Brenden Brusberg

Data Science Specialist

Hoboken, NJ, USA
✆ *1-201-401-6723*
✉ *brenbrus@gmail.com*
⎙ *brendenbrusberg.com*
in *brendenbrusberg*
⊕ *brusberg*

Summary

I am a data scientist at McKinsey & Company, focusing on the crossroads of data science and software engineering for applications in machine learning, graphs, natural language processing, and model interpretability. What I enjoy the most in my role is the ability to interact with the entire practical implementation of a model. When presented with a problem, I am a part of the process from ideation to when it is a fully serviceable model deployed in the cloud.

Work Experience

Data Science Specialist, McKinsey & Company, *NYC, NY*   Dec 2023 – Current
- Developing graph based algorithms in conjunction with Neo4j generative AI to preform advanced people analytics.
- Democratizing NLP use around the Firm with custom language models, BERTopic models and LLMs. Implementing RAG, chain of thought, and data extraction with GPT-4 and open source models like Llama-2 and Mistral.

Senior Data Scientist, McKinsey & Company, *NYC, NY*   Dec 2021 – Dec 2023
- Expanding our People Analytics capabilities with GenAI, natural language processing tools and machine learning pipelines.
- Integrating explainable AI tools like SHAP to make analyses more interpretable and impactful.
- Responsible for identifying emerging data science methods and leading team deep-dives implementing new packages and papers.
- Enabling better cross-team development with CI/CD pipelines in GitHub Actions and AWS CloudFormation.

Data Scientist, McKinsey & Company, *NYC, NY*   Intern: May 2019 – Feb 2020 | Full-time: Feb 2020 – Dec 2021
- Python development in Kedro building production-grade machine learning analytic pipelines. Outperformed vendor model saving the Firm over a million dollars annually.
- Created scalable AWS infrastructure, scaling model inference to be invoked over 1.2 million times annually without error.
- Developed topic models using Sentence-BERT and BERTopic to leverage vector space with sentence and document embeddings.
- Implemented custom objective functions for XGBoost. Developed embeddings for high cardinality categorical features.

Education

Master of Science, Machine Learning, Stevens Institute of Technology, *Hoboken, NJ*   Dec 2021
GPA – 3.97
- Degree with a focus on machine learning, analytics, and statistics. Courses involved deep dives into deep learning, distributed learning, game theory, NLP, and advanced statistical methods behind machine learning models.

Bachelor of Science, Computer Science, Stevens Institute of Technology, *Hoboken, NJ*   May 2021
GPA – 3.82, **Mathematics Minor, Literature Minor**
- **TA**: CS-392 Systems Programming, CS-677 Parallel Programming of Many-Core Processors
- **Upsilon Pi Epsilon**: Member of the International Honor Society for the Computing and Information Disciplines
- **Honors**: Dean's List, Presidential Scholarship, and Accelerated Masters Program

Projects and Skills

Evaluation of Fine-Tuned Language Transformers
- Fine-tuned BERT and Sentence-BERT in PyTorch with over 1.5 TB of Reddit comments with additional vocabulary tokens to compare the effectiveness of different training techniques.
- Evaluated the models through the performance of the produced embedding space and other downstream classification tasks.

Word2Vec Embeddings Clustered with K-Means in CUDA
- Understanding the optimization process of taking algorithms off the CPU and into GPU memory: K-Means clustering on large word embeddings provided by Glove (2.4 Million Vectors with 300 latent dimensions).

Software/Tools: Linux, Docker, Visual Studio, Copilot, Kedro, Jupyter, Flask/FastAPI, AWS Services & SAM, OpenGL, Git
Languages: Python, CloudFormation, C++, C, CUDA, Rust
Acheivements: Eagle Scout, Order of the Arrow, President of the Stevens Rock Climbing Team