





Dr. Yu-Ting Shen

 yu-ting-shen-6b730b160  ytshen2020@gmail.com
 ytatus94  (405).200.2633

SKILLS

- **Programming:** Python, SQL, C/C++, Spark, Scala, Bash shell script, VBA,
- **Machine learning, Deep learning, Reinforcement learning:** Scikit-learn, Keras, TensorFlow, PyTorch, Gym, Stable-Baselines, Ray,
- **Analysis & Modeling:** ETL, EDA, Predictive modeling, Time series forecast, Anomaly detection, Monte Carlo simulation, RandomForest, XGBoost, LightGBM, LSTM, clustering,
- **Database:** SQL, Big Query, PostgreSQL, Incorta,
- **visualization:** Matplotlib, Seaborn, Dash, Bokeh, Google DataStudio, Tableau, Power BI,
- **Cloud:** Google Cloud Platform, Microsoft Azure, Amazon AWS, Cloudera,
- **Others:** Git, Docker, Anaconda, Jupyter, Databrick, Visual Studio Code, IntelliJ, Jira,
- **Soft skills:** Collaboration, Communication, Problem-solving, Leadership,

EXPERIENCE

Senior Data Scientist **Seeloz Inc, San Jose, CA** **2019/04 - present**

- Created data sanity, analysis, and ETL pipelines using **Python, SQL**, and **PySpark** in **Jupyter** and **Databrick**.
- Created data visualization and reports using **Matplotlib, Seaborn, Bokeh**, and **Plotly**.
- Created dashboard using **Tableau, Power BI, Google DataStudio** for stakeholder to help decision making.
- Built predictive models for demand forecast using **Scikit-Learn, H2O, XGBoost, LightGBM**, and **TensorFlow**.
- Built time series forecasting models using **ETS, ARIMA, Exponential-Smoothing** and **Prophet**.
- Built **anomaly detection** models to detect anomaly in orders and deliveries in supply chain.
- Implemented **Reinforcement Learning** model to optimize supply chain management and increase 44% annual turnover rate.
- Implemented **EOQ, TPOP**, and **ROP** supply chain models to build resource planing recommendations.
- Developed Python API offering cross-platform capability for accessing storage blobs on **GCP, Azure, S3**.
- Developed Shell script to submit training jobs to **Google AI platform, Azure VMSS** and **on-premises** clusters.
- Led junior data scientists on the team and facilitated resolution of project-related challenges.

Data Scientist **CERN, Geneva, Switzerland** **2015/03 - 2018/03**

- Improved the electron isolation efficiency from 83% to 99% (a 19% increase) by restricting the transverse energy and momentum distributions within a topological cone of 0.2 in spherical coordinate. The outcome set a new benchmark for all analysis at CERN.
- Designed, optimized, and implemented a high-performing classification model for real leptons across multiple energy scales, leveraging both statistical and machine learning methodologies. This resulted in a significant improvement in the model's recall, which increased from 62% to 98%.
- Conducted a comprehensive analysis of an extensive 400 TB dataset and employed decision tree, multi-dimensional regression, and statistical models, to deliver sophisticated solutions that effectively addressed complex project requirements.

Research Scientist **Academia Sinica, Taipei, Taiwan** **2009/07 - 2011/07**

- Created a Monte Carlo simulation model using C++ and increased 20% precision.

R&D Engineer **TSMC, Hsinchu, Taiwan** **2006/12 - 2009/02**

- Performed rigorous statistical analysis to develop models for advanced IC devices with cutting-edge technology.

EDUCATION

Ph.D. in Physics **University of Oklahoma, Norman, OK** **2011 - 2018**