

Contact

www.linkedin.com/in/tanpengshi
(LinkedIn)
medium.com/@tanpengshi (Blog)
github.com/tanpengshi (Portfolio)

Top Skills

Python
Data Visualization
Public Speaking

Languages

German (Elementary)
English (Native or Bilingual)
Mandarin (Native or Bilingual)

Certifications

Python for Everybody -
Specialization
Machine Learning
The Complete SQL Bootcamp 2020:
Go from Zero to Hero

Honors-Awards

DSTA Undergraduate Scholarship

Tan Pengshi Alvin

Data Scientist at Miracom Pte Ltd
Singapore

Summary

Passionate and curious Data Scientist, with a flair for programming and data story-telling, and a repertoire of skills in exploratory data analysis, statistical analysis, machine learning algorithms, NLP, recommender systems, computer vision and neural networks. Conversant in SQL, Tableau and Python packages such as Pandas, Matplotlib, Seaborn, Scikit-Learn, Keras and OpenCV, I am highly results driven and I look forward to interesting data science projects, in which I can deliver innovative solutions, business value and impact for the company. In the process, I plan to work towards acquiring the skills of a Full-Stack Data Scientist.

In my available time, I also like to pick up additional data science knowledge through navigating Medium, textbooks and learning from Coursera, Udemy. Other hobbies include reading non-fiction, meditation, gaming, hitting the gym, Table-tennis and going for long walks in the neighborhood.

Data Science Blog: <https://medium.com/@tanpengshi>
GitHub: <https://github.com/tanpengshi>

Experience

Miracom Pte Ltd
Data Scientist
December 2020 - Present (5 months)
Singapore

Lead data science projects at a startup in the area of Computer Vision, in particular Facial Recognition and Car Number Plate Recognition. Achievements include:

- Facial Recognition: Develop a highly accurate real-time face recognition (with/without mask) neural network model with IP cameras using OpenCV, ArcFace or Google's FaceNet for embeddings and MTCNN/DNN/DSFD for

face detection. The model is engineered to train on the fly with existing targets autonomously, and with new targets manually. The authentication model is also trained against spoofing attacks.

- Car Number Plate Recognition: Model is currently undergoing development.
- Data Engineering: Set up protocol buffers and APIs for client-server gRPC connection.

Metis

Data Scientist

September 2020 - November 2020 (3 months)

Singapore

Trained under an accredited 12-week immersive apprenticeship in full-stack data science, covering topics such as communication and design, programming, data acquisition, exploratory data analysis, statistical inference, supervised and unsupervised learning, time series analysis, Natural Language Processing, Recommender Systems, Neural Networks, SQL and noSQL databases and Spark. I worked on a total of 5 data science projects as follows:

- Project 1: Exploratory data analysis of traffic flow trends with NYC MTA turnstile data
 - Techniques Used: Highly customized Heatmaps, Histograms, Barplots and Boxplots
 - GitHub: https://github.com/tanpengshi/Metis_Project_1_MTA_Analysis
- Project 2: Predicting Market Value of FIFA Soccer Players with Regression
 - Techniques Used: Regularization, CV and LASSO, Ridge, ElasticNet, Polynomial Regression
 - GitHub: https://github.com/tanpengshi/Metis_Project_2_FIFA_Players
- Project 3: Predicting Satisfaction of Airline Passengers with Classification
 - Techniques Used: kNN, Logistic Regression, Naive Bayes, SVM, Decision Trees, Random Forest
 - GitHub: https://github.com/tanpengshi/Metis_Project_3_Airline_Satisfaction
 - Flask App: <https://flight-satisfaction-prediction.herokuapp.com/>
- Project 4: Strategic Analysis of Trump Rallies with NLP and Time Series
 - Techniques Used: Non-negative Matrix Factorization, K-Means, SARIMA Time Series Prediction
 - GitHub: https://github.com/tanpengshi/Metis_Project_4_Trump_Rallies

- Project 5: A Content-Based Recommender System for E-Commerce Webstore
- Techniques Used: Cosine Similarity, Matrix Product, TfidfVectorizer, MinMaxScaler
- GitHub: https://github.com/tanpengshi/Metis_Project_5_Ecommerce

National University of Singapore

Research Engineer

March 2019 - August 2020 (1 year 6 months)

Singapore

Lead Engineer on the government-funded defence project, involving optimization of physics-based and numerical human thermoregulatory models using MATLAB to predict the evolution of SAF soldiers' core temperatures during route marches, in anticipation of potential heat injuries during military exercises. Other achievements include:

- Enhanced model prediction by identifying and innovating new scaling parameters, based on physiological principles through literature research and data-driven insights.
- Simulated 5-Fold Cross Validations across multiple datasets based on Root Mean Square Error (RMSE) and Standard Deviation fall-in Percentage (SDP) to test robustness across models.
- Collaborated with professors and engineers to integrate the prototype model into the final product with user interface.

Self-Employed

JC Specialist Tutor

June 2015 - March 2019 (3 years 10 months)

Singapore

Ventured into building a tutoring business by lecturing Math/Statistics and Physics to a range of students from secondary to junior colleges, in both classroom and one-to-one settings. Other achievements include:

- Created a residential classroom studio with a dedicated website that attract student enrollment through online advertisements
- Designed and conducted various Physics experiments to both O'Levels and A'Levels students.
- Participated in curriculum development in Physics and Math teaching/ learning materials.

STATS ChipPAC Ltd., Wafer Process Department

Lead Planner

January 2015 - June 2015 (6 months)

- Supervised real-time operations in the production line, planning the flow of production units in a complex manufacturing environment to meet tight deadlines.
- Built from scratch a project by drafting and simulating a new productivity-tracking system for manufacturing operators, using system data and Microsoft Excel formulae.

National Library Board, Singapore Memory Project

Assistant Associate

August 2014 - December 2014 (5 months)

Singapore

- Sourced for memory contributors by conducting phone and face-to-face interviews, before transcribing them into an article on the Singapore Memory portal and iremember.sg blog.
- Managed and supervised other administrative executives on data entry and uploading of memory contributions

Education

Nanyang Polytechnic

Specialist Diploma, Applied Artificial Intelligence · (2021 - 2022)

National University of Singapore

Bachelor of Science (Merit), Physics · (2010 - 2014)

Raffles Junior College

GCE 'A' Levels · (2005 - 2006)

Victoria School (Singapore)

GCE 'O' Levels · (2001 - 2004)