Serap Ogut

ogutoykuserap@gmail.com | linkedin.com/in/serapogut/ | github.com/oykuserap

EDUCATION

Southern Methodist University - North Texas Scholar   Dallas, TX
Bachelor of Science: Computer Science and Data Science — GPA: 3.9   *Exp. Graduation: Dec. 2025*

SUMMARY

AI/ML-focused CS & DS senior graduating **Dec 2025**; experience in NLP/RAG, PyTorch, LangChain, Snowflake/Databricks. Delivered production-ready analytics and search systems. **Seeking full-time Data Scientist / ML Engineer roles (Jan 2026 start)**.

TECHNICAL SKILLS

Languages: Python, R, JavaScript, C++, SQL, HTML/CSS
Frameworks/Libraries: scikit-learn, PyTorch, Django, LangChain
Databases: PostgreSQL, Neo4j, Snowflake
Tools & Platforms: Google Cloud Platform, Azure DevOps, Databricks, Tableau, VS Code, Visual Studio, PyCharm, IntelliJ, Eclipse

PROFESSIONAL EXPERIENCE

Research Assistant   May 2025 – Present
Southern Methodist University   *Dallas, TX*
- Generated a Retrieval-Augmented Generation chatbot to query City of Richardson council meeting agendas from the early 2000s, enabling semantic search and context-aware responses.
- Processed and embedded **1500+** large-scale public records data, and presented capabilities to faculty and stakeholders.

Data Engineer Intern   May 2025 – Aug. 2025
CBRE   *Dallas, TX*
- Automated reporting workflows and built a metadata tagging/query catalog framework, projected to save **4+ hours** weekly for **93%** of users.
- Leveraged Snowflake, dbt, Python, SQL, and Azure DevOps to improve discoverability and governance for CBRE's industry-leading billions of rows of global real estate and investment data.
- Led the "Voices of Leadership – Women in STEM" event, interviewing senior professionals to share insights with interns.

Data Science Intern   May 2024 – Aug. 2024
Parkland Center for Clinical Innovation   *Dallas, TX*
- Collaborated on an innovative Suicide Screening project, utilizing a unique, large-scale dataset of over **3M** encounters, marking a first.
- Conducted comprehensive data analysis on Databricks and developed a machine learning model with logistic regression, achieving **75%** balanced accuracy and **83%** sensitivity.
- Facilitated weekly meetings with stakeholders to present data-driven insights, leading to successful project launch.

Research Assistant   May 2024 – Aug. 2024
Southern Methodist University   *Dallas, TX*
- Integrated an LLM using OpenAI's ChatGPT-3.5 and LangChain to identify best-matching professors for research interests across departments, enabling seamless natural language queries in a Neo4j database.
- Led a comprehensive study involving **184** researchers across 14 departments at Southern Methodist University.

Data Science Intern   Jan. 2024 – Jul. 2024
Richardson Adult Literacy Center   *Richardson, TX*
- Conducted data analysis and produced impact reports to demonstrate the value of literacy programs to stakeholders.
- Led a team in creating data visualizations and interactive dashboards using Tableau to present to the City of Richardson.

LEADERSHIP EXPERIENCE

Student Coordinator | *Thrive Scholars - Southern Methodist University*   May 2025 – Present

- Oversaw **25+** student mentors supporting **120+** freshmen and transfer students in engineering and related majors, fostering a cohesive program for 150+ participants.
- Organize and lead weekly coordinator–mentor meetings, resolving challenges and aligning on program goals with campus resources.

Founder & President | *SMU Dialogue Society*   Jan. 2023 – Present

- Established a student organization to promote intercultural and interfaith understanding on campus.
- Created a safe space for everyone's voice by organizing collaborative and interactive weekly workshops that promote meaningful connections.

PROJECTS

Search Engine | *C++*   Dec. 2022 – Jan. 2023

- Leveraged POSIX multi-threading, adhered to Google C++ coding & testing standards, and employed advanced C++ tactics to deliver scalable high-performance data-indexing and querying services.
- Designed an article indexing service using a multi-threaded producer–consumer paradigm with load-balancing mechanics to reduce data-indexing times.