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

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

- Analyzed clients' supply chain raw data using Python and SQL to create data ETL using PySpark.
- Developed the data schema and table architecture for SCAS metadata model using the raw data from Oracle ERP.
- Implemented a reinforcement learning model to optimize supply chain management and inventory control. Lowered annual inventory by 26% and increased annual turnover rate by 44%.
- Built time series forcasting model using ETS, ARIMA and Random Forest and achieved model accuracy from 83% to 97% across products.
- Created interactive dashboard using Python Dash and deployed on Azure using Docker. Experienced and proficient in Tableau, Power BI, Google DataStudio to visualize and analyze data.
- Developed a Python API offering cross-platform capabilities for accessing storage blobs on GCP, Azure, and AWS.
- Developed a Sell script for submitting model training jobs to Google AI platform, Azure VMSS, and on-premises cluster.
- Led Junior data scientists on the team and facilitated resolution of project-related challenges

CERN (Organisation Européenne pour la Recherche Nucléaire) **Geneva, Switzerland**
Data Scientist *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. Employed various modeling techniques, such as decision tree, regression, and statistical models, to deliver sophisticated solutions that effectively addressed complex project requirements.

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

- Created a Monte Carlo simulation model using C++ and GEANT4 that resulted in a 20% increase in precision.

TSMC (Taiwan Semiconductor Manufacturing Company) **Hsinchu, Taiwan**
R&D Engineer *2006/12 - 2009/02*

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

EDUCATION

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