Samson Qian

La Jolla, CA 92093 | (408) 656-4984 | saqian@ucsd.edu | linkedin.com/in/samsonq | kaggle.com/samsonqian

SKILLS

Technical Skills:

- Python *(Pandas, Matplotlib/Seaborn, NumPy, SciPy)*
- TensorFlow/Keras/sklearn
- AWS SageMaker
- R *(RStudio)*
- Hadoop
- Java
- Tableau
- SQL *(MySQL)*
- MongoDB
- Spark

Languages: English, Chinese (Mandarin)

PROFESSIONAL EXPERIENCE

Business Data Analyst — **Sunnyvale, CA**
TribioScience Inc. — 06/2018 – Present

- Created company database in MongoDB and wrote SQL queries to retrieve company sales data
- Generated Tableau visualizations to observe trends of reagent transactions for different customers
- Designed regression models to predict number of reagents that customer will purchase
- Used A/B testing to evaluate statistical hypotheses and determine best sales bundles for customers
- Reported insights learned from data to marketing team and offered business suggestions

Marketing Manager — **Sunnyvale, CA**
TribioScience Inc. — 06/2016 – 05/2018

- Produced/distributed flyers of reagents to promote products and expand network of customers
- Managed marketing team of 8 and divided tasks and formulated weekly goals
- Contacted potential customers and negotiated product bundle deals
- Outsourced sales from other biotech companies and sold products through online market
- Completed transactions of over 2000 products in August 2017

EDUCATION

University of California, San Diego — **La Jolla, CA**
Bachelor of Science in Data Science; Minor in Business — Expected 2021

- **GPA:** 3.8/4.0
- **Honors:** Revelle College Honors, Provost Honors
- **Leadership/Activities:** Undergraduate Economics Society – *VP of Technical Committee,*
  Data Science Student Society (DS3) – *Marketing Committee*, Undergraduate Investment Society – *Officer*
- **Relevant Coursework:** Concepts of Machine Learning, Deep Learning, Data Structures & Algorithms, Data Visualization

PROJECTS

Store Traffic/Sales Time-Series Forecasting — 05/2019
Percolata

- Cleaned and reformatted customer traffic and sales transactions data for pharmaceutical stores; contains information about number of customers currently in store and number of purchases made in 15-minute time intervals for a span of 3 years
- Implemented stateful LSTM model for multistep time-series forecasting of store transactions and traffic
- Forecasted values one month into the future with customer traffic predictions RMSE of 0.9 & sales predictions RMSE of 10.5
- Deployed model with AWS SageMaker for efficient model training and a scalable Forecasting As A Service framework
- Used online learning to update model with real-time store data and learn new trends for more accurate forecasts

PocketDoc — 10/2018
UC Health Hackathon

- Created database to manage hospital patient information including alcohol consumption, age, gender, geographical location, and medical conditions like diabetes or injures
- Wrote Pandas algorithms to impute missing data, handle extraneous values, and set up training/testing datasets for modelling
- Designed classification model to predict whether a patient will show up to their scheduled appointment or not, based on medical features of patients; achieved 85% accuracy on testing set
- Combined model with mobile app for patients and doctors to use for improved patient experience
- Generated Tableau visualizations to display associations of medical conditions and no-shows; discovered trends between age, alcohol consumption, as well as severity of condition with appointment no-shows