

# Alicia Chong Tsui Ying (She/Her)

[alicia.chong.data@gmail.com](mailto:alicia.chong.data@gmail.com) | [LinkedIn](#) | [GitHub](#) | +60122596163 | Kuala Lumpur, Malaysia

Recent graduate with a degree in Data Analytics and expertise in **Data Analysis** and **Machine Learning**. Seeking an entry-level **Data Scientist / Data Analyst** position, bringing strong analytical skills and a passion for extracting insights from data.

## EDUCATION

**Sunway University, School of Engineering and Technology**

**Petaling Jaya, Selangor.**

*Bachelor of Information Systems (Honors) (Data Analytics)*

*Graduation Date: Aug 2023*

- **GPA:** 3.73/4.00 (*First Class Honors*)
- **Capstone Project Focus:** Natural Language Processing (Text Classification Model)
- **Relevant Coursework:** Applied Statistics, Data Mining, Business Intelligence Systems, Database Management Systems, Big Data Analytics in the Cloud, Linear Algebra, Analytics Engineering.
- **Organizations/Awards:** Undergraduate Student Representative, Scholarship recipient for academic performance, Dual Award by Lancaster University, Public Relations Director of Sunway Career Peer Advisor (student leadership body).

## PROFESSIONAL EXPERIENCE

### CARSOME

**Petaling Jaya, Selangor.**

*Data Scientist*

*Sep 2023 - Nov 2023 (Laid-off)*

- Collaborated cross-functionally with marketing and customer service teams to devise targeted campaigns based on data-driven insights, leading to a significant increase in customer engagement metrics.
- Conducted extensive data analysis utilizing **SQL** to identify high-intention lost leads for retargeting by the customer service team, contributing to a 3% increase in customer engagement metrics.
- Revamped and implemented a **SQL** script that suggest alternative car options for leads lost due to unavailability of their car of interest. Achieved a satisfactory result of 5% leads-to-appointment conversion rate during a successful pilot test.
- Delivered timely insights on competitor data by revising **SQL** and **Python** scripts for data extraction and cleaning. Empowered business stakeholders to make informed decisions.

### airasia academy

**Kuala Lumpur, Malaysia**

*Data Analyst Intern*

*Dec 2022 - Mar 2023*

- Developed and maintained dashboards using **Google Looker Studio** to provide real-time insights by combining data from multiple sources including **MySQL** databases, **Google Analytics** and Facebook Ads for company stakeholders to access timely and accurate data for making data-driven decisions.
- Performed comprehensive data analysis on the company's website employing **web scraping** techniques and **Python** libraries (Selenium, Pandas, Matplotlib). Identified trends, patterns, and areas for improvement, contributing valuable insights to strategic decision-making.
- Supported the team with ad-hoc tasks involving the cleaning, merging, and formatting of tabular data utilizing **Python** and **Microsoft Excel**.

## PROJECTS

**Natural Language Processing** - Detection of DeepFake Text on Twitter with Semantic, Emoji, Sentiment and Linguistic Features (Published Journal with IEEE)

[Published Journal with IEEE](#) | [GitHub](#)

- Developed a **classifier** to distinguish between machine-generated text (MGT), also known as deepfake text, and human-written text (HWT) on Twitter.
- Fine-tuned a pre-trained transformer model, **BERT**, and integrated it with emoji and linguistic features to construct a comprehensive feature vector for robust classification.
- Achieved an impressive **accuracy rate** of **88.3%** and identified distinct characteristics of MGT, contributing to the field of machine-generated text detection on social media platforms.

- Enhanced the TweepFake dataset by including additional MGT from the latest text generative models such as GPT-3 and ChatGPT.

**Machine Learning** - Car Sourcing with Data-Driven Segmentation (*1 of the only 13 Winning Solutions for Microsoft Code; Without Barriers 2023 Hackathon*)

[Hackathon Submission](#) | [GitHub](#) | [Featured Newsletter on Sunway University Website](#)

- Achieved top honors as the winning solution among 7 submissions for a problem statement by Carsome, showcasing the effectiveness of the proposed solution.
- Developed a K-prototype cluster model that effectively segments cars based on shared car characteristics, delivering useful insights to Carsome for optimizing their car sourcing strategy.

## LEADERSHIP EXPERIENCE

**Sunway Career Peer Advisor**

*Public Relations Director*

**Petaling Jaya, Selangor.**

*Mar 2022 - Mar 2023*

- Led a dynamic team of 6 subordinates, fostering a collaborative and high-achieving work environment.
- Leveraged data analytics techniques, employing **Microsoft Excel** and **Python**, to automate the tracking of social media presence growth, ensuring efficient and data-driven decision-making.
- Crafted and presented monthly social media benchmark reports to organization members and advisors, providing actionable insights for informed decision-making regarding content creation strategies.
- Built and managed the organization's website using **Bootstrap** and hosted it on **GitHub Pages**, contributing to a modern online presence and increased brand awareness.
- Devised and executed a strategic social media content plan, resulting in a notable 43% increase (274 followers) in the organization's Instagram account within a span of 7 months.

## TECHNICAL SKILLS & TOOLS

**Skills:** Statistical Analysis, Data Mining, Machine Learning, Natural Language Processing, Database Design and Management Systems, Big Data Analysis, Data Storytelling, Cloud Infrastructure, Web Scraping, Website Design

**Programming Languages:** SQL (MySQL, RedShift, Oracle, BigQuery), Python, R, SAS (SAS Enterprise Miner, SAS Enterprise Guide, SAS Studio)

**Data Visualization Tools:** Python (Matplotlib, Seaborn), Tableau, Google Looker Studio, QlikSense

**Machine Learning Libraries:** Python (Scikit-learn, TensorFlow, PyTorch, Numpy, Pandas)

**Spreadsheet Software:** Microsoft Excel, Google Sheet (Advanced formulas, Pivot Table)

**Other Tools:** GitHub, AWS (EC2, S3), Apache Hadoop (MapReduce, Spark, Hive), Bootstrap Studio, Adobe Photoshop, Adobe Premier Pro

## LICENSES & CERTIFICATES

- Cloud Infrastructure and Architect Complete Reskill Program - airasia academy (Issued 09/2022)
- Digital Transformation Program - Alibaba Business School Global Digital Talent (GDT) (Issued 04/2022)
- A1900 Azure AI Fundamentals - Microsoft (Issued 08/2021)
- DP900 Azure Data Fundamentals - Microsoft (Issued 08/2021)
- The Data Scientist's Toolbox – Coursera (Issued 01/2021)
- International English Language Testing System (IELTS) (Band: 7.5) (Issued 11/2019)

## ADDITIONAL INFORMATION

**Languages:** English (*Proficient*), Chinese (*Proficient*), Malay (*Proficient*), Cantonese (*Proficient*)

**Interests:** Boulderling, Photography, Video Editing, Arts and Crafts, Playing Saxophone & Piano