Nicholas WAN

DATA SCIENTIST

nicholaswan@live.com.sg · http://www.nicholaswan.me/ · https://github.com/nicholas-wan

Education

National University of Singapore *Singapore*

BS.C BUSINESS ANALYTICS (HONS), HIGHEST DISTINCTION *Aug. 2017 – Dec 2020 · 3.5 yrs*

- NUS Merit Scholarship
- Cumulative Average Point: 4.66/ 5.00
- NUS Overseas Colleges (Silicon Valley)

Stanford University

NON GRADUATING OPTION *Aug. 2019 – Apr 2020 · 9 mos*

- Mining Massive Datasets – Grade Obtained: A
  Topics Include: Frequent itemsets and Association rules, Near Neighbor Search in High Dimensional Data, Locality Sensitive Hashing (LSH), Dimensionality reduction, Recommendation Systems, Clustering, Link Analysis, Large-scale Supervised Machine Learning, Data streams, Mining the Web for Structured Data, Web Advertising.

Skills

Programming Python · R · Java · HTML/CSS · JavaScript · Spark
Database SQL
Visualization Tableau · Sisense
Cloud Amazon Web Services · Solutions Architect Associate

Experience

Defence Science and Technology Agency *Singapore*

ENGINEER (CYBER AI) *Jan 2021 – Present· 1 yr 7 mos*

- AI and Machine Learning solutions to defend against Cybersecurity threats.
- Developed models to detect Phishing based on URLs, logos, and Website Screenshots, which achieved a validation F1 score of 0.83 on a public dataset from PhishTank. Automated email alerts in the forms of PDF reports to be sent to stakeholders.
- Created a Tech Showcase (Interactive Website) on the topic of <Adversarial AI Attacks on Smart Home devices> for BrainHack 2021, an annual DSTA organised event. The showcase was voted by participants as the top showcase out of 27 showcases that were created by other departments in DSTA.

Shopee *Singapore*

DATA SCIENCE INTERN *Apr 2020 – Nov 2020 · 8 mos*

- Trained NLP models to perform multiclass classification on product titles and descriptions to predict product attributes.
- Implemented NLP in multiple languages including English, Chinese, Thai, Indonesian and Vietnamese.

The League *San Francisco*

DATA SCIENCE INTERN *Aug 2019 – Apr 2020· 9 mos*

- Created and productionized a Machine Learning model to predict catfishers and bad users leading to the customer support team saving 30% of time reviewing profiles and improving the trust and safety in the community.
- Set up an alert system for over 20 key company metrics such as Revenue, DAUs, Churn, using Periscope / Sisense to notify stakeholders through Slack.
- Performed Statistical modelling through regression models such as Diff in Diff, IV Regression, Fixed Effect models to advise product decisions.
- Conducted AB Testing experiments with power analysis and statistical significance tests to inform and influence product decisions, which included several product features.
- Experience with creating, scheduling, monitoring and debugging jobs with Airflow. .

Project

StreetCredit bit.ly/streetcredit

- Conceptualized and prototyped a solution that allows crowdfunded micro-loans for students.
- Won 1st Place in Financial Inclusion Hackathon organized in San Francisco.

Hack4Climate bit.ly/blizzard-yale

- Created a website to show how stakeholders are affecting the environment, using Data Visualisations with Plotly.
- Won 1st place in YaleNUS Hackathon in the Data Science Category, Hack4Climate.