Date :

Form Name: Monthly Report/Activity Report Issue No: 01

**MONTHLY REPORT / ACTIVITY REPORT**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Project Title** | **Ilmu** | | | |
| **Prepared By:** | **Hamza Khaled Mahmoud** | | | |
| **Report Frequency** | **MONHTLY PROGRESS REPORT** | | | |
| **Reporting Period:** | **From** | **1st August/2025** | **To** | **31th August/2025** |

**Activities Completed for the reporting period:**

|  |  |
| --- | --- |
| **Week** | **Description Of Work Done & New Skills Learnt** |
| 1st week  (2/8/2025 – 6/8/2025) | Designed and implemented a Prefect-Orchestrated Pipeline: Architected a robust document processing workflow using Prefect, ensuring reliability, concurrency, and automated retries for performance.  Integrated Multi-Modal Processing: The pipeline incorporates a multi-stage process to handle diverse data types:  High-Fidelity Conversion: Utilized MinerU for superior PDF-to-Markdown conversion, detecting images , tables , equations.  VLM Enhancement: Integrated a Vision-Language Model (Gemini 2.5 Flash) to analyze images within documents and generate descriptive text.  Semantic Chunking: Implemented context-aware text segmentation to maintain the logical coherence of the source data.  Established the Core Knowledge Infrastructure:Implemented LightRag , changed the parameters to best suite our use case , extracting relationships from documents.  Skills Demonstrated: Data Engineering, System Architecture, Workflow Orchestration (Prefect), Multi-Modal Data Processing, AI Infrastructure Setup. |
| 2nd week  (9/8/2025 – 13/8/2025) | Developed a Multi-Agent System: Built an advanced agentic workflow using LangGraph for state management. The architecture features specialized agents (Query, Answering, Corrective) that collaborate to handle user requests with adaptive logic.  Built an Automated Evaluation System: In parallel with development, I created the "Evaluation Providers" system—a dedicated framework to automate model testing and analysis.  Integrated with MLOps Tooling: The evaluation system was integrated directly with Weights & Biases (W&B), enabling seamless experiment tracking, visualization, and comparative analysis from the outset.  Skills Demonstrated: Agentic AI System Design (LangGraph), MLOps (W&B), Test-Driven Development, API Integration, Python Development. |
| 3rd week  (16/8/2025 – 20/8/2025) | Developed a Comprehensive EDA Suite: Created a suite of scripts for Exploratory Data Analysis (EDA) to systematically generate the high-quality visualizations and statistical summaries used in the final reports.  Finalized Two Major Reports: Completed the writing and formatting of the two distinct reports from Week 3, ensuring they presented a clear, data-driven narrative.  Synthesized a Business Case for Investment: Strengthened the reports by integrating case studies on BloombergGPT, GatorTronGPT, and Harvey AI, contextualizing the technical findings and justifying the need for strategic investment in AI infrastructure.  Consolidated Project Code and Documentation: Organized all developed systems, pipelines, and analysis tools into a clean, well-documented repository for future use and handover, not pushed on gitlab yet.  Skills Demonstrated: Technical Writing, Data Visualization (EDA), Business Case Development, Strategic Recommendation, Project Completion & Documentation. |
| 4th week  (23/8/2025 – 31/8/2025) | Conducted a comprehensive comparative analysis of leading cloud models (OpenAI GPT-4o Mini, Gemini 2.0 & 2.5 Flash).  Evaluated models on accuracy, latency, and cost, creating a clear cost-performance value proposition.  Authored the "Comparative Analysis of Leading Language Models" report, formally recommending Gemini 2.0 Flash as the optimal choice for the project.  Task 2: On-Premise Hardware Viability Assessment  Executed a rigorous stress test on the NVIDIA L40S GPU to determine its suitability for hosting open-source LLMs.  Conclusively proved that the hardware was inadequate due to VRAM saturation (91.3%), leading to an 80% failure rate under load.  Authored the "L40S GPU Stress Test Results" report, providing definitive evidence that a significant hardware upgrade is necessary for on-premise deployment. |

Reviewed by:

(HOU/SV signature)

---------------------------------------

Name

Designation

Unit