

	BẢNG MÔ TẢ CÔNG VIỆC	
--	-----------------------------	--

Chức danh	Chuyên viên Phát triển AI	Cấp bậc	Chuyên viên
Phòng/ban	Phòng Công nghệ thông tin	Địa điểm làm việc	38, Nguyễn Huệ, phường Sài Gòn, Thành phố Hồ Chí Minh
Báo cáo trực tiếp cho	Trưởng phòng CNTT	Báo cáo chức năng cho	Giám đốc CNTT

NHIỆM VỤ CHÍNH

Nhiệm vụ 1: Phát triển chatbot và AI Agent trên nền tảng LLM phục vụ nhu cầu nội bộ.
Nhiệm vụ 2: Xây dựng pipeline RAG và tích hợp LLM API vào hệ thống của công ty.
Nhiệm vụ 3: Phát triển backend, API và tích hợp giải pháp AI vào các hệ thống nghiệp vụ.
Nhiệm vụ 4: Kiểm thử, tối ưu, vận hành và viết tài liệu kỹ thuật cho các ứng dụng AI.
Nhiệm vụ 5: Thực hiện các nhiệm vụ khác do Quản lý Bộ phận AI phân công.

CHI TIẾT CÔNG VIỆC

Chi tiết công việc	Tỉ trọng
Nhiệm vụ 1: Phát triển chatbot và AI Agent trên nền tảng LLM - Lập trình chatbot/trợ lý ảo nội bộ bằng Python trên nền tảng LLM (OpenAI, Anthropic Claude, Gemini hoặc tương đương); - Thiết kế và triển khai AI Agent có khả năng function/tool calling để tự động hóa tác vụ và quy trình nghiệp vụ; - Xây dựng, tinh chỉnh prompt và system prompt; đánh giá chất lượng đầu ra và xử lý các trường hợp lỗi (hallucination, sai định dạng); - Phát triển giao diện hội thoại và tích hợp chatbot vào các kênh nội bộ (web, Zalo, Telegram, MS Teams...) theo yêu cầu; - Ghi nhận phản hồi người dùng, cải tiến sản phẩm theo từng phiên bản và báo cáo tiến độ lên Quản lý.	25%
Nhiệm vụ 2: Xây dựng pipeline RAG và tích hợp LLM API - Xử lý dữ liệu, chunking, embedding và quản trị vector database (pgvector, Chroma, Qdrant hoặc tương đương);	30%

- Xây dựng pipeline RAG truy xuất tài liệu nội bộ; đo lường và cải thiện độ chính xác truy hồi (retrieval); - Sử dụng các framework LLM (LangChain, LlamaIndex hoặc tương đương) để điều phối workflow xử lý; - Tích hợp LLM API vào ứng dụng: xử lý streaming, rate limit, retry và lưu vết (logging) hội thoại; - Theo dõi, kiểm soát chi phí token và tối ưu việc lựa chọn model cho từng bài toán.	
Nhiệm vụ 3: Phát triển backend, API và tích hợp hệ thống - Viết API/service (FastAPI, Flask hoặc Node.js) phục vụ các ứng dụng AI; - Kết nối, truy vấn và xử lý dữ liệu từ database (SQL/NoSQL) và các hệ thống nội bộ (ERP, CRM, HRM...); - Quản lý mã nguồn bằng Git; đóng gói và triển khai ứng dụng bằng Docker theo hướng dẫn; - Phối hợp với bộ phận hạ tầng và an toàn thông tin để đảm bảo bảo mật dữ liệu, tuân thủ quy định của công ty.	20%
Nhiệm vụ 4: Kiểm thử, tối ưu, vận hành và tài liệu kỹ thuật - Viết unit test, thực hiện kiểm thử chức năng và đánh giá (evaluation) chất lượng đầu ra của model; - Theo dõi log, xử lý sự cố và tối ưu hiệu năng – chi phí của các hệ thống AI đang vận hành; - Viết tài liệu kỹ thuật, tài liệu hướng dẫn sử dụng và bàn giao cho các phòng ban; - Cập nhật công nghệ AI mới, chia sẻ kiến thức và tham gia đào tạo nội bộ theo định hướng của Quản lý.	15%
Nhiệm vụ 5: Thực hiện các nhiệm vụ khác do Quản lý Bộ phận AI phân công	10%

MỐI QUAN HỆ TRONG CÔNG VIỆC

Nội bộ (Các phòng ban nội bộ trong công ty):	Bên ngoài (Các đối tác bên ngoài):
Quản lý Bộ phận AI Các phòng/ban liên quan Phối hợp tác nghiệp và hỗ trợ ứng dụng AI theo yêu cầu	Nhà cung cấp nền tảng AI (OpenAI, Anthropic, Google AI...) Đối tác cung cấp công cụ và plugin AI Cộng đồng AI khi tham khảo tài liệu và công cụ mới

YÊU CẦU:**1. Kiến thức, Kinh nghiệm, Kỹ năng chuyên môn**

- Tốt nghiệp Đại học các ngành Công nghệ thông tin, Khoa học máy tính, Khoa học dữ liệu, Toán – Tin, Kỹ thuật phần mềm hoặc các ngành liên quan;
- Ưu tiên ứng viên mới tốt nghiệp hoặc dưới 02 năm kinh nghiệm, có đam mê AI và đã có sản phẩm/dự án thực tế (đồ án, portfolio, GitHub);
- Thành thạo lập trình Python; nắm vững cấu trúc dữ liệu, thuật toán và lập trình hướng đối tượng;
- Có kiến thức và kinh nghiệm thực hành với LLM: gọi API (OpenAI, Anthropic, Gemini...), prompt engineering, function/tool calling;
- Đã từng xây dựng chatbot hoặc AI Agent (dự án học tập, cá nhân hoặc thực tập);
- Hiểu nguyên lý RAG: embedding, vector database, chunking và cách đánh giá chất lượng truy hồi;
- Biết sử dụng framework LLM (LangChain, LlamaIndex...) và thư viện xử lý dữ liệu (pandas, numpy);
- Biết phát triển API với FastAPI/Flask, làm việc với database SQL và các định dạng REST/JSON;
- Sử dụng được Git; có kiến thức cơ bản về Docker, Linux là lợi thế;
- Biết sử dụng công cụ automation/no-code AI (n8n, Make, Dify hoặc tương đương) là lợi thế;
- Đọc hiểu tốt tài liệu kỹ thuật tiếng Anh.

2. Kỹ năng mềm

- Tư duy logic, khả năng phân tích và giải quyết vấn đề (debug, truy tìm nguyên nhân gốc);
- Khả năng tự học nhanh, chủ động nghiên cứu và thử nghiệm công nghệ mới;
- Kỹ năng giao tiếp, trình bày và diễn giải vấn đề kỹ thuật cho người không chuyên;
- Kỹ năng làm việc nhóm, phối hợp hiệu quả với các bộ phận nghiệp vụ;
- Cẩn thận, có trách nhiệm với chất lượng mã nguồn và tính bảo mật dữ liệu;
- Sử dụng thành thạo tin học văn phòng (Word, Excel, PowerPoint, Google Workspace).

3. Yêu cầu khác (nếu có)

- Độ tuổi: Ưu tiên ứng viên mới ra trường hoặc dưới 02 năm kinh nghiệm.
- Có portfolio/GitHub với sản phẩm AI đã triển khai hoặc chứng chỉ liên quan (AI/ML, Cloud) là lợi thế.