Data Scientist

Data SMB Enterprise

We are hiring a Data Scientist to apply statistical modeling, machine learning, and experimentation to solve high-impact business problems at [Company Name]. You will work with large datasets, build predictive models, and partner with product and engineering teams to bring data-driven solutions into production. This role blends rigorous analytical thinking with practical software engineering to deliver measurable business outcomes.

Key Responsibilities

- Develop and validate predictive models (classification, regression, clustering, time series) to solve business problems like churn prediction, recommendation, demand forecasting, or fraud detection
- Design, implement, and analyze A/B tests and multi-variant experiments with proper statistical methodology
- Collaborate with product and engineering teams to define model requirements, integrate models into production systems, and monitor model performance over time
- Perform exploratory data analysis on large datasets to identify patterns, generate hypotheses, and inform business strategy
- Build and maintain reproducible analysis pipelines using Python, Jupyter notebooks, and version-controlled workflows
- Communicate model results, limitations, and recommendations clearly to both technical and non-technical stakeholders
- Stay current with advances in ML/AI research and evaluate new techniques for applicability to business problems

Required Skills & Experience

- 3+ years of data science experience with a track record of deploying models that impacted business outcomes
- Strong programming skills in Python (pandas, NumPy, scikit-learn) or R for data analysis and modeling
- Solid foundation in statistics: hypothesis testing, Bayesian inference, regression analysis, experimental design
- Experience with machine learning algorithms: ensemble methods, gradient boosting (XGBoost, LightGBM), neural networks
- Advanced SQL skills for data extraction and feature engineering from large datasets
- Experience with model evaluation, cross-validation, and techniques to prevent overfitting
- Ability to communicate technical results to non-technical stakeholders through clear visualizations and narratives
- Master's or PhD in a quantitative field (Statistics, Mathematics, Computer Science, Economics, Physics) or equivalent work experience

Nice-to-Have

- Experience deploying ML models to production using MLflow, SageMaker, or Vertex AI
- Familiarity with deep learning frameworks (PyTorch, TensorFlow) for NLP or computer vision tasks
- Experience with causal inference methods (difference-in-differences, instrumental variables, propensity scoring)
- Knowledge of Bayesian modeling and probabilistic programming (PyMC, Stan)
- Experience with feature stores or ML platforms (Feast, Tecton)

Tech Stack

Python | scikit-learn | XGBoost | PyTorch | SQL | Jupyter | pandas | MLflow | Airflow | Snowflake | AWS SageMaker | Git

What We Offer

- Competitive salary and equity at [Company Name]
- Access to GPU compute resources and modern ML infrastructure
- Conference and research paper budget (NeurIPS, ICML, KDD, etc.)
- Comprehensive health, dental, and vision insurance
- Flexible remote work arrangement with async-first culture
- Opportunity to solve high-impact problems with real business outcomes

Interview Process

1. Recruiter phone screen (30 min) — background, experience, and role expectations
2. Technical screen (60 min) — statistics fundamentals, ML concepts, and a Python coding exercise
3. Case study (take-home, 3-4 hours) — analyze a realistic dataset, build a model, and present findings
4. On-site or virtual loop (3 hours) — case study presentation, ML system design, and coding deep-dive
5. Hiring manager conversation (45 min) — research interests, collaboration style, and career goals
6. Reference checks and offer