Tan Shih Jen

+65 8145 4991

e1094875@u.nus.edu

www.linkedin.com/in/tan-shih-jen-99546b306

https://github.com/shihjen

Summary

With 8 years of hands-on research lab experience, I excel in managing lab operations and research funds. Passionate about learning new technologies, I apply them in both work and daily life. Recognizing the rise in data creation and the importance of data-driven decisions, I've enhanced my data literacy through analytics coursework at NUS and am pursuing a part-time Master's in Biomedical Informatics to boost my healthcare analytics skills. I understand the value of extracting insights from medical data to improve patient care and operational efficiency. With a strong foundation in analytical methods and skills in machine learning, data visualization, and natural language processing, I am ready to drive innovation in healthcare.

Education

National University of Singapore

Master of Science - Biomedical Informatics (Analytic Specialisation)

August 2022 - July 2024

University Putra Malaysia

Bachelor of Science (Honors) - Biochemistry

September 2011 - June 2015

Experience

Laboratory Technologist (Jan 2019 - Present)

Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore

Singapore

- Led undergraduate practical classes for LSM3232 Microbiology (Practical 1-3) and LSM3225 Molecular Microbiology in Human Diseases (Practical 4), providing technical and logistics support to faculty, instructors, and teaching assistants.
- Supported the Assistant Professor Chris Sham Lok-To in managing research grants by overseeing budget allocation, tracking expenditures, and preparing financial reports. Implemented Python programming to automate the tracking and management of oligomer orders and sequencing service usage in the research lab.
- Served as the safety lead in Assistant Professor Chris Sham Lok-To's laboratory. Conducted laboratory safety induction training for new students and staff before they commenced their projects.
- Prepared, reviewed, and maintained Standard Operating Procedures (SOP) and Risk Assessments (RA) for the PI research lab and department core facilities.
- Implemented and monitored laboratory safety practices, ensuring compliance with University and national regulations.

Research Assistant (Sep 2015 - Jan 2018)

Laboratory of Natural Products, Institute of Bioscience, University Putra Malaysia

Selangor, Malaysia

- Conducted metabolomics study on natural products.
- Wrote research proposal and assisted in the grant application.
- Ensured proper maintenance and management of the husbandry and animal facility.
- Managed procurement tasks, including sourcing and purchasing laboratory supplies.

QA/AC Intern (Jun 2014 - Aug 2014)

Rida Fruits Sdn Bhd

Batu Pahat, Johor, Malaysia

- Performed QC inspection and routine laboratory tests of incoming raw materials, intermediate products, packaging, and finished goods with established sampling and testing protocols.
- Prepared QA and QC reports.
- Performed sensory evaluation on raw materials and finished products.
- Executed food safety standards in compliance with HACCP and HALAL requirements.

Skills

- Programming Language: Python, R, SQL
- Machine Learning: Scikit-Learn, RapdMiner, SIMCA
- Data Visualization: Matplotlib, Seaborn, Plotly, Tableau
- Data Manipulation & Analysis: Pandas, MS Excel
- Natural Language Processing: spaCy, Gensim, NLTK
- Web Development: HTML, CSS, Streamlit

Selected Projects

Early Detection of Non-alcoholic Fatty Liver Disease (NAFLD)-Associated Hepatocellular Carcinoma (HCC) by Machine Learning
Description: Capstone Project of MSc Biomedical Informatics. Utilized the NUHS Discovery AI platform for a cancer prediction project focused on early detection of NAFLD-associated hepatocellular carcinoma. Queried multi-domain medical databases using SQL. Conducted data wrangling and preprocessing with pandas, addressing imbalanced data, imputation, encoding, and scaling. Developed multiple machine learning models, including logistic regression, decision tree, random forest, SVM, and adaptive boosting. Evaluated model performance using F1 score, precision, recall, and AUROC, with the SVM model achieving the highest scores across all metrics.

Kidney Failure Risk Prediction in Intensive Care Unit (ICU) Patients
Description: This project involved developing a strategy to predict the risk of kidney failure in ICU patients using real-world EMR data from the MIMIC-IV database. The process included thorough data inspection to address duplicate records and missing values, followed by preprocessing steps such as encoding, scaling, and imputation. Predictive models were trained using various machine learning algorithms, focusing on tree-based methods and boosting algorithms. To achieve optimal model performance, several feature selection techniques were applied, including filter methods, embedded methods, and wrapper methods such as recursive feature elimination, stepwise feature selection, and genetic algorithms. Eleven models were trained using different feature subsets and learning algorithms. The goal was to construct a model capable of accurately estimating the likelihood of ICU patients developing acute kidney injury (AKI), with a preference for models exhibiting high average precision. Notably, all models demonstrated comparable AUC (ranging from 0.694 to 0.739) and average precision scores (ranging from 0.801 to 0.847).

Automated Tracking System for Oligomer and Sequencing Services
Description: Developed a web application in Python to automate tracking and recording of oligomer and sequencing service usage in a research laboratory. The application extracts unstructured data from invoices and uploads it to Google Sheets for record-keeping. It features a usage dashboard displaying expense breakdowns and details such as purchase order numbers, sales order numbers, invoice numbers, and delivery orders. By automating data collection through invoice uploads, the system eliminates manual data entry, reducing repetitive tasks and minimizing errors.

BCEAD Research Laboratory Website (Dr. Chris Sham's Laboratory)
Description: Developed a static website using HTML and CSS to inform the public about the lab research activities. The site includes comprehensive details on research scope, publications, lab members, and events, ensuring timely updates and accessibility of the lab's latest work.

Certification & Professional Development

- CITI Health Information Privacy and Security (HIPS) | CITI Program
- AICP Python for Artificial Intelligence and Machine Learning | National University of Singapore
- Data Literacy Programme (DLP) - Advanced: Survey Analytics | National University of Singapore
- Artificial Intelligence Competency Course (Intermediate) | National University of Singapore
- Data Literacy Programme (DLP) - Intermediate: Applied Regression Models using R | National University of Singapore
- Data Literacy Programme (DLP) - Intermediate: Customer Analytics | National University of Singapore
- Professional Certificate in Basic Artificial Intelligence and Machine Learning | National University of Singapore
- Data Visualization Begins With Me | National University of Singapore
- Database Creation, Manipulation, and Querying with SQL | National University of Singapore
- Artificial Intelligence Competency Course (Foundation) | National University of Singapore
- Python Programming | National University of Singapore
- Data Literacy Programme (DLP) - Basic | National University of Singapore

Publications

- Chong SG, Ismail IS, Ahmad Azam A, Tan SJ, Shaari K, Tan JK. Nuclear magnetic resonance spectroscopy and liquid chromatography-mass spectrometry metabolomics studies on non-organic soybeans versus organic soybeans (*Glycine max*), and their fermentation by *Rhizopus oligosporus*. J Sci Food Agric. 2023;103(6):3146-3156. doi:10.1002/jsfa.12355
- Tan SJ, Ismail IS. Potency of Selected Berries, Grapes, and Citrus Fruit as Neuroprotective Agents. Evid Based Complement Alternat Med. 2020;2020:3582947. Published 2020 May 30. doi:10.1155/2020/3582947
- Marikkar, Nazrim & Tan, S.J. & Salleh, A. & Azlan, Azrina & Shukri, M.A.M.. (2016). Evaluation of banana (*Musa* sp.) flowers of selected varieties for their antioxidative and anti-hyperglycemic potentials. 23. 1988-1995.