Lê Minh Triết

Thực tập sinh AI Engineer

Conversely triet the strict of the strict o



Quận Hoàng Mai, Hà Nội

MUC TIÊU NGHÊ NGHIỆP

Tôi là Lê Minh Triết, một sinh viên yêu thích Toán học, lập trình, dữ liệu và thiết kế hệ thống. Mặc dù chưa có nhiều kinh nghiệm, tôi tự tin với khả năng tự học, tự đọc, thành thạo với dữ liệu và AI. Tôi muốn theo đuổi vị trí trong ngành AI, nơi có thể áp dụng kiến thức về học máy, dữ liệu lớn và kiến trúc hệ thống để xây dựng, triển khai và tối ưu hóa các giải pháp AI vào ứng dụng thực tế, góp phần giải quyết các vấn đề kinh doanh. Tôi tin rằng với đam mê và kỹ năng của mình và khả năng chịu áp lực tốt, tôi sẽ hoàn thành tốt mọi nhiệm vụ được giao với năng suất cao

HOC VÂN

ĐẠI HỌC BÁCH KHOA HÀ NỘI 2022 - Hiện nay Khoa học máy tính • CPA: 3.74/4.0 Trường THPT Chuyên sơn la 2019 - 2022 Chuyên Toán Tốt nghiệp loại giỏi

KINH NGHIỆM

Công ty cổ phần VCCorp 03/03/2025 - 30/05/2025

Thực tập sinh Data Engineer

- Trau dồi các kiến thức nền tảng về máy tính, hệ điều hành, mạng máy tính, bảo mật.
- Luyện tập và nâng cao kiến thức, kĩ năng về OOP, Java backend, thiết kế, xây dựng chuẩn hóa cơ sở dữ liệu.
- Xây dựng ứng dụng cơ bản về đồng bộ file bằng Restful API trong hệ thống client-server.
- Làm quen cách làm việc trong môi trường doanh nghiệp, giao tiếp hiệu quả với leader và đồng nghiệp.

CÁC KỸ NĂNG

| CAC KI NANG | | |
|----------------------------------|--|--|
| Lập trình và nền táng | Ngôn ngữ lập trình: C, C++, Java, Python, HTML, CSS, Javascript, Typescript Kiến thức chuyên sâu về Cấu trúc dữ liệu và giải thuật và lập trình hướng đối tượng (OOP). | |
| Cσ sở dữ liệu | Cơ sở dữ liệu: Thành thạo truy vấn SQL và có kinh nghiệm sử dụng các cơ sở dữ liệu quan hệ (MySQL, PostgreSQL, SQL Server) và phi quan hệ (MongoDB, Redis) Có kinh nghiệm xây dựng các quy trình dữ liệu (ETL) và chuẩn hóa, xây dựng mô hình dữ liệu. | |
| Backend | Biết cách xây dựng, quản lý phiên bản các API kiểu RESTful. Có hiểu biết về kiến trúc Microservices và cách quản lý các dịch vụ và tài nguyên một cách hợp lý. Có kinh nghiệm sử dụng FastAPI của Python để quản lý phụ thuộc hiệu quả. | |
| AI - ML | Machine Learning: Sử dụng thành thạo scikit-learn để làm việc với dữ liệu và huấn luyện mô hình Deep Learning - NLP: Hiểu về các mạng sâu (CNN, RNN và GAN), về Transformers, LLM và biết cách s dụng Pytorch, Hugging Face để xây dựng và fine-tune các mô hình DL và mô hình ngôn ngữ (LLMs). | |
| Big Data | Hiểu biết về xử lý dữ liệu theo Batch và xử lý dữ liệu Real-time. Hiểu biết về kiến trúc, hoạt động và cách triển khai các công cụ/hệ thống dữ liệu lớn như Hadoop, Spark, Kafka, Hive, HBase. | |
| Kiến trúc và triển khai hệ thống | Có khả năng thiết kế và xây dựng kiến trúc hệ thống phần mềm dễ bảo trì và mở rộng. Biết cách sử dụng Linux, Docker, Kubernetes để triển khai hệ thống. Sử dụng thành thạo Git, GitHub để kiểm soát và quản lý mã nguồn. Có kiến thức cơ bản về Cloud và Google Cloud Platform. | |
| Kỹ năng mềm | Kỹ năng thuyết trình và làm việc nhóm, giải quyết vấn đề, tự học, tự đọc tài liệu. Biết cách sử dụng AI để tăng tốc độ và hiệu suất làm việc. | |

HỆ THỐNG HỖ TRỢ ĐẦU TƯ CỔ PHIẾU - ITAPIA (MVP)

6/2025 - 8/2025

Người phát triển dự án

Mã nguồn dự án: https://github.com/triet4p/itapia

Số lượng người tham gia: 1

Công nghệ sử dụng: Python (numpy, pandas, scikit-learn, fastapi, pytorch), Jupyter Notebook, PostgresSQL, Redis, Docker, Hugging Face, SpaCy, TALib

Mô tả: ITAPIA (Intelligent & Transparent AI-Powered Personal Investment Assistant) là hệ thống hỗ trợ đầu tư full-stack dựa trên kiến trúc microservices, với triết lý cốt lõi là Khả năng Giải thích (Explainability - XAI) và cá nhân hóa sâu.

- Thiết kế và xây dựng hệ thống phân tích đa luồng: Tự động thu thập, xử lý và tổng hợp tín hiệu từ nhiều nguồn dữ liệu cho các mục đích khác nhau:
 - Chỉ báo kĩ thuật: Sử dụng TA-Lib để tính toán chỉ báo kĩ thuật, mẫu hình nến, phát hiện các mức hỗ trợ, kháng cư.
 - Phân tích tin tức: Sentimen Analysis (BERT), NER (Spacy), Keyword-Based Impact Assessment.
 - Dự báo kĩ thuật: Sử dụng mô hình ML (LightGBM) để hỗ trợ đưa ra các dự báo theo các task (dự đoán chạm ngưỡng, dự đoán phân phối).
- Phát triển một Rule Engine minh bạch và dễ truy vết: Trái tim của hệ thống là một hệ thống các Rule để tổng hợp các loại tín hiệu phân tích khác nhau thành quyết định (lời khuyên) cuối cùng cho người dùng.
 - Thiết kế một "ngôn ngữ" quy tắc nội bộ dựa trên Cây Biểu thức Tượng trưng (Symbolic Expression Trees), cho phép biểu diễn các chiến lược phức tạp, hỗ trợ định kiểu ngữ nghĩa mạnh để tạo các Rule có ý nghĩa (Semantic Typing).
 - Hệ thống không chỉ đưa ra quyết định mà còn cung cấp bằng chứng có thể truy vết, giải thích rõ ràng "tại sao" đằng sau mỗi lời khuyên về Ouvết định, Rủi ro và Cơ hôi.
- Xây dựng và triển khai Backend API hiệu suất cao:
 - Sử dụng FastAPI và Pydantic để tạo ra các API được định kiểu mạnh (strongly-typed), tự động validate dữ liệu và tự sinh tài liệu OpenAPI.
 - Kiến trúc được thiết kế theo các nguyên tắc phần mềm hiện đại, tách biệt rõ ràng các lớp Service và CRUD, đảm bảo code sạch sẽ, dễ kiểm thử và có khả năng mở rộng cao.
- Xây dựng Giao diện Người dùng (SPA) hoàn chỉnh:
 - Phát triển một ứng dụng frontend bằng Vue.js & TypeScript, cho phép người dùng tương tác với hệ thống, quản lý các hồ sơ đầu tư cá nhân hóa và xem các báo cáo phân tích chi tiết.
 - Triển khai luồng xác thực an toàn end-to-end sử dụng OAuth2 và JWT.
- Được thiết kế để sẵn sàng tích hợp các thuật toán tiến hóa và Agentic LLM trong tương lai.

XÂY DỰNG HỆ THỐNG BIG DATA TRÊN KUBERNETES

12/2024 - 4/2025

Người phát triển dự án

Mã nguồn dự án: https://github.com/triet4p/app-log-process-k8s

Số lượng người tham gia: 1

Công nghệ sử dụng: Python, Docker, Kubernetes, Hadoop, Spark, Kafka, Hive

Mô tả:

- Xây dựng hệ thống Big Data real-time xử lý log streaming với kiến trúc Kafka → Spark Streaming → HDFS → Hive, triển khai hoàn toàn trên Kubernetes cluster.
- Thiết kế và containerize các thành phần chính (Hadoop, Spark, Kafka, Hive) thành custom Docker images, tối ưu hóa cấu hình để tương thích với môi trường K8s.
- Triển khai Hadoop Cluster distributed trên K8s với NameNode, DataNode, YARN ResourceManager/NodeManager, cấu hình HDFS replication và ACL permissions
- Tích hợp Spark Operator với Kubernetes để quản lý Spark applications, xử lý streaming data từ Kafka topics và lưu trữ định dạng Parquet trên
- Cài đặt Hive Metastore với PostgreSQL backend, tạo external tables để query dữ liệu từ HDFS, cung cấp data warehouse layer cho business intelligence.

CHÚNG CHỈ

| Problem Solving (Immediate), SQL (Immediate), Software Engineer, Python, Java. | 2023 (By HackerRank) | | |
|--|----------------------|--|--|
| DANH HIỆU VÀ GIẢI THƯỚNG | | | |
| Đạt danh hiệu Expert trên Kaggle. | 2024 | | |
| Giải Nhì kỳ thi HSG Quốc gia môn Toán. | 2022 | | |
| Giải Ba kỳ thi HSG Quốc gia môn Toán. | 2021 | | |
| | | | |