

Job Description: Junior Data Scientist – AI

About the Role

We are looking for a **Junior Data Scientist** to join our AI/ML team and contribute to **Proof-of-Concept (POC) projects** that explore how emerging AI technologies can be applied to real-world enterprise problems.

You will work closely with senior Data Scientists, MLEs, and AI Architects to build and validate prototypes in areas such as **RAG (Retrieval-Augmented Generation)**, **semantic search**, **prompt engineering**, **document clustering**, and **LLM-based automation**.

This is a hands-on learning role — ideal for someone passionate about building, experimenting, and rapidly turning ideas into working prototypes.

Key Responsibilities

- **POC Development**
 - Assist in building experimental models, scripts, and pipelines for AI/ML POCs.
 - Explore new ideas using embeddings, transformers, vector databases, and re-ranking models.
 - Evaluate models for accuracy, recall, and efficiency on sample datasets.
- **Data Exploration & Preprocessing**
 - Clean, analyze, and prepare datasets for model training and evaluation.
 - Work with both structured and unstructured data (text, documents, logs, etc.).
- **Experimentation & Benchmarking**
 - Conduct small experiments using pre-trained LLMs (OpenAI, Hugging Face, local models).

- Measure and report improvements across benchmarks (Recall@K, NDCG, precision, latency).
 - **Collaboration & Documentation**
 - Collaborate with engineers and business analysts to translate ideas into working demos.
 - Document experiment setups, results, and learnings clearly for internal sharing.
-

Requirements

- **Education:** Bachelor's or Master's in Computer Science, Data Science, AI/ML, or related fields.
 - **Technical Skills:**
 - Good understanding of Python, Pandas, NumPy, and Scikit-learn.
 - Basic knowledge of NLP and deep learning concepts.
 - Familiarity with Hugging Face, LangChain, or vector databases (Pinecone, FAISS, Chroma) is a plus.
 - Comfortable using Jupyter/Colab for experiments.
 - **Mindset:**
 - Curious, experimental, and self-driven.
 - Enjoys reading papers, testing new models, and finding practical ways to apply them.
 - Strong problem-solving attitude — not afraid to debug or try multiple approaches.
-

Nice-to-Have (Bonus Points)

- Exposure to **LLMs**, **RAG pipelines**, **semantic search**, or **embedding models**.
 - Experience with Git, Streamlit, or simple API integration for POC demos.
 - Basic understanding of evaluation metrics (precision, recall, MRR, etc.).
 - Participation in hackathons, Kaggle competitions, or personal ML projects.
-

What We Offer

- Opportunity to work directly with **cutting-edge AI/LLM tech** from day one.
- Mentorship from experienced data scientists and AI architects.
- Exposure to real-world enterprise problems across multiple domains.
- Fast-paced growth path toward **Applied ML Engineer** or **Research Data Scientist** roles.
- A learning-first culture where curiosity is valued more than perfection.