Senior Data Scientist - GenAI, LLM & NLP

Build Production AI & LLM Solutions

募集職種

人材紹介会社
SThree株式会社

求人ID
1619506

業種
インターネット・Webサービス

雇用形態
派遣

勤務地
東京都 23区

給与
時給制

勤務時間
9am to 5:30pm (hybrid)

更新日
2026年10月29日 11:57

応募必要条件

キャリアレベル
中途経験者レベル

英語レベル
ビジネス会話レベル

日本語レベル
日常会話レベル

最終学歴
大学卒： 学士号

現在のビザ
日本での就労許可は必要ありません

募集要項

Our client is a leading technology-focused organization operating one of the largest digital ecosystems in Japan. Their data platform team is responsible for managing and enriching large-scale product and business data used across multiple services and platforms.

The team leverages machine learning, natural language processing, large language models, and cloud technologies to automate data enrichment, improve data quality, and drive business value through AI-powered solutions.

This is an opportunity to work on real-world AI applications at scale, collaborating closely with engineering, product, and business stakeholders while bringing cutting-edge AI solutions from concept through production.

Senior Data Scientist (GenAI / NLP / LLM)

We are looking for an experienced Data Scientist to help design, develop, and deploy AI-powered solutions that improve and enrich large-scale data platforms.

This is a highly hands-on role where you will work across the entire data science lifecycle, from exploratory analysis and hypothesis testing through model development, evaluation, deployment, and continuous improvement.

Responsibilities

- Analyze large-scale structured and unstructured datasets to identify patterns, opportunities, and data quality issues
- Translate business challenges into practical machine learning and AI solutions
- Design and develop NLP and LLM-based applications
- Build data enrichment and automation solutions using state-of-the-art AI technologies
- Evaluate model and LLM performance using statistical methodologies and evaluation datasets
- Develop scalable data processing pipelines using cloud technologies
- Partner with engineering teams to integrate AI solutions into production environments
- Drive data-informed decision making through analysis, experimentation, and stakeholder communication
- Improve existing manual processes through intelligent automation and AI-driven workflows

Technology Stack

- Python
- SQL
- GCP / BigQuery
- Azure OpenAI / OpenAI
- Generative AI
- LLMs
- NLP
- RAG
- Git
- Cloud-based data platforms

Why Join?

- Work on cutting-edge AI and GenAI initiatives
- Build real-world production systems rather than research-only projects
- Large-scale datasets and impactful business challenges
- International working environment
- Long-term project with strong growth potential

Location

Tokyo, Japan

Work Style

- Hybrid work environment

スキル・資格

Required Experience
- 5+ years of experience in Data Science, Machine Learning, AI Engineering, or related fields
- Strong Python development skills
- Advanced SQL skills and experience working with large datasets
- Practical experience with Machine Learning and NLP
- Experience designing and implementing production AI solutions
- Hands-on experience with LLM-based applications
- Ability to perform exploratory data analysis and hypothesis-driven problem solving
- Strong communication skills and ability to work with both technical and non-technical stakeholders

Preferred Experience
- Azure OpenAI or OpenAI APIs
- Retrieval-Augmented Generation (RAG)
- Prompt Engineering
- AI Agents
- GCP / BigQuery
- LLM Evaluation Frameworks
- LLMOps
- Search & Retrieval Systems
- Knowledge Graphs
- AI-assisted software development tools

Language Requirements (Important)
Japanese: Conversational level minimum (Business level preferred)

English: Business level

Please note that due to project requirements, we will only contact candidates who possess at least conversational-

level Japanese language ability. Applications from candidates without Japanese communication skills will not be considered.

会社説明

Our client is a leading technology-focused organization operating one of the largest digital ecosystems in Japan. Their data platform team is responsible for managing and enriching large-scale product and business data used across multiple services and platforms.

The team leverages machine learning, natural language processing, large language models, and cloud technologies to automate data enrichment, improve data quality, and drive business value through AI-powered solutions.

This is an opportunity to work on real-world AI applications at scale, collaborating closely with engineering, product, and business stakeholders while bringing cutting-edge AI solutions from concept through production.