Nyack, NY
Portfolio: github.io.jbecker7326

Jennifer Becker

Email : jbecker7326@gmail.com
Mobile : +1-845-826-4836

Education

- **Georgia Institute of Technology** — Atlanta, GA
  Master of Science in Computer Science; GPA: 4.0 — *Aug. 2022 – Present*
- **State University of New York at New Paltz** — New Paltz, NY
  Bachelor of Science in Environmental Geochemistry; GPA: 3.5 — *Aug. 2013 – May 2017*

Experience

- **Anchor QEA** — Woodcliff Lake, NJ
  Data Scientist — *Mar. 2020 - Aug. 2022*
  - Collaborated with project executives to comprehend business problems and propose custom dashboard designs with real-time database integration, tailored to client-specific requirements.
  - Leveraged Python to develop data extraction, cleansing, visualization, and statistical analysis algorithms, enabling the delivery of meaningful insights and innovative business solutions.
  - Maintained strong accountability for data completeness and quality control in relational databases, emphasizing meticulous attention to detail and data management skills.
  - Strategically optimized ETL pipelines to gather, prepare, cleanse, and transform client data, resulting in significant efficiency gains in data workflows between data science and project management teams.
- **Grid Logistics** — Kearny, NJ
  Data Analyst — *Aug. 2018 - Mar. 2020*
  - Developed algorithms with R and LaTeX for generating statistical analysis tables, formatted spreadsheets, and PDF summary reports, contributing to data-driven decision-making processes.
  - Designed SQL queries and analyzed environmental compliance reports to provide actionable recommendations for stakeholders, demonstrating strong problem-solving and consulting skills.
  - Fostered client relationships through proactive site visits, frequent meetings, and meticulous management of soil logistics, ensuring seamless communication and customer satisfaction.

Projects

- **Molecular Graph Network**
  Links: GitHub, Report — *July 2023*
  - Implemented a Graph Neural Network (GNN) architecture using PyTorch Geometric for precise node classification within the protein-protein interaction (PPI) dataset.
  - Designed a novel combination of Molecular Fingerprint Convolution (MFConv) and Graph Isomorphism Network (GIN) layers, resulting in an F1 score within 5 points of the benchmark while reducing runtime by 67%.
- **Yelp Fake Review Analysis**
  Links: GitHub, Video Presentation — *May 2023*
  - Deployed various machine learning classifiers, including SVM, KNN, Naive Bayes, Logistic Regression, and Random Forest, via scikit-learn to effectively classify fake reviews within the Kaggle Yelp dataset.
  - Engineered innovative features for sentiment analysis with TF-IDF and implemented oversampling with SMOTE, resulting in a notable 12% increase in recall.
  - Developed a user-friendly, interactive visualization application employing D3.js as an analytical tool to uncover sectors negatively impacted by fake reviews within suburban populations.

Skills

- **Languages**: Python 3 (Matplotlib, NumPy, Pandas, PySpark, PyTorch, scikit-learn, SciPy, SQLite), R (data.table, envstats, ggplot2), SQL (PostgreSQL, T-SQL), JavaScript, Java, Scala, C#, HTML5, CSS.
- **Technologies**: Amazon Web Services (AWS), DataBricks, Docker, Git, Google Cloud Platform (GCP), Tableau.