NGUYỄN MINH TRUNG

MLOPS ENGINEER

NGUYỄN MINH TRUNG

1997

0363692327

nghetrung@gmail.com

A2 Vinhomes Gardenia, Hàm Nghi, Hà Nội

nghetrung.github.io

CAREER OBJECTIVE

Finding a position as a Data Scientist/ML Engineer at a Data-driven organization that has a well-defined data analysis pipeline. Be able to perform analytics about sales, products, cost, profit, and customer behavior. Have a chance to practice building end-to-end data driven system and applying Deep Learning models, MLOps and Cloud.

SKILLS

Machine Learning

- Sklearn, Keras, Pytorch, PySpark

Programming Language

- SQL, Python and R for Data Analysis and Modelling

BI Tools

- Tableau, PowerBI, Apache Superset

Data Pipeline

Apache Airflow, PostgreSQL, Redis, Milvus, MinIO, Kafka

MLOps

Kubernetes, Docker, MLFlow, SeldonCore, Google Cloud Platform

EDUCATION

University of Sydney

Bachelor of Computer Science and Technology

03/2015 - 07/2018

Graduated with Distinction grade, GPA 7.9:

- High Distinction grades in Math and Statistics subjects
- High Distinction grade in Algorithm and Complexity subject

University of Melbourne

Master of Data Science

03/2019 - 01/2021

Graduated with High Distinction grade, GPA 8.1

- Distinction grade in Statistical Machine Learning including knowledge about the Math behind different Machine Learning model such as Linear & Logistic regression, Support Vector Machine, Random Forest and Neural Network
- Distinction grade in Natural Language Processing including knowledge about Text processing, N-gram model, POS tagging, Contextual representation and Context-Free Grammar
- Multiple statistics modelling subjects taught using R about different distributions, significance test, confidence interval
- Data Science project to help non-profit organization analysing their volunteer. Built dashboard using R-Shiny.

EXPERIENCE

Viettel AI

Machine Learning Engineer

12/2022 - Present

Data Analytics

- Analysing user watching behaviours on Myclip.vn
- Clean and prepare log data
- Design A/B testing strategy for video recommendation system
- Design database and data pipeline for the recommendation system
- Collect user clickstream behaviours data on website to measure the performance of recommendation system
- Experiment building Kedro pipeline and MLflow for Machine Learning

Deployment

- Design pipeline to migrate recommendation system from on-prem to Kubernetes clusters
- Deploy Kafka for data ingestion
- Deploy Postgresql, Redis, Milvus for data storage on Kubernetes using Helm Chart
- Deploy Airflow and run batch jobs to generate artifacts for recommendation model
- Deploy streaming jobs to receive data from kafka topics
- Write Helm chart to deploy Recommendation API server application
- Setup Prometheus and Grafana for performance monitoring
- Build an API Gateway app to split traffic for two different recommendation models.
- Apply ArgoCD for automatic deployment

BIDV - Bank for Investment and Development of Vietnam

Data Analyst at Digital Banking Center

03/2021 - 11/2022

- Perform data cleaning and transformation to prepare data as input for the model using Python and PySpark
- Using SQL to query, and prepare data for further analytical tasks
- Analyze and build a Machine Learning model to get insights into Customer churn behavior
    - Analyze customer churn for current account product
    - Perform EDA to explore the distribution of the classes and different features. Discovered that churn group tends to have decreasing balance in the recent months
    - Random Forest model to train and predict with 90% F1-Score
- Design and build analytic dashboards to provide information for the business team (Tableau)
    - Monthly reports on the Bill Payment system performance
    - Overall dashboard for business performance on different channels
- Perform customer segmentation analysis using RFM techniques and showcase results on an interactive dashboard
- Applying OCR technology to extract information from images using PyTorch framework

- Research and apply face recognition framework to build an attendance system