

SUMMARY

I am interested in Machine learning, CV/NLP, Modern Data Architecture, Cloud Computing and Algorithms. My goal is to become an excellent programmer, applying artificial intelligence to improve the quality of our lives.

EXPERIENCE

- **Techcombank - DnA Division** Hanoi
Data Engineer - Full time Jun 2022 - present 2022
 - **Integration to CRM system project:** Main responsibility for design and implement Bulk API (both push and pull) to ingest data into CRM (Salesforce), ETL pipeline
 - Optimizing ETL job (AWS glue, S3, Crawler), data processing, implement speed layer pipeline with Lambda and Kinesis, migrating datawarehouse to the cloud. Having knowledge about SCD, CDC, Data Lake, DataMart
 - **Techstack:** AWS service: Glue, Kinesis, S3, Step Function (for Orchestration), Lambda function, DynamoDB, PostgreSQL engine, SNS, SQS, Crawler; Monitoring (Athena, Grafana); Pyspark, pandas; Terraform (Entry Level)
- **Pionero JSC** Hanoi
Machine Learning Engineer - Full time April 2021 - Jun 2022
 - **Water flow control system:** Main responsibility for planning, building system design and development of functions: Hypothesis testing (statistics and visualization), Flow rate prediction, processing signal from IoT device (MQTT protocol), system design (using AWS service), dashboard.
 - **Image colorize:** Be responsible for research, make solutions, and complete the PoC of image coloring system based on an improvement of an existing paper.
 - **SSO:** Develop account management services with AWS Serverless service. Be responsible for API development, database design, infrastructure building, monitoring. Techstack: Cloudformation, CI/CD, DynamoDB, AWS Lambda, Keycloak, ...
 - **Social listening:** Collect data from twitter and answer sheets, analyze by emotional frame STAR.
 - **Pionero Demo:** Building demo projects to demonstrate the company's technological strength, for sales: face emotion recognition, chatbot (Rasa), face recognition and monitoring system for webservice.
 - **Techstack:** AWS service: IoT core, Kinesis, S3, EC2, ECR, Lambda function; Monitoring (Prometheus, Grafana); Deep Learning framework: Pytorch, Keras, Sklearn
- **Viettel Big Data Analytics Center - DAC** Hanoi
Data Scientist - Full time July 2020 - Mar 2021
 - Be responsible for choosing, developing ML models at telco projects to boost campaign performance. Include app traveling, up-sell specific services, recommender systems,... Algorithm range could be traditional ML (tree-based, SVM, KMeans, PCA,...) or Deep learning (MLP, Deep factorization machine, word2vec, CNN, RNN,...)
 - Optimizing models, data processing, feature engineering process, performance and speed boosting.
 - Data processing and preparation, build ETL data flow (on pentaho kettle) for data preprocessing.
 - Optimizing for model development process, AutoML, reduce ML development effort while improve performance.
 - Familiar tools: Python (pandas, numpy, sklearn, ...), Tensorflow, Pyspark, SQL (Hive, MySQL), Spark/Hadoop Ecosystem.
 - **Project:** Potential customer (app MyVT), Voucher classification, Voucher recommendation system, Product Selection (Onme)
- **Ftech CO., LTD** Hanoi
NLP Engineer - Intern Mar 2019 - Dec 2019
 - **Babe bot:** Building a virtual assistant to support customer care for stores, e-commerce sites etc... Custom Action and NLU model (classification and NER)
 - **Womie bot:** Building chatbot to support meeting room management, work report management.
 - Building Deep Learning model for Language Understandings tasks (Classification, NER and Spell Correction) in Vietnamese.
- **VCCorp** Hanoi
NLP Engineer - Intern Apr 2018
 - Learn and implement about basic problem in Natural Language Processing: tokenization, part of speech (POS), named entity recognition (NER), summarization (single and multi document), ... and machine learning
 - Building keyword extraction model for news
- **Hanoi University of Science and Technology** Hanoi
Information System - ICT - Student Oct 2016 - Jan 2021
 - Study how computers work, write software, and learn about data structures, classical and evolutionary algorithms.
 - Main members in major exercises in algorithms, artificial intelligence, machine learning / deep learning. exercises: Text Classification, Sentiment Analysis, NER, POS, Tokenizer, Text Summarization, Web Crawler, Information Retrieval, Machine Learning from scratch, Add Tone, etc.

EDUCATION

- **Hanoi University of Science and Technology** Hanoi
Engineer's degree; Information System - Information Technology of SoICT *Oct 2016 - Jan 2021*

SKILLS

- Algorithms, Competitive programming, Data Structures.
- Machine learning, Deep learning, Natural language processing, Data Science, Crawl Data.
- Languages: Python, Javascript, Java (basic), SQL, Latex.
- AI frameworks: Pytorch, TensorFlow, Scikit-learn, Keras, Rasa framework, Spark/Hadopp ecos, Pyspark, etc.
- Technologies: Flask, Git, Linux, Scrappy, Docker, PostgreSQL, AWS (Serverless, DynamoDB, SNS, SQS, ...), Monitoring (Prometheus, Grafana).

ACHIEVEMENTS

- **The Tech Savvy at Pionero JSC.** - Jan 2022