

## EXPERIENCED DATA SCIENTIST

---

I am a data scientist with nearly a decade of combined experience in industry and academia, creating data-driven solutions with passion, curiosity, and skill. I keep myself updated by learning from new articles and technologies, and collaborate well with others. I also have a strong sense of ownership, business focus, project management, and mentorship skills. I communicate science effectively to both technical and non-technical audiences, using expert data presentation and story telling skills. I have presented at several international conferences and published three papers in reputable peer-review journals.

## KNOWLEDGE

---

### Technical Skills

- **Programming Languages** - Python (*expert*), SQL (*intermediate*), shell (*intermediate*), HTML (*working*), CSS (*working*)
- **Data Science Software** - NumPy, SciPy, Pandas, Polars, Spark/PySpark, Sci-Kit Learn, LightGBM, Streamlit, Plotly/Dash, MS Excel, Power BI, H2O.ai, Azure ML
- **Infrastructure** - Databricks, Jupyter (Lab, NB), MLFlow, Docker, Cloud services: Azure (Blob Storage) and AWS (S3), Google Big Query, Kubernetes, Airflow, Github Actions, Azure Dev Ops
- **Machine Learning** - Linear/Multi-Linear Regression, Tree-based Regression (e.g., RF, GBM), Clustering (e.g., KNN, K-Means), Projection Methods (e.g., PCA, T-SNE), Time-Series Analysis and Forecasting, Digital Signal Processing, Linear Algebra, Inverse Methods, Factor Decomposition, Multi-Variate Analysis, Statistical/Empirical/Physical Modeling, Simulations, Statistical Analysis/Inference, Hypothesis Testing, Experimental Design, Hyper-parameter Tuning, Decision Making, Risk Analysis

### Soft Skills

- **Project Management** - Planning (Jira / Azure Dev Ops), ownership, daily stand-up, book-keeping (Excel, Word, OneNote, etc.)
- **Communication** - Verbal and non-verbal presentations, technical report writing, Zoom/Slack/MS Teams
- **Mentorship** - Expertise in explaining technical subjects to technical/non-technical audiences, history of mentoring colleagues and students for entire project life-cycles

## EXPERIENCE

---

### Coyote Logistics (A UPS Company)

Remote, USA

*Data Scientist, 3PL Logistics*

*Jun. 2022 - Sept. 2023*

- Supported and improved a live API multi-model service that predicts costs for truck loads on an auction market, generating on the order of \$10s of millions in revenue each year.
- Led our pricing team and improved data science communication standards with internal business partners.
- Expanded model to Canadian trucking lanes, forecast to increase revenue by about \$500,000 (20%).
- Developed, obtained buy-in, and deployed an add-on linear pricing model, which reduced pricing bias from -30 - 10%, to 0% Mean Percent Error.

### Bayer AG, Crop Sciences

Remote, USA

*Data Scientist, Agronomic Sciences*

*Oct. 2021 - Jun. 2022*

- Headed research and development efforts to improve and propose new tree-based machine learning models that predicted corn seed yield using historical growth and weather data.
- Created an automated ML model testing and performance bench marking framework for my team, in Python, to improve experimenting efficiency, cadence, and capabilities.
- Presented research updates to our partnering teams each week and gave quarterly progress updates to our stakeholders.

## EDUCATION

---

### University of Liverpool, UK

*Ph.D., Geophysics*

*Sept. 2015 - Oct. 2019*