Tenzin Choezin

tenzin.p.choezin@gmail.com
510-981-9734
linkedin.com/in/tenzinchoezin/
Berkeley, CA

WORK EXPERIENCE

IBM (Extreme Blue Program) May 2022 – Aug. 2022
Data Science Intern (Watson AIOps); San Jose, CA
- Prototyped an extension **toolkit** that enables client data scientists to **enhance** Watson AIOps' existing capabilities using their company's own additive domain- and business-specific data, AI/ML models, and insights
- Developed a Python **API client library** to cut live/historical change ticket and metrics data extraction time from **Elasticsearch** and **Kafka** by **3x**
- Created Jupyter Notebook workflows that demonstrate how to **enrich** AIOps data with financial data to minimize business costs by up to **50%**
- Built a live *AIOps x Business Analytics* **dashboard** with Plotly Dash and used Red Hat OpenShift to **containerize** and deploy the web application

Pacific Gas and Electric Company (PG&E) Jun. 2021 – May 2022
Data Engineer Intern (PSPS); San Francisco, CA
- Designed automated data **pipelines** and **querying** processes to decrease the turnaround time needed to procure Public Safety Power Shutoff (PSPS) event data for **800,000+** customers and **30+** counties by **65%**
- Built a PSPS **metrics dashboard** that visualizes updated event data and core business KPIs, saving **24 hours per month** of manual data fetching
- Engineered **data access** solutions with **PySpark** and **Spark SQL** to reduce PMO reporting and data request completion times by **70%** in my division
- **Spearheaded** the development of **KPIs** to drive actionable insights needed to improve PG&E's future customer experience with PSPS events

UC Berkeley Department of Statistics Jan. 2021 – May 2021
Teaching Assistant; Berkeley, CA
- Tutored students for a **330-student** course on probability and statistics
- Helped develop and graded assignments; hosted weekly office hours
- Prepared topical exam-review sessions on distributions, expectation, etc.

UC Berkeley Division of Data Science Aug. 2020 – Dec. 2020
Teaching Assistant; Berkeley, CA
- Created **curriculum** for an economic models and data science course
- Held office hours to assist students with **programming** assignments

DATA SCIENCE PROJECTS

Heart Failure Predictor Jan. 2022 – May 2022
Course: Data, Inference, and Decisions (Machine Learning)
- Trained and built binary **classification** models on a dataset of over **310,000** individuals to best predict people that have or don't have heart disease
- Implemented **logistic regression**, **random forest**, **decision tree**, and **KNN** models to achieve a **91%** accuracy on a test set of **80,000+** data points

Stock Sentiment Tracker Jan. 2022 – May 2022
Course: Data Science with Venture Applications
- Developed a tool that provides holistic financial and sentiment data on popular stocks via comprehensive leaderboard, interactive charts, and automated email alerts, achieving a **95%** approval rating across **50**+ users
- Leveraged the Twitter, Reddit, Yahoo Finance, and Hugging Face **APIs**, Anvil for UI development, and Amazon **EC2** for **data storage**

EDUCATION

University of California, Berkeley
Majors: Data Science, Economics
Overall GPA: 3.58 / 4.00
Graduation: August 2022
Relevant Coursework: Principles and Techniques of Data Science, Data Mining and Analytics, Data Engineering, Data Science with Venture Applications, Data Structures, Machine Learning, Probability Theory, Mathematical Statistics, Linear Algebra, Econometrics, Macro/Micro-economics, Product Development
Certifications: SCET Certificate in Entrepreneurship and Technology

SKILLS

Programming Languages: Python (Pandas, NumPy, SciPy, Plotly, Dash, SciKit-Learn, PySpark, Seaborn, Matplotlib, Beautiful Soup, Statsmodels), SQL (MySQL, Spark SQL, PostgreSQL), Java, R, Scheme
Skills: Exploratory Data Analysis, Machine Learning, Data Visualization, Hypothesis Testing, Data Wrangling, Data Pipelines, Feature Engineering, Data Modeling
Tools: Apache Spark, Jupyter Notebook, Git, Palantir Foundry, Red Hat OpenShift, Kafka, Elasticsearch, MS Excel, Tableau

ACTIVITIES

California Actuarial League
Actuarial Data Analyst Fall 2020
- **Led** a team of 4 to process, explore, and analyze catastrophic claims and health insurance datasets in **R** and **MS Excel**
- Created **linear regression models** to quantify risk exposure and develop estimates for suitable rate relativities

Data Science Society
Data Analyst Spring 2020
- Used **Pandas** to analyze **1000s** of Kickstarter projects and determine the key factors behind successful companies
- Synthesized main points through **visualizations** using **Seaborn** and presented overall findings at a research symposium attended by **150+** people