ONG JIUN JER

+6016-233-9860 | jiunjer0326@gmail.com | Selangor, Malaysia | GitHub | LinkedIn | Kaggle

As a Computer Science graduate with a First-Class degree and a specialization in Data Analytics, I have developed a strong grounding in programming (Python, SQL, R, C, and Java), data analysis, statistical analysis, data visualization, and building analytical models. I'm now ready to begin my professional journey in a dynamic, data-driven environment.

My growing interest in machine learning, deep learning, and artificial intelligence motivates me to apply my skills to solve meaningful real-world problems. I hope to contribute to impactful projects, support data-informed decision-making, and continue learning within an innovative, forward-thinking organization.

I am seeking a full-time role that values creativity, continuous learning, and the strategic use of data analytics.

WORK EXPERIENCE

1) Corporate Insight Analyst Intern | IWISERS

April 2024 – August 2024

- Cleaned and categorized large datasets using Excel and proprietary tools
- Developed monthly PowerPoint reports on digital marketing performance
- Used Pulsar and ExportComment to extract and analyze social media data
- Collaborated with colleagues to deliver social media performance insights
- Shared knowledge and handed over tasks to help the team continue work smoothly

EDUCATION

University: Asia Pacific University of Technology & Innovation (APU)

Course: Bachelor of Science (Hons) in Computer Science (Specialism in Data Analytics)

CGPA: 3.51

Graduated: September 2025

CERTIFICATE

1) IBM Data Analyst Professional Certificate (June 2024)
2) Google Data Analytics Professional Certificate (June 2022)
   a. Process Data from Dirty to Clean (April 2022)
   b. Analyze Data to Answer Questions (April 2022)
   c. Share Data Through the Art of Visualization (May 2022)
   d. Data Analysis with R Programming (May 2022)
   e. Google Data Analytics Capstone: Complete a Case Study (June 2022)
3) Introduction to Probability and Data with R by Duke University (July 2022)

PROJECT

1) Final Year Project: Game Review Classification Using NLP

Tools: Python, Google Colab, Streamlit, Gemini API.

Project Overview:

This project involved building and deploying a Natural Language Processing (NLP) pipeline to classify Steam game reviews into distinct topics. Clustering was first applied to generate topic labels, transforming the dataset for supervised learning. Subsequently, various BERT, Machine Learning (ML), and Deep Learning (DL) models were developed and evaluated. The final system was deployed using Streamlit and integrated with the Gemini API for enhanced interactive features.

- Conduct data understanding to analyze the dataset.
- The raw dataset was preprocessed to ensure quality, consistency, and proper formatting.
- Conducted unsupervised clustering using K-Means to assign topic labels for supervised model training.
- Built and trained four BERT model variants (Tiny, Mini, Small, Medium) for topic classification using the Hugging Face framework.
- Developed and evaluated four ML/DL models: CNN, LSTM, Naive Bayes, and Logistic Regression for comparison with BERT models.
- Implemented model evaluation techniques such as learning curve, classification report, and confusion matrix to review model performance.
- Deployed the best-performing model using Streamlit.
- Integrated the Gemini AI API to build an AI chatbot that can answer questions given by users and do interpreting chart images.
- Achieved high classification accuracy (up to 97%) with Logistic Regression.

2) Certificate Project: IBM Data Analyst Capstone Project

Tools: Python, Jupyter Notebook, Power BI.

Project Overview:

This capstone project was completed as part of the IBM Data Analyst Professional Certificate program. It focused on applying the entire data analytics workflow. From data collection and preparation to visualization and dashboard creation to derive insights from a real-world dataset.

- Collected data through API and web scraping.
- Performed data cleaning using Python in Jupyter Notebook.
- Conducted exploratory data analysis (EDA) to understand dataset patterns
- Created informative data visualizations.
- Developed a multi-tab Power BI dashboard to present insights interactively.

3) House Rent Dataset Analysis

Tools: R

Project Overview:

This project focused on analyzing rental housing trends across major cities in India using linear regression models. The goal was to identify key factors influencing rental prices and generate actionable insights for property owners, tenants, and analysts.

- Conducted exploratory data analysis (EDA) to uncover trends and anomalies in rental pricing across cities, including frequency analysis, descriptive statistics, correlation analysis, and hypothesis testing using t-tests.
- Performed data cleaning and preprocessing to prepare the dataset for analysis.
- Built and evaluated linear regression models, grouping rental prices by multiple influencing factors.
- Example Influencing factors included city, furnishing status, contact type, and carpet area.
- Calculated and interpreted R-squared and Adjusted R-squared values to assess model accuracy and generalization.

4) Predicting Player Engagement with Deep Learning

Tools: Python

Project Overview:

Developed a deep learning model using Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU) to classify player engagement levels based on online gaming behavior. Achieved 90% prediction accuracy after optimizing the CNN and GRU architectures.

- Explore the Predict Online Gaming Behavior dataset from Kaggle.
- Preprocessed data by cleaning, encoding, normalizing, and balancing classes.
- Split data into training and testing sets (80/20).
- Built two models: CNN (Conv1D + Dense) and GRU (GRU + Dense).
- Tuned hyperparameters using Keras Tuner (RandomSearch) for both models.
- Evaluated performance with learning curves, classification reports, and confusion matrices.

5) Sentiment Analysis by Using Naive Bayes

Tools: Python

Project Overview:

Built a sentiment classification model using Naive Bayes to analyze the Women's E-Commerce Clothing Reviews dataset, categorizing reviews into positive, negative, and neutral sentiments. Achieved overall accuracy of 76%.

- Conducted data exploration and understanding on the Women's E-Commerce Clothing Reviews dataset.
- Cleaned and preprocessed text data by removing irrelevant columns, null values, special characters, punctuation, and stop words; converted text to lowercase and created a sentiment target variable.
- Transformed text into numerical features using TF-IDF and Count Vectorization techniques.
- Handled class imbalance using random over sampling and random under sampling methods.
- Trained and validated a Naive Bayes model for sentiment classification.
- Performed hyperparameter tuning using Grid Search to optimize model performance.
- Evaluated results using classification report, confusion matrix, and cross-validation.

LANGUAGE PROFICIENCY & SKILLS

Language Proficiency	Skills
English: Good (read, write, speak) Mandarin: Native (read, write, speak) Malay: Good (read, write, speak)	• Power BI / Looker Studio • Excel • Power Point • Python • SQL • Java • C/C++ • R

ADDITIONAL INFORMATION

I contributed as a puzzle designer for escape room events, collaborating closely with my teammates to design and execute immersive challenges and to ensure successful and engaging experiences.