

# TRUC-VIET LE

## FULL-STACK DATA / AI SCIENTIST

+65 9420 8417 | [vietexob@gmail.com](mailto:vietexob@gmail.com) | [trucvietle.me](http://trucvietle.me) | [in://truc-viet-le](https://in://truc-viet-le)

### PROFESSIONAL SUMMARY

A seasoned full-stack data/AI scientist with 7+ years' experience in developing and deploying impactful AI solutions across diverse industries. Experienced in banking, financial services, e-commerce, and technology sectors. Expertise in entity resolution, record linkage, data mastering, and deduplication problems, with approved US patents. Enthusiastic about GenAI and LLM technologies to solve real-world problems. Strong advocate for ethical and responsible AI use.

### TECHNOLOGIES & SKILLS

<b>Programming</b>	Python, Java, Scala, HTML5, JavaScript, Bash, SQL
<b>AI Frameworks</b>	scikit-learn, spaCy, Hugging Face, PyTorch, TensorFlow, Langchain, LlamaIndex
<b>Cloud &amp; Data</b>	Hadoop, Spark, PostgreSQL, MongoDB, Redis, Chroma, Neo4j, Docker, Azure, GCP
<b>Skills &amp; Interests</b>	Payment & network analytics, transaction monitoring, knowledge graphs, entity resolution, ETL processes, clustering, anomaly & fraud detection, natural language processing (NLP), LLMs
<b>Soft Skills</b>	Problem-solving, critical thinking, effective communications, creative & strategic thinking, Teamwork, mentorship, stakeholder & project mgmt. (agile, scrum), cultural sensitivity & inclusivity

### EXPERIENCE

**Senior Data Scientist (VP)** | EFG International | Singapore, Singapore **Jan 2024 – Present**

*EFG International is a Swiss boutique private bank with a global reach. The bank is known for its modern entrepreneurial approach to wealth management blended with the Swiss age-old craft of private banking and client confidentiality.*

- Developed bank-wide digital capabilities using GenAI and LLMs on private cloud (e.g., RAG system for internal knowledge search, chatbot), boosting productivity across the organization.
- Leading effort to develop AI solutions for AML investigations (i.e., entity resolution and knowledge graph models of transactions), saving 2-3 man-hours per case.
- Helped establish bank-wide AI governance framework, policy and procedure for AI applications.
- Using anomaly detection on KYC and transaction behaviors to timely update a client's risk profile, resulting in reduced risk and savings of ~600 man-days per year (in periodic review).

**Senior Data Scientist (AVP)** | Credit Suisse AG (Part of UBS Group) | Singapore, Singapore **Dec 2021 – Dec 2023**

- Developed machine learning (ML) solutions for entity resolution for AML purposes, achieving high level of automation and accuracy (>95%) in transaction monitoring (compared to traditional rules-based approach).
- Using advanced NLP, knowledge graphs and network analysis for external entity ontology and data mastering.

**Data Scientist** | Agoda.com (Part of Booking Holdings Group) | Singapore, Singapore **Sept 2019 – Nov 2021**

*Agoda.com is an online travel agency with dominant market share in the Asia-Pacific region. Agoda strives to become a one-stop shop for all travel needs, including accommodations, flights, activities and land transport with competitive pricing.*

- Developed matching algorithms to map and deduplicate properties across different suppliers (e.g., Airbnb, Booking.com, Agoda), improving search and recommendations.
- Deployed in production (on a Spark cluster) mapping pipeline that processes millions of pairs of properties daily with high precision (>97%) and recall (>80%), resulting in ~80% fewer duplicate listings.
- Developed image deduplication algorithms that merge property galleries to improve content and conversion rate.

**Machine Learning Developer** | SAP Asia Pte. Ltd. | Singapore, Singapore

Jul 2017 – Jul 2019

- Part of team effort to automate the enterprise accounting processes (cash application, invoice digitization) using machine learning, resulting in [a successful commercial product](#).
- Invented methods to solve the multi-match problem in cash application (i.e., matching one bank statement to many invoices) using deep learning with three approved U.S. patents.
- Generated structured data from unstructured bank statement (BS) memo lines using named entity recognition (NER), boosting accuracy of matching BS to open invoices to >98%.

## EDUCATION

---

**Ph.D. in Information Systems** | Singapore Management University | Singapore, Singapore

Jan 2013 – Jul 2017

- Thesis: *An integrated framework for modeling & predicting spatiotemporal phenomena in urban environments*

**Visiting Ph.D. Scholar** | Carnegie Mellon University (CMU) | Pittsburgh, USA

Aug 2014 – Jul 2015

- Did coursework, research and teaching assistant at Heinz College of Information Systems and Public Policy

**M.Sc. in Mathematical Sciences** | Nanyang Technological University | Singapore, Singapore

Jan 2011 – Jul 2012

- Thesis: *Pareto stable matchings: An empirical study*

**B.Eng. in Computer Engineering** | Nanyang Technological University | Singapore, Singapore

Aug 2005 – Jul 2009

- Final Project: *A mathematical model of hospital length of stay*

## PUBLICATIONS

---

**Truc Viet Le**, Baoyang Song & Laura Wynter. **Real-time Prediction of Length of Stay Using Passive Wi-Fi Sensing**. The 2017 IEEE International Conference on Communications (IEEE ICC 2017), Internet of Things (IoT) Track, Paris, France.

**Truc Viet Le**, Richard Oentaryo, Siyuan Liu & Hoong Chuin Lau (2017). **Local Gaussian Processes for Efficient Fine-Grained Traffic Flow Prediction**. IEEE Transactions on Big Data (TBD) Special Issue on Urban Computing, 3(2), 194--207.

**Truc Viet Le**, Siyuan Liu & Hoong Chuin Lau. **A Reinforcement Learning Framework for Trajectory Prediction Under Uncertainty**. The 22<sup>nd</sup> European Conference on Artificial Intelligence (ECAI 2016), The Hague, The Netherlands.

**Truc Viet Le**, Siyuan Liu, Hoong Chuin Lau & Ramayya Krishnan. **Predicting Bundles of Spatial Locations from Learning Revealed Preference Data**. The 14<sup>th</sup> International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Istanbul, Turkey.

## PATENTS

---

**Truc Viet Le**, Sean Saito, Rajalingappaa Shanmugamani & Chaitanya Joshi. **A Graphical Approach to the Multi-match Problem**. Issued Jun 11, 2020. Patent no.: US20200184281A1.

Sean Saito, **Truc Viet Le**, Rajalingappaa Shanmugamani & Chaitanya Joshi. **Representing Sets of Entities for Matching Problems**. Issued Jun 4, 2020. Patent no.: US20200175559A1.

Rajalingappaa Shanmugamani, Chaitanya Joshi, Rajesh Arumugam, Sean Saito & **Truc Viet Le**. **Utilizing Embeddings for Efficient Matching of Entities**. Issued Jun 18, 2020. Patent no.: US20200193511A1.

## PROFILES

---

PERSONAL WEBSITE

<https://trucvietle.me>

LINKEDIN

<https://www.linkedin.com/in/truc-viet-le/>

GITHUB

<https://github.com/vietexob>