

BIN XUAN KONG

(+60)126112151 ◊ binxuankong@gmail.com ◊ github.com/binxuankong

Skills

Data Science	Machine Learning, Deep Learning, Computer Vision, Large Language Models (LLMs), Natural Language Processing (NLP), Recommendation Systems, Data Visualization
Programming	Python, SQL, Java
Libraries	NumPy, Pandas, SciKit-Learn, TensorFlow, PyTorch, Transformers
Infrastructure	dbt, AWS (ECS, EC2, Lambda), Docker, Kubernetes, NoSQL
Soft Languages	Problem Solving, Project Management, Adaptability, Fast Learner, Teamwork English (Native or Bilingual), Chinese (Simplified) (Native or Bilingual), Malay (Professional Working), Cantonese (Elementary)

Work Experience

GXBank

Data Scientist, Senior Specialist

Jul 2023 - Present

Selangor, Malaysia

- Developed a name matching model to predict name equivalence, capable of handling localized variations.
- Created a profanity image scanning model to filter out inappropriate images.
- Explored various LLM solutions including OpenAI GPT-4, Google PaLM, Anthropic Claude, and Meta Llama for generative AI applications.
- Built a customer service RAG chatbot using OpenAI GPT-4, featuring functionalities such as security checks, customer support redirection, and FAQ context search.
- Led the VoC360 project: extracted customer comments/reviews from multiple sources, built data pipelines for consolidation, and performed analytical tasks, including developing an AI assistant for insights generation and query answering based on reviews.
- Conducted analysis and optimization on the name screening model to determine the optimal threshold for AML compliance.
- Performed ad-hoc analysis and generated actionable insights.
- Set up AWS infrastructure, including ECR, ECS, VPC, and IAM, to deploy prototypes for internal testing.

Naluri

Jun 2022 - Jun 2023

Data Scientist

Kuala Lumpur, Malaysia

- Developed a user segmentation model to cluster users based on activity patterns.
- Created recommendation models to provide coaches with optimal call-to-actions for different user segments.
- Conducted analysis on user health data to predict the probability of achieving improved health outcomes.
- Built a fitness AI prototype with computer vision capabilities to count exercise repetitions (e.g., squats, push-ups, sit-ups, planks).
- Developed and fine-tuned models, including a web scraper to gather food images for model enhancement.
- Oversaw AWS EC2 and ECS services for cronjobs and APIs.
- Managed and maintained ETL pipelines utilizing dbt, Prefect, and AWS Redshift.
- Designed and maintained interactive dashboards using Metabase.

The Center of Applied Data Science (CADS)

Mar 2020 - Jun 2022

Data Science Specialist

Kuala Lumpur, Malaysia

- Build pipeline to migrate data from several different databases into a clean integrated database to allow the ease of analysis
- Create regression models to predict probability of closing opportunities on Salesforce
- Build recommendation models for jobs, learning paths and skills to users, and deploy them as API using Flask
- Perform web scraping and text mining on job posting websites to get current skill trends for various jobs
- Work as Technical Assistant to provide assistance and consultation to clients on upskilling data literacy skills

Maxis	Jul 2017 - Aug 2017
<i>Business Intelligence Intern</i>	Kuala Lumpur, Malaysia
· Extracted and managed data in the database using Oracle SQL	
· Created reports based on customer requirements using Qlik Sense	
· Analysed and compared the back-end data with business data	

Education

Imperial College London	Sep 2018 - Sep 2019
<i>MSc Advanced Computing</i>	<i>Result: Distinction</i>
University of Manchester	Sep 2015 - Jun 2018
<i>BSc(Hons) Computer Science and Mathematics</i>	<i>Classification: First Class</i>
Taylor's College Subang Jaya	Jan 2014 - Jun 2015
<i>Cambridge International A Level</i>	<i>Chemistry (A *), Physics (A *), Mathematics (A *)</i>

Projects

Deep Image Reconstruction from Brain Activity	Apr 2019 - Sep 2019
· Functional magnetic resonance imaging (fMRI) and electroencephalogram (EEG) used to record people's brain activity of subjects viewing different images	
· Developed novel approach of using a dual generative adversarial network (DualGan) to generate images from recorded brain activity	
· Implemented the current state-of-the-art method of reversing a pre-trained convolutional neural network (CNN) to investigate results on own dataset	
Poetry Generation using Deep Learning	Sep 2017 - Apr 2018
· Program to generate poems using recurrent neural network and gated recurrent units	
· Distinct poems can be generated based on a user-supplied topic or image	
· Built with Python, in particular Theano and TensorFlow to implement the neural networks, and external APIs	
EventLite	Feb 2017 - May 2017
· Web application to create, list, search and manage events that are going on in Manchester	
· Has two main roles, event seekers for those looking for events and event managers for those who create events	
· Created by a team of six using Spring frameworks and external APIs	
Stendhal Game	Sep 2016 - Dec 2016
· Stendhal is a multi-player online adventure open source game	
· Tested, debugged, built, developed and deployed a multi-user, multi-threaded, client-server open source game	
· Automation of builds and tests done using Eclipse, Git, Apache Ant, Jenkins, JUnit and SonarQube	
Study Buddy	Feb 2016 - Apr 2016
· Web application to assist users in studying and help them be aware of procrastination	
· Users can input quizzes, which would pop-up periodically for them to complete	
· Created by a team of six using HTML, CSS, PHP, MySQL and JavaScript	

Awards and Achievements

Certificate of Excellence	Jun 2018
Awarded to the top 10% of graduating students in the School of Computer Science	
International Excellence Undergraduate Scholarship	Dec 2015
Awarded based on outstanding academic performance by the University of Manchester	

References available on request
