

WONHA SHIN

Phone: +1 551-245-2356 | Email: leahnote01@gmail.com
Linkedin: <https://www.linkedin.com/in/wshin7> | Project Portfolio: <https://leahnote01.github.io/>

EDUCATION

University of Rochester

Master of Data Science | GPA 3.43 / 4.0

Rochester, New York

Aug. 2023 – Dec. 2024

Pusan National University

Bachelor of Arts in International Business and Economics | GPA 3.73 / 4.0

Busan, South Korea

Mar. 2011 – Aug. 2018

PROFESSIONAL EXPERIENCE

VOESH New York

Administrative Accounting Specialist, Admin/Accounting

South Plainfield, New Jersey

Oct. 2018 – Mar. 2023

- Transformed sales analytics for 10,000+ clients by developing a predictive & time-series framework, integrating Oracle DB for insightful data management that guided strategic actions for 60+ sales reps. Achieved a 15% boost in sales effectiveness and precision in forecasting trends.
- Revamped the company's CRM and ERP systems, reducing data errors by 70% and increasing cross-team efficiency. Strategically structured workflows to prepare the company for future technical insights, w/ forward-looking approaches.

EASTERN AMERICAN CDC (Non-Profit Federal Community Bank)

Credit Analyst Intern, Credit Assessment

Englewood, New Jersey

May. 2018 – Oct. 2018

- Drove financial stability by assessing risk profiles for 100+ SBA 504 loan applications, applying deep credit analysis to support robust decision-making and secure community business funding. Secured new clients with active analysis.

PROJECTS & INTERESTS

- LLM-Based Real-Time Translation System with Kafka and MLOps Automation (On-going)**
 - Spearheading a cutting-edge, real-time multilingual translation system that harnesses Kafka for seamless streaming and Kubernetes for on-demand scaling. Integrating Hugging Face models to deliver instant, high-accuracy translations, setting a new standard for accessible language processing.
 - Designing with resilience and adaptability, using Prometheus and Grafana for live monitoring and Weights & Biases (W&B) for real-time drift detection and retraining, ensuring consistent, high-quality language translation.
- Real-Time Tweet Sentiment Analysis Pipeline :** Engineered an advanced sentiment analysis pipeline using PySpark, Databricks, and AWS, managing high-velocity tweet streams to deliver instant sentiment insights. Optimized data flow with Delta Lake, enabling real-time analysis and fostering proactive decision-making across key data touchpoints.
- Data Mining: Cross-Cultural NLP Analysis of Luxury Hotel Reviews in Europe- LDA Topic Modeling :** Led a transformative NLP project, dissecting over 515,000 hotel reviews to reveal nuanced cultural patterns across Europe's diverse traveler demographics. Leveraged LDA Topic Modeling and Scikit-Learn to unlock insights, empowering hospitality businesses to align services with cross-cultural preferences and expectations.
- Multidimensional Analysis of Video Game Sales and Global Market Trends:** Directed a comprehensive data analysis project using RStudio, scrutinizing over 16,000 video game titles with advanced statistical methods
- ClinicOps: Integrated DBMS Solution:** Developed a robust DBMS using PHP and SQL, significantly enhancing data accuracy in clinical operations. Celebrated as a top technical project for its real-world impact on data management.
- Tri-States Area Quality of Life Analysis:** Delivered deep insights into living standards across NY-NJ-CT through data visualization in Tableau. Identified and showcased socio-economic factors impacting regional quality of life.

SKILLS

- Big Data & Database:** Apache Spark, AWS, Databricks, SQL
- Machine Learning & MLOps:** NLTK, NLP, TensorFlow, Keras, PyTorch, Scikit-Learn, RandomForest, SVM, Kafka, Kubernetes, Prometheus, Grafana