YUNPENG WANG

◇ 302-509-7658 ◇ yw3954@columbia.edu ◇ https://www.linkedin.com/in/wyunpeng0302

EDUCATION

Columbia University *New York, US*
Master of Operations Research; Current GPA: 3.93 *Dec 2022*

· **Coursework: Optimization; Stochastic Models; Simulation; Deep Learning; Forecasting: A real-world application**

University of Delaware *Newark, DE*
Bachelor of Applied Mathematics; Minor in Computer Science & Statistics; GPA: 3.75 *May 2021*

· **Coursework: Discrete Mathematics; Data Structures; Real Analysis; Database Management; Ordinary Differential Equation; Regression Analysis**

TECHNICAL SKILLS

Programming:Python (Numpy, Pandas, Scipy, Tensorflow, Keras), R, Java, C++, html,
Data Management: SQL, Github
Data Visualization: Matplotlib, Seaborn, Tableau, Plotly, Kepler.gl
Machine Learning Algorithm: Regression, classification, KNN, SVM, Random Forest
Deep Learning Algorithm: Multilayer Perceptron, CNN, RNN, GAN, Transformer

PROFESSIONAL EXPERIENCE

Teradata Information System *Beijing, CN*
Data Engineer, Intern **(Python, C++)** *Feb 2022 - Apr 2022*

· **Maintained ETL opening of Bank of China credit card data warehouse, including but not limited to daily business work order, uploading, and downloading, warehousing, and model design**
· **Reviewed orders and transactions according to project manager's requirements and assisted in acceptance process**
· **Collaborated with colleagues from various departments, including the Business Department, to optimize correcting technical solutions based on unmet business needs**
· **Acquired crucial professional knowledge regarding Teradata financial data models and developed a general understanding of credit card data warehousing supervision**

Diamond Equity Research *New York, US*
Data Analyst, Intern **(Python)** *Sep - Oct 2021*

· **Extracted and aggregated sales data from company's internal databases with SQL, conducted data cleaning tasks including outlier detection and replacement, missing value imputation and variable type conversion, etc.**
· **Leveraged Business Intelligence tools including QlikView and Tableau to build trend line visualizations to examine underlying patterns behind sales revenue fluctuations in past 6 quarters**
· **Located anomalies based the Interquartile Range Rule and performed root cause analysis to uncover potential risk factors**
· **Built simulation models in Excel to analyze impact of fluctuations in macroeconomic and industry-wide factors on sales volume and revenue respectively, provided customized revenue growth strategies to stabilize earning performance**
· **Established a multiple linear regression model to pinpoint top influencing factors driving sales conversion rate, recommended corresponding sales operations solutions for maximized pipeline conversion**

ACADEMIC PROJECT

Geospatial Data Analytics *Sep - Dec 2022*
Deloitte **(Python)**

· **Retrieved Census demographic and economic data via API and created advanced filters to select and aggregate data from different census periods**
· **Converted shapefiles of census blocks to proper formats and visualized data of interest via multiple GIS tools such as Kepler.gl and OpenStreetMap**
· **Implemented translations between blocks over different years based on the target-density weighting method**

Biological Data Engineering *Sep 2022 - Jan 2023*
Frederick National Laboratory **(Python, C++)**

· **Developed a pipeline in the shell script that automated simulation of ligand (drug) - receptor (protein) process on different RAS systems via DOCK6.9 (molecular docking application), and prepared a detailed tutorial for user reference**
· **Tested different metrics evaluating binding process, including Hawkins GB/SA score function, Amber score function, and enrichment, on different RAS systems**
· **Visualized test results utilizing Python to analyze metrics' performance by analogy and proposed underlying issues within each metric**