%	78.5%
Trend Analysis	91.7%	74.3%
Anomaly Detection	88.9%	70.1%
NL Query Response	96.1%	N/A
Report Generation	92.5%	65.0%

These results confirm the superiority of LLM-based orchestration for analytical tasks, particularly in scenarios involving ambiguous user intent and multi-source data integration. The natural language reporting feature was rated highly by domain expert evaluators who emphasized the accessibility and clarity of the generated insights.

VI. CONCLUSION AND FUTURE ENHANCEMENT

This paper presents an AI Data Analysis Agent powered by Large Language Models that enables natural language interaction with complex datasets. The proposed system successfully bridges the gap between domain expertise and data science proficiency by automating analytical workflows through LLM-driven orchestration of specialized tools. Experimental results demonstrate high accuracy, interpretability, and user satisfaction across diverse analytical tasks.

The proposed architecture is domain-agnostic and extensible, making it applicable to healthcare analytics, financial intelligence, educational research, and agricultural planning. The conversational memory mechanism enables progressive analysis sessions that mimic the working style of professional data analysts.

Future Work

Future enhancements will focus on integrating real-time streaming data sources through Apache Kafka connectors, enabling the agent to monitor live dashboards and issue proactive alerts. Fine-tuning domain-specific LLMs on curated analytical corpora is expected to further improve accuracy on specialized queries. Additional work will explore multi-agent frameworks where specialized sub-agents collaborate on large-scale analytical tasks. Development of a mobile companion application and voice-based query interface will increase accessibility for field practitioners.

REFERENCES

[1] Brown, T. B., et al., "Language Models are Few-Shot Learners," Advances in Neural Information Processing Systems (NeurIPS), 2020.

[2] Wei, J., et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," NeurIPS, 2022.

[3] Schick, T., et al., "Toolformer: Language Models Can Teach Themselves to Use Tools," arXiv, 2023.

[4] Yao, S., et al., "ReAct: Synergizing Reasoning and Acting in Language Models," ICLR, 2023.

[5] OpenAI, "GPT-4 Technical Report," arXiv:2303.08774, 2023.

[6] Anthropic, "Claude: A Family of Large Language Models," Technical Report, 2024.

[7] Chase, H., "LangChain: Building Applications with LLMs," GitHub Repository, 2023.

[8] Pedregosa, F., et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.

[9] Taylor, S. J., & Letham, B., "Forecasting at Scale," The American Statistician, vol. 72, no. 1, pp. 37-45, 2018.

[10] Liu, H., et al., "AgentBench: Evaluating LLMs as Agents," arXiv:2308.03688, 2023.