**Vuong Tuan Khanh**

khanhvuongtuan37@gmail.com | (+84) 587 601 494

https://www.linkedin.com/in/vuong-tuan-khanh/

**PROFESSIONAL EXPERIENCE**

| **FPT SOFTWARE - HO CHI MINH** | 06/2022 - Present |
| --- | --- |
| *Full-time Mid-Level AI Ecosystem Engineer* |  |

* Engineered AI platforms using Dataiku and Squirro AI platforms.
* Remodeled an Intent Classification Bot with a complete logging system. Data handled using spaCy, Tf-Idf, Bag of words and digest into variety of classification models, including: Logistic Regression, SVC, MultinomialNB, MLPClassifier, Decision Tree, Gradient Boost, Random Forest, Calibrated Classifier.
* Digest data into Neo4J and ElasticSearch to store data. Built difference procedure to query data out of the storage.
* Streaming data from Oracle to Azure Synapse using Azure Data Factory by building different data pipelines and stored procedures.
* Standardized different pipelines and data flows to digest data from Denodo, Mendix into Synapse and SQL Server. Connect different data sources like Blob Storage and Data Lake. Trigger difference Data Factory jobs through Rest API. Do data normalization/denormalization and build different data schema for data mart.

| **VELA CORP. - UBOX CO LTD** | 10/2022 - Present |
| --- | --- |
| *Full-time Mid-Level Data Engineer* |  |

* Operated a My-SQL serverless platform developer with PlanetScale and Neon, aims at horizontal scaling using Vitess, Python SDK, and Prisma as the bridge for UBox app development.
* Streamlined data and multi-Cloud self-hosted Database (AWS and GCP) using YugaByteDB. Branches managed for projects that use DoltDB.
* Distributed SQL developer on CockroachDB. Developed a Full-text search engine using XataDB.
* Graph-DB engineer and front-end developer, focus on modeling the relationship for a recommender system using Neo4J and 8Base. Currently migrating to Memgraph for building a real-time analytic platform.

| **GOBIZ.VN** | 01/2023 - Present |
| --- | --- |
| *Full-time Mid-Level Data Engineer* |  |

* Self-developed the entire Gobiz Data Management Tool project, using Flask Framework for UI, Firebase for Authentication and role-assign feature, and BigQuery for data storage.
* Graph-DB engineer and front-end developer, focus on modeling the relationship for a recommender system using Neo4J and 8Base. Currently migrating to Memgraph for building a real-time analytic platform.
* Self-hosted SurrealDB as a multi-model database, mostly focused on graph and key-value features that served as a distributed cache together with KeyDB on top of some persistence data layer.

| **ATP Corp - OfficeComponent, ComponentPro, Incredium** | 04/2019 - 06/2021 |
| --- | --- |
| *Data Scientist, Web Developer, Shopify Liquid App Developer* |  |

* Front-end developer for https://www.officecomponent.com/ and https://www.severium.com/. Data Analysis and VueJS developer for FireClick, which is a web traffic management system.
* Logical handling of complex query string and data for ProsperTasks - Issues tracker online service.
  + Reference: https://prospertasks.com
* Liquid Shopify Back-end Developer for The Ultimate Search & Filter project.
  + Reference: https://sobooster.com

| **MICROSOFT** | 04/2020 - 07/2020 |
| --- | --- |
| *Engineering Intern at Microsoft* |  |

* Applies engineering principles to solve complex problems through sound and creative engineering.
* Quickly learns new engineering methods and incorporates them into work processes.
* Demonstrates skill in time management and completing software projects in a cooperative team environment.

| **SENDO + CODERSCHOOL** | 10/2019 - 12/2021 |
| --- | --- |
| *Data Scientist Intern + Back-End Developer* |  |

* Mine and analyze data from company databases to drive optimization and improvement of product development. Selecting features, building and optimizing models, measuring the similarity using the Cosine Distance.
* Data mining using state-of-the-art methods. Handled an extremely huge dataset, consisting of 7 Billions records as well as assessed the effectiveness. Created automated recommender systems and constant tracking of its performance.
* Develop custom data models and algorithms to apply to data sets. Implemented predictive models to increase and optimize customer experiences, revenue generation, ad targeting and other business outcomes.

**EDUCATION**

| **HCMC UNIVERSITY OF EDUCATION AND TECHNOLOGY** | 01/2017 - 1/2022 |
| --- | --- |
| *Bachelor of Engineering - Information Technology* |  |

| **FUNIX UNIVERSITY** | 10/2020 - 10/2021 |
| --- | --- |
| *Bachelor Degree - Valedictorian Graduated - Machine Learning* |  |

| **CODERSCHOOL** | 10/2019 - 10/2021 |
| --- | --- |
| *Engineer Degree - Data Science* |  |

**PERSONAL PROJECT**

| **AIRBORNE OBJECT DETECTION** | 01/08/2021 - 24/08/2021 |
| --- | --- |
| *FUNIX Capstone Project* |  |

* Airborne Object Detection project 🚀 is the capstone project on object recognition through images and videos, inspired by the Airborne Object Tracking Challenge.
* Working on 13TB dataset, including 6.0M+ images and 164 hours+ of videos.
* Building YOLOv3, YOLOv4, YOLOv5, as well as implementing the newest SiamMOT paper
* Visualization using TensorBoard, WanDB and NetronViewer. Deployed YOLOv5 Web Service to Heroku using Flask and Docker Image.
  + Docker Image: https://hub.docker.com/r/khanhvuongtuan/yolov5\_web
  + Github Repo: https://github.com/VuongTuanKhanh/Funix-Capstone-Project

| **BRAIN MRI - TUMOR SEGMENTATION** | 07/01/2021 - 07/01/2021 |
| --- | --- |
| *Brain MRI GAN reached top 4 in the National FPT Edu Hackathon 2021* |  |

* Brain MRI GAN is an application that generates MRI T2 Contrast images based on T1 MRI images, reducing half of the time and cost of detecting brain tumor processes using Pytorch, pix2pix and CycleGan.
  + Github Repo: https://github.com/VuongTuanKhanh/Brain-MRI-GAN

| **KANTA - SENDO RECOMMENDER SYSTEM** | 28/09/2019 - 28/11/2019 |
| --- | --- |
| *Verified by Sendo, CoderSchool, and Facebook Developer Circle Viet Nam* |  |

* KANTA is a React Native application that recommends new items based on 7,000,000 user historical activity records.
* KANTA was developed with natural language processing methods, based on TF-IDF and BoW to represent feature matrices, as well as calculate the correlation between products based on different distance metrics.
* Constructed a time-series database architecture of 7,000,000 records from Sen Do E-Commerce Platform using Python.
* Implemented two new Hybrid Recommender Systems and Percentile Range based index method.
  + Github Repo: https://github.com/VuongTuanKhanh/DevC\_2019\_Sendo

**ACHIEVEMENT**

| **Full Scholarship from Facebook Developer Circles** | 2021 |
| --- | --- |
| **Top 1% Full Scholarship from FUNiX AI and KALAPA** | 2021 |
| **Top 4 National FPT Edu Hackathon Artificial Intelligent** | 2021 |
| **SBA Scholarship 2022 from SBA, Intel, and Jabil Vietnam** | 2022 |
| **Top 10 Australia - Top 50 Global Hack4Growth** | 2021 |
| **Top 6 Hackathon Gameloft South East Asia** | 2019 |
| **Top 2 Hackathon HCMUTE** | 2019 |
| **Top 1 Arena Hack Brain from Gameloft South East Asia** | 2018 |
| **Top 1 Shark Project - HCMUTE** | 2018 |
| **Top 4 HCMUTE Hackathon** | 2018 |

**CERTIFICATION**

| * IBM Enterprise Data Science * IBM Enterprise-Grade AI * AWS Software Engineering * Dataiku Advanced Designer * AWS S3 Basic * Microsoft Software Engineer * Getting Started with Power BI | * AWS Solution Architect * Meta Facebook - Databases * Dataiku Developer | * Dataiku Core Designer * Dataiku MLOps * Dataiku Practitioner |
| --- | --- | --- |
| * JPMorgan Chase & Co. Software Engineering * Azure Synapse SQL Pool - Implement Polybase * Facebook Developer Circles Vietnam Certified Data Scientist | |

**TECHNICAL SKILL**

| * TensorFlow * Pytorch * Keras * Hadoop / Spark * Python * R * HTML/CSS/JS * Flask * PlanetScale Database * Yugabyte Database | * Dataiku * Squirro * AWS Cloud * Azure Cloud * Neon Database * Dolt Database * 8Base * EdgeDB * Surreal DB * KeyDB | * TensorBoard * WanDB * Docker * Github / Gitlab * Cockroach Database * Foundation + Titanium Database * CloudFlare Database * Xata Database * Neo4J / Memgraph Database * MindsDB |
| --- | --- | --- |

**OTHER**

* *Computer Science Teacher at Algorithmics Global and MindX*
* *Founder & President of SKY UNI - TOP 10 project in the Australia Hack4Growth 2021*
* *https://github.com/VuongTuanKhanh/Certificates-of-Achievements*
* *https://vnexpress.net/game-thu-hoc-cong-nghe-de-theo-duoi-dam-me-ai-4270629.html*
* *https://funix.edu.vn/chan-dung-xter/tu-nghien-game-den-nam-sinh-cong-nghe-dam-me-tri-tue-nhan-tao/*
* *https://www.sggp.org.vn/trao-36-suat-hoc-bong-sba-cho-sinh-vien-co-thanh-tich-hoc-tap-va-hoat-dong-xa-hoi-noi-bat-816790.html*