

Khanh Duong TRAN

Data Scientist Intern | Île-de-France

 raydiwill  Khanh Duong TRAN  duong.tranhn1102@gmail.com

 +33 6 62 48 81 52

SUMMARY

A "Two-for-one" hire with expertise as both Data scientist and ML (Machine learning) engineer, able to handle both job types. Open-minded, results-driven, and business-centric, with experience in telecom (telecommunications) ML use cases. Bilingual in English with intermediate French understanding, eager to solve real-world problems.

PROJECTS

BANKING CUSTOMERS CHURN PREDICTION

| DATA SCIENTIST

Objective: To built and deploy ML models into production.

- Conducted Exploratory Data Analysis on customer data.
- Classified churners using tree-based classification models.
- Utilised scikit-learn pipelines for efficient feature engineering.
- Implemented feature selection based on feature importance to reduce training time, and feature generation to enhance model performance.
- Applied SMOTE and class weight adjustments to tackle class imbalance.
- Implemented Optuna for hyperparameter tuning across models for optimal performance.

CHURN PREDICTION ANALYTICS PLATFORM

| ML ENGINEER

Objective: Researched state-of-the-art deep learning models to develop a Minimum Viable Product for telecoms churn prediction.

- **Model Development:** Implemented a hybrid ensemble DNN and BiLSTM model, Logged models metadata via MLflow.
- **Model Deployment:** Deployed the model into a web application using FastAPI for real-time predictions.
- **Data Pipeline Inetegration:** Integrated data quality validation using Great Expectations in production pipeline.
- **Process Automation:** Automated model prediction tasks using Apache Airflow Docker image.
- **Stakeholder Engagement:** Developed, presented interactive dashboards using Grafana and Streamlit to demonstrate model performance.

EXPERIENCE

VIETTEL GROUP

April 2022 - June 2022

| ASSOCIATE DATA ANALYST

| Hanoi, Vietnam

- Analyzed telecoms customer usage data using SQL.
- Assisted in reporting the ML applications' effectiveness in telecoms usecases.
- Researched tree-based ML algorithms for customer churn predictions.

SKILLS

PROGRAMMING

Proficient:

Python • SQL

Familiar:

R • Java • Shell Script •
C/C++ • Javascript • HTML
& CSS

LIBRARIES/Frameworks

Scikit-learn • Tensorflow •
Numpy • Pandas • Polars •
FastAPI • Optuna •
Matplotlib • Streamlit

TOOLS/PLATFORMS

Docker • AWS • Git •
GitHub • MongoDB •
Apache Airflow • MLflow •
MySQL • Grafana • Tableau
• Dataiku

OPERATING SYSTEM

Unix

LANGUAGES

- English (C1)
- French (B1)
- Vietnamese

SOFT

- Teamwork
- Problem solving
- Inquisitive
- Adaptability
- Autonomous

EDUCATION

EPITA

M.S.C IN COMPUTER SCIENCE

March 2023 - Present | Le Kremlin-Bicêtre

INTERESTS

- Photography
- Guitar
- Cue sports