Max Boonjindasup

Los Angeles | 818-428-0901 | mboon1228@gmail.com | Portfolio | Kaggle | Github | LinkedIn

Skills

Languages: Python, SQL, R, Excel, Matlab

Statistics: Hypothesis & A/B Tests, Regression, Classification

Tools: AWS, BigQuery, Databricks, Docker, Git, Langchain, Power BI, PostgreSQL, PySpark, Spark SQL, Spotfire, Tableau

ML: Deep Learning, Gen AI, k-means, Neural Networks, PCA, RAG Pipelines, Random Forest, XGBoost, Clustering

Relevant Experience

Data Scientist, *Upwork*   02/2023 – present | Remote

- **Developed a *RAG*-based system on *AWS* using *Bedrock* and *Sagemaker* that automated document and image parsing and enabled natural language question answering,** reducing query response time by 40%
- Built a data pipeline utilizing *AWS Lambda, Textract, S3, and Bedrock* (*Cohere Embed V3*) to store and retrieve vector embeddings in *DocumentDB*, supporting semantic NLP responses with *Bedrock* (*Claude V3 LLM*)
- **Applied anomaly detection techniques** (*graph analysis, k-means, DBSCAN, & hierarchical clustering*) and *PySpark* on *AWS Sagemaker, NumPy, & SciPy* **to identify fraudulent overcharging in 16,000 vehicle service records,** visualizing results with *Matplotlib, Streamlit, & Tableau* and achieving 30% cost savings
- Engineered end-to-end pipelines with Apache Spark and SQL in Databricks on AWS to streamline data acquisition, clean corrupt data, and accelerate data delivery by 80% for faster business-critical insights

Data Science Consultant, *Catenactio*   05/2024 – 09/2024 | Los Angeles, CA

- **Developed an HR chatbot powered by *Generative AI* (*Gemini API*) on *Google Cloud Platform*,** reducing repetitive inquiries by 40% and cutting HR operational costs by 25% through optimized synthetic training data
- **Architected automated data pipelines using *Python* and *SQL*,** enabling real-time cybersecurity threat detection and a 30% reduction in response times
- Delivered high-impact analytics solutions by collaborating with cross-functional stakeholders, driving a 20% increase in operational efficiency and accelerating business growth

Operations Manager, *Beverly Hills Arthritis Associates*   11/2021 – 11/2022 | Beverly Hills, CA

- Spearheaded the development of an **automated data cleaning pipeline using *Excel* and *Python* scripts, boosting efficiency by 60% and uncovering $100,000** in profit from unpaid accounts within 6 months
- Implemented a new Epic EHR data documentation system, reducing errors by 15% and ensuring accuracy in over 1,000 patient records, potentially improving patient care through enhanced data quality
- Directed a team of analysts to process insurance claims, ensuring efficient management of over 120 complex med-legal cases

Research Associate II, *Cedars-Sinai* (*Smidt Heart Institute*) 03/2021 – 10/2021 | Los Angeles, CA

- **Improved data collection and analysis with 3 custom *Excel* templates, reducing data processing time by 80%**
- **Applied advanced statistical techniques in *Python*, including correlation analysis and regression modeling**, to quantify gene-protein associations, contributing to a 30% improvement in experimental reproducibility and validation

Associate Scientist, *Amgen*   07/2019 – 01/2021 | Thousand Oaks, CA

- **Developed an end-to-end system that automated Amgen's workflow by integrating *Biovia ELN* for assay templating and generating *Spotfire* reports, reducing analysis time and resource use by 50%, which earned VP recognition and a promotion to project lead**
- Authored technical reports documenting methodologies and findings, supporting regulatory submissions and internal reviews
- Oversaw the development of 9 drugs from discovery to FDA approval and eventual market release

Projects

Machine Learning Approaches to Mental Health Diagnosis 🔗

- **Visualized mental health trends in 300,000 patient records using *Tableau* and *Python*,** revealing key intervention areas, including that 54% of struggling students are US males
- Built a 98% accurate *XGBoost* model that classifies susceptibility to mental health difficulties as low, medium, and high

Predicting Employee Retention 🔗

- **Engineered an employee churn model** (*XGBoost - 98% accuracy & precision*) using *Python* that identified 5 key factors in employee turnover and **projected a 20% improvement in resource management and employee satisfaction**
- **Analyzed a 5-year employee dataset using *Pandas, Seaborn, & Tableau*** to uncover insights and identify a strong correlation between assigning > 4 projects and a 200% surge in employee attrition

Certificates

- Google Advanced Data Analytics
- IBM Data Science
- AI/ML in Precision Medicine (Stanford Med)
- Data Science in Healthcare

Education

Bachelor of Science, *University of California San Diego*   09/2014 – 06/2018 | La Jolla, CA

Majors: Cognitive Science (Machine Learning & Neural Computation) | Biochemistry & Cell Biology