HENRY TAPPA

Data Scientist

htappa@gmail.com
(571) 481-6310
htappa.github.io

SUMMARY

Seasoned Data Scientist with an M.S. in Data Science and over 8 years of experience in technology and renewable energy sectors. Proficient in Python, SQL, and R, with expertise in machine learning, statistical modeling, experiment design, and project management. Proven track record of leading significant data-driven projects, developing data infrastructure, and collaborating with cross-functional teams to drive business growth and innovation.

SKILLS

General Skills: Machine Learning, Deep Learning, Data Visualization, Data Preparation, Data Analysis, Statistical Analysis, Project Management.

Programming Languages: Python, R, SQL, HTML, CSS, JavaScript, Visual Basic.

Software Systems: Pytorch, Tensorflow, Scikit-Learn, Pandas, Numpy, Tableau, Git, Visual Studio Code, Google Cloud Platform, Apache Cassandra, MongoDB, Azure Machine Learning, Google Analytics, GIMP, SharePoint, MS Access, MS Excel.

EXPERIENCE

Allegheny Science & Technology
Data Scientist, August 2023 – June 2024
Client: New York State Energy Research and Development Authority (NYSERDA)

- Led the development and deployment of machine learning models to estimate the number of 2-8 person households across New York counties, utilizing multivariate analysis techniques.
- Designed and implemented advanced database systems using SQL to support clean energy component supply chains, enhancing data-driven decision-making processes.
- Developed and maintained analytical dashboards in Tableau to provide actionable insights into clean energy supply chain segments and potential job growth impacts.
- Authored a comprehensive Consumer Protection Plan for the EmPower+ initiative, ensuring alignment with DOE's HEAR program requirements to enhance customer experience and program integrity.

Allegheny Science & Technology
Data Scientist, February 2020 – January 2021

- Spearheaded the development of the Energy Benchmarking & Monitoring Tool, utilizing SQL for data storage and HTML/JavaScript for interface development, as part of a new business development initiative while simultaneously working as Senior Program Data Analyst.
- Conducted hands-on analysis of building energy usage data, creating statistical models to benchmark energy efficiency performance.
- Collaborated with product managers and engineers to transition proof-of-concept models into production, optimizing models for real-world applications.
- Designed and conducted experiments to gather data on energy usage patterns, performing statistical analysis to drive insights and recommendations.

Allegheny Science & Technology
Senior Program Data Analyst, January 2018 – August 2023
Client: U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy

- Co-managed the Small Business Innovation Research (SBIR) program for the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE), developing strategy for small business engagement, training federal clients, and coordinating EERE's 3.65% RD&D investment in the program.
- Led the data science efforts for the SBIR program, developing machine learning regression models to predict future government funding allocations.
- Built interactive dashboards and analytical frameworks to provide senior leadership with data-driven insights, supporting strategic decision-making.
- Managed end-to-end machine learning projects, from data sourcing and preprocessing to model training, validation, and deployment.
- Provided mentorship and training to junior analysts, fostering a culture of continuous learning and professional development within the team.

Allegheny Science & Technology
Data Analyst, February 2016 – December 2017
Client: U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy

- Delivered comprehensive project management, data analysis, and database engineering expertise to the Technology to Market (T2M) team for the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE).
- Developed relational databases and SQL queries for extensive data mining, providing client with detailed insights into T2M activities across various states and congressional districts.
- Created statistical models and interactive data analysis packages to inform management on the impact of Small Business Vouchers Pilot and the Lab Embedded Entrepreneurship Program.
- Identified the need for improved collaborative tools and project management processes within the T2M team and assembled two working groups, having since led the testing of multiple software systems to address T2M needs and challenges in the technology space.

EDUCATION

- Master of Science, Data Science, George Washington University, 2020
- Bachelor of Arts, International Affairs, Eastern Washington University, 2009

CERTIFICATIONS

- Microsoft Professional Program Data Science Certificate, Microsoft, 2017