





Justin Schulberg

 schulbergjustin@gmail.com
 www.linkedin.com/in/schulbergjustin
 www.github.com/jschulberg
 www.datacracy.us

EXPERIENCE

Rhombus Power

Washington, D.C.

Data Scientist

11/22 – present

- Developed, optimized, and fine-tuned *GPT-based LLMs* from Huggingface, leveraging embedding-based methods, customizing attention mechanisms, and fine-tuning transformer models for extractive semantic search
- Spearheaded the implementation of embedding-based methods for extractive semantic search, leading to a significant enhancement in information retrieval accuracy and reduced staff hours dedicated to laborious search and retrieval tasks
- Implemented and optimized supervised NLP-based machine learning models using *NLTK* and *spaCy*, incorporating feature transformation, model training, hyperparameter tuning, and model evaluation using tools like *scikit-learn*.
- Collaborated with DevOps to containerize NLP models using Docker, ensuring seamless integration into production environments and enabling scalable deployment.
- Conducted comprehensive analyses of NLP model performance, identifying areas for improvement and implementing enhancements for increased accuracy using model interpretability measures like Shapley values.

Akima

Washington, D.C.

Data Scientist

01/21 – 10/22

- Built and trained an entity resolution, machine learning model using the *dedupe* library in Python to synthesize ~63 million data points into a few thousand of interest, saving innumerable staff hours' worth of review time
- Created a self-service pipeline for the ingest of new data into the above machine learning model using *pyinstaller* and *tkinter*, and incorporation into live Power BI dashboards to ensure the continuity of ML models by non-technical users
- Developed a geospatial- and time-based clustering algorithm using *DBSCAN* in Python to add meaningful structure to an otherwise nebulous data challenge that stakeholders had deemed impossible to solve
- Presented the results of the aforementioned machine learning models in nearly a dozen settings across the Bureau to various audiences including live, virtual, technical, and non-technical individuals, with audience sizes as large as 275 people
- Developed interactive, mission-oriented Power BI dashboards, which have been used by nearly 100 stakeholders, including an Executive-level dashboard synthesizing operational data, which has been shown weekly to the Director of the FBI
- Automated a monotonous, data entry process that extracted information from recurrent emails, used the requests package in Python to search for additional information in the organization's online databases, and wrapped the code into a single executable file using *pyinstaller*, saving stakeholders nearly 20 hours/week of staff time

Deloitte Federal Consulting

Washington, D.C.

Strategy & Analytics Consultant

09/17 – 12/20

- Developed analytical tools using SQL, R, VBA, and Tableau that bring critical insights to new visa policies to Department of State (DoS) leadership, enabling quick reactions to emerging trends and allowing DoS to lead critical interagency efforts
- Collaborated with DoS leadership to translate analytics solutions into policy by supplying DoS Consular Affairs leadership and Congressional stakeholders with data-driven insight on U.S. visa policies using advanced modeling techniques in R
- Employed Natural Language Processing (NLP) techniques in R, including sentiment analysis, TF-IDF, and topic modelling, on ~50,000 social media posts to support the repatriation of American citizens during the beginning of the COVID-19 pandemic
- Translated DoS business logic into 5,000-line SQL script, generating new business measures from 100,000,000+ row dataset

- Built an enterprise-wide data strategy with a Chief Data Officer at Department of Homeland Security (DHS), which involved strategizing and developing an Analytics Measures Registry to innovate the current data and measure intake process at DHS
- Strategized and developed an Analytics Measures Registry to innovate the current data and measure intake process at the DHS by creating a central repository to track measures, metadata, and data
- Founded Deloitte's first Recruiting Business Intelligence team, engineering a data ingestion pipeline in R and building a Tableau dashboard to improve recruitment strategy, leading to a 40% YoY increase in acceptance rates

LEADERSHIP

Rutgers University Student Assembly (RUSA)

New Brunswick, NJ

Student Body President

05/14 – 05/17

- Served as the leader to represent and advocate for 33,000 undergraduate Rutgers University students, overseeing an executive team of 15 and an organization of 80 individuals
- Spoke in front of crowds ranging in size from groups of a few hundred students to Convocation with 7,500 students

EDUCATION

Georgia Institute of Technology

Master of Science, Data Analytics

Graduated: Dec. 2023

GPA: 3.8/4.0

Rutgers University

School of Arts and Sciences Honors Program

Sep. 2013 - May 2017

GPA: 3.85/4.0

Majors: Mathematics, Political Science

Minor: Spanish

Study Abroad: *La Academia Latinoamericana de Español* (Cusco, Perú)

ADDITIONAL INFORMATION

Certifications: Tableau Desktop Specialist (03/2020), Microsoft Azure Fundamentals (05/2020)**Tools:** Expert in Tableau, Advanced in Python, Advanced Proficient in R, Advanced Proficient in SQL, Intermediate in VBA, Intermediate in Power BI, Beginner in Power Automate**Languages:** Advanced Proficient in Spanish, Intermediate in Mandarin, Beginner in Hindi**Clearances:** FBI Top Secret SCI w/ CI Polygraph, DOD Top Secret, DHS Public Trust