CareerCross

G Talent — Professional IT Recruitment | 外国人ITエンジニアの転職支援
- Many IT jobs for global companies and high-profile startups!
- IT specialty bilingual consultants support your career change!

【MLOps Platform Engineer】 Data Strategy Office

Listed Company★Flextime&Remote Work OK

Job Information

Recruiter
G Talent at Bizmates, Inc.

Hiring Company
Development of SaaS platform

Job ID
1489481

Industry
Internet, Web Services

Job Type
Permanent Full-time

Location
Tokyo - 23 Wards

Salary
7.5 million yen ~ 12 million yen

Work Hours
9:30~18:30 *Flextime

Refreshed
August 5th, 2024 16:59

General Requirements

Minimum Experience Level
Over 3 years

Career Level
Mid Career

Minimum English Level
Business Level

Minimum Japanese Level
Business Level

Minimum Education Level
High-School or Below

Visa Status
Permission to work in Japan required

Job Description

------------- 【About the company ...】 -------------

The company was established in May 2012 to solve the money issues of individuals and companies through the services, and to make users' lives more positive and forward-thinking.
Through their services, they aim to become the No. 1 "money platform" in Japan by relieving individuals of their money worries and anxieties, improving corporate management, and increasing domestic productivity.

------------- 〖 Job Description 〗 -------------

In this position, you will be at the forefront of the latest technologies in container orchestration, cloud services, and CI/CD pipelines to ensure efficient development, training, and deployment of ML models.
You will design and implement optimization strategies, operate and maintain a scalable and robust infrastructure for ML projects, and autonomously empower ML engineers throughout the MLOps cycle.

・As an MLOps Platform Engineer, you will play a key role in enabling a team of ML engineers to efficiently develop, train, and deploy ML projects using the latest technologies in container orchestration, cloud services, and CI/CD pipelines for data collection, model training, and monitoring in production environments. Play a key role in enabling the ML engineering team to efficiently develop, train, and deploy ML projects using the latest technologies for data collection, model training, and monitoring in production.
・Build and maintain a scalable infrastructure to run ML projects, committed to results and user value
・Design and implement strategies for prediction servers, data and training pipelines to ensure efficient use
・Design and build inference platforms to ensure reliability and high performance
・Provision and monitor infrastructure resources Build and maintain ML workflows and pipelines
・Deploy and maintain monitoring services
・Ensure compliance with security best practices

【Development Environment】
AI/ML: Bigquery,Vertex ai,AI Ops,Redshift,Databricks,Glue,SageMaker,TensorFlow,PyTorch,Kubeflow
AWS: SageMaker,EKS,ECS,Lambda,Step Functions
Middleware: Docker,Terraform,Kubernetes
Programming languages: Python, Go

【Tools】
Groupware: Google Workspace
Repository management: GitHub
CI/CD: CircleCI, GitHub Actions
Monitoring: Datadog, Grafana
Communication: Slack
Ticket management: Jira

【 Working time 】

9:30~18:30 *Flextime

【 Welfare 】
■Full Social Insurance
■Commuting Allowance
■Defined Contribution Pension
■Employee Shareholding System
■Medical Checkup
■Training/Self Development Allowance
■Housing allowance for the neighborhood and celebration for moving to the neighborhood
■Influenza vaccination
■Subsidy for purchasing books
■Preferential treatment when using the following services (*limited to the company's contracted businesses)
  - Rental brokerage
  - Housekeeping service
  - Babysitting
  - Online English conversation school

【 Holiday 】

  - Saturday/Sunday/National Holiday
  - Annual Paid Leave
  - New Year Holiday
  - Summer Holiday
  - Winter Holiday, etc

Required Skills

【**Required**】
Experience deploying and operating AI/ML models
Experience in development and operation of data pipelines and jobs that process large amounts of data
Experience with AWS infrastructure design and construction
Experience with AWS cloud services Experience using Kubernetes to deploy and manage containerized applications with high availability and performance
Effective communication skills for collaboration with technical and non-technical team members

【**Preferred**】
Experience in AI/ML model development
Contributions to open source projects
Experience working to improve AI/ML models and model performance, including fine tuning
Experience with CI/CD practices
Proficiency in programming (Python or Go) Excellent problem solving skills and ability to work in a dynamic environment

【Language】
English: Business level proficiency (TOEIC score of 700 or above)
Japanese: N3 level proficiency

Company Description