Tran Huu
Tuan (Toby)

SOFTWARE ENGINEER

INFORMATION

- Feb 16th ,1998
- Hanoi, Vietnam

CONTACT INFO

- +84 32 657 8866
- tuanhuu162@gmail.com

SOCIAL

in Huu Tuan

S @tuanhuu162

LANGUAGES

- ⭐ Vietnamese
- 🇬🇧 English

ORIENTATION

I always pay all of my energy and concentration to become a responsible and reliable Software Engineer. I am looking forward to the opportunity to work in a competitive environment where I can utilize my skills and realize y my potential, while contributing to company's productivity.

EDUCATION

2020 | 2016

Bachelor's Degree of Information Technology
GPA 8.875 / 10 | 3.55 / 4
University of Engineering and Technology, Vietnam National University Hanoi

ACTIVITIES / EXPERIENCES

Now | AUG 2021

BACKEND ENGINEER IN PARCEL PERFORM PTE LTD.
- Working as Mid-level Backend Engineer to develop e-commerce SaaS with scalable systems with microservices-based using Flask, Django and FastAPI
- Maintainer of synchronization CDC service that sync main database to Data WareHouse(Redshift, OpenSearch) and DataLake(Apache Hive, IceBerg) used for Logistics Intelligence and Data Analytics, as well as design high-traffic data synchrony pipeline using Flink, Dagster and Debezium
- In charge of expect delivery date and SLA features that generate required data for customer
- Working with multiple teams across world-wide offices(Singapore, Netherlands, US,...) to develop mutiple features to integrate with e-commerce merchants and their carrier partners in scraping, receiving, standardizing and storing data
- Taking part in the solution design step and writing technical design for multiple epics.
- Co-operating with Data Science team to deploy ML model with high availability and reliability using MLflow and FastAPI
- Mentoring and code review for junior member of team to maintain the code quality and team's velocity

AUG 2021 | MAR 2018

MLops ENGINEER IN Aimesoft JSC.
- Experiencing models about text and speech for emotion recognition, text to speech, happiness score evaluate, speech recognize.
- Investigating, preprocessing data to build ML model about psychologic, fintech, etc and serving them using FlaskAPI.
- Implementing web scraping and data cleaning flow for e-commercial website like adidas, Uniqlo, Nike, etc. then build searching engine using Flask API and Solr.
- Cooperating with other teammates to create and maintain AI automatic solutions for advertisement, customer care, etc.
- Implementing POC demo for company

PROJECTS

2023

Asynchronous file upload

Designing technical solution and implementing Web service allow user submit file with millions of rows to API interface for scraping web pages asynchronously under Python threading before save result to PostgreSQL for further analytics. This pipeline using Django, Django ninja, Kafka MQ and PostgreSQL, along with existed service written by Scrapy

2023

Email extraction service

Designing technical solution, mentoring and implementing with teammates a data pipeline that process authorized email content to extract shipping information using regex, xpath and OpenAI LLM, sync directly from Google API to our internal service. Then the email content is pushed to KafkaMQ before send to extraction service to extracting shipment data then added to our database. The extraction service is written with Django and Python Kafka consumer, using PostgreSQL for storing pattern, having Django Ninja to have CRUD and data validation for patterns management and deployed by K8s Job for easy scaling

2023

ML API service for ML model serving

Designing technical solution and implementing API for serving ML model, with high response time and less resources. Using MLFlow as model management and Model serving server, FastAPI service as a web proxy for redirecting request to deployed model, authorizing and payload verification. The server is deployed all on AWS with K8s as cloud management.

Redshift and Datalake data pipeline by WAL CDC

2022

Implementing CDC pipeline with teammates to adapt syncing 4-5 millions of records per day between DBs by WAL CDC. The pipeline consumes CDC messages from Kafka which is triggered via Debezium connector, stream to table API with Scala Flink before importing collected data-by-batches to Redshift DB with Dragster sensor. Another part of the pipeline will have Dagster jobs trigger Spark clusters to compact Hive output data from Flink job before import them to Aws Athena for query as a Datalake

2022

Maintain and update legacy code of Flink pipeline

Working with code base of 22 Flink Jobs to maintain the data flow work well in high traffic, resolve system issues and developing new feature in data stream.

2022

Maintain and update legacy code of Logistic Intelligence service

Working with legacy Django code base to refactor and provide new features for charts/reports management system, with multiple API using Django, Django Rest Framework, Jaspersoft(For chart visualize) and PostgreSQL.

2021

Speech Synthesis service(http://tts.gpu01.aimesoft.com/)

Researching and applying speech synthesis models like FastSpeech and Tacotron to build real-time speech synthesis service for Japanese, Vietnamese, English and Korean using FlaskAPI, ReactJS and Pytorch with minimum resource requirement (with/without GPU). The speech is streaming using websocket

2020 | 2019

Automl platform(https://forecastinge.com/index_e.html)

Cooperating with other teammates to create and maintain AI automatic solutions for advertisement, customer care, etc. By taking advantages of pandas, the service can handle file data up to 5gb with single upload with support modification, append rows/columns, enrich data by fetching from other sources before predicting result by h2o library and AWS Sagemaker to train multiple ML models then visualize it into confusion matrix or charts with plotly. The CRUD API is implement using Flask and serving on a EC2 machine using docker-compose. The UI created with ReactJS

SKILLS

Foundational skills	- Using visualization tools like Mermaid, drawio to creating diagrams, charts for system designing and documentations. - Technical writing skill for system design and solution. - Having knowledge about design pattern, entity framework and etc. - Be able to work on environment Unix/Linux based and development tools such as Docker, Git - Be able to work with AWS, Kubetnetes and Infrastructure as Code using Terraform to get the most out of infrastructure - Be able to work with monitoring tools like Grafana, Prometheus to maintain services, and sensitive in debugging - English communication with product team/customers for requirements extraction
Python	- Familiar with Python 3 on development and maintenance backend services - Creating Web API from scratch using Django, Flask, FastAPI with full CI-CD pipeline using Gitlab CI and deploy via Docker, Docker compose and k8s for multiple purpose like Web API, scheduled job, migration tools, ... - Using Python and its extensions / libraries for Web Application Development in common frameworks: Django, Flask, FastAPI, etc. and data processing framework: Pandas, Numpy, etc.
Database	- Be able to work with multiple databases about both SQL(PostgreSQL) and NoSQL(Redis, Elasticsearch) - Working with PostgreSQL for data management with billions of records, Redshift for Data Warehouse
Message Queues	- Having knowledge about Kafka MQ in microservice communications and operating data pipeline incoming system
Distributed Processing Engine	- Working with Flink framework to build and maintain pipeline which can process TBs/day - Having knowledge about streaming processing, event driven architecture and scaling application with Flink and K8s
Machine Learning	- Working in Machine Learning and Deep learning with knowledge in NLP and Speech processing field - Building and executing ML/DL models and apply them into service using Python Web framework like Django, Flask, FastAPI with model serving framework like MLFlow
Orchestrate Task	- Working with Dagster to create scheduled jobs/sensor for annually reports/data extraction - Setting up Dagster for syncing database to database

ALWAYS READY AND WILLING TO LEARN NEW THINGS!

AWARDS / HONORS

2020 | - Scholarship of VinGroup for talent student