

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

- Developed a predictive & time-series framework, integrating Oracle DB for insightful data management, transforming sales analytics for 10,000+ clients, and improving forecasting accuracy by 15%. Revamped CRM and ERP systems, reducing data errors by 70% and increasing cross-team efficiency by designing structured workflows for future scalability.

EASTERN AMERICAN CDC (Non-Profit Federal Community Bank)

Credit Analyst Intern, Credit Assessment

Englewood, New Jersey

May. 2018 – Oct. 2018

- Conducted credit analysis for 100+ SBA 504 loan applications, mitigating financial risks and supporting funding decisions for small businesses. Assisted in securing new clients through active financial analysis and risk assessment.

PROJECTS & INTERESTS

- **LLM-Based Real-Time Translation System with Kafka and MLOps Automation:** Designed and implemented a real-time multilingual translation pipeline using **Kafka** for distributed streaming, **FastAPI** for high-speed API orchestration, and **Hugging Face** models for neural machine translation. Developed a robust **MLOps** framework, integrating **Prometheus** and **Grafana** for real-time monitoring and **Weights & Biases (W&B)** for drift detection and automated retraining, ensuring model adaptability.
 - Achieved sub-second latency for translations, maintaining response time into .001 sec, and enhanced translation quality across multiple languages with dynamic load handling. Ensured seamless scalability and maintainability, supporting millions of real-time translation requests per hour.
- **100 Days Learning Posting Challenge :** Successfully completed an intensive 100-day posting journey, mastering foundational and advanced topics in AI and deep learning, including linear regression, PCA, decision trees, SVMs, ensemble methods, CNNs, transformers, and LLMs. Explored innovative concepts such as Efficient Transformers, bias mitigation strategies, AI ethics, and conformal prediction, with a focus on real-world applications.
 - Documented daily progress, sharing technical insights on GitHub and LinkedIn to engage with the AI community. Developed and fine-tuned models like ResNet, CLIP, and GPT-based architectures, implementing practical use cases for visual and text-based AI systems.
- **E-Cigarette Perception Capstone Project (Chief Operative Manager):** Directed all coding, modeling, and analysis for a social media data project on e-cigarette perceptions. Leading a team applying NLP, sentiment analysis, and topic modeling to generate insights, guiding data-driven public health strategies.
- **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 empowering insights.
- **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.