



GEMINUS

720 University Avenue Suite 200  
Palo Alto CA94301 U.S.A.  
Ph: (650) 234-0400

<https://geminus.ai/>  
Twitter: @Geminus\_AI  
Email: info@geminus.ai

## Geminus: Data Scientist

### **About Geminus**

Geminus is the world's first industrial process centric digital twin product that delivers self-optimized design, predictive operational intelligence and asset performance enhancement to operators and engineering service providers. The product leverages physics-constrained artificial intelligence to deliver high-fidelity O&M intelligence and to continuously optimize the designs of assets & processes. Geminus integrates with operational data through SCADA systems, historian software, sensor gateways and enterprise asset management to trigger physics-constrained AI modeling of the end-to-end process for every significant time slice of data. This generates high-fidelity operational intelligence and design recommendations to prevent multi-step failures, adverse shifts in yield, structural process faults and degraded remaining useful lifetime. The product is currently under deployment in automotive, energy and hi-tech manufacturing verticals.

### **About the role**

The Data Scientist will drive the design and development of key artificial intelligence (AI) components of the platform that integrate with its multi-fidelity physics-based modeling capabilities. The spectrum of applications of AI includes black-box IoT modeling, consolidated error modeling across multiple physics-based models, supervised & unsupervised training of parameters & data-driven functions in hybrid physics-based models and meta-learning for performance boosting.

### **Responsibilities**

As a Data Scientist of a fast-growing startup, the successful candidate will be leading the development of key aspects of Geminus' product:

- Model training, model analytics and serving using state-of-art TensorFlow based orchestration techniques
- Selecting features, building & optimizing predictive accuracy and error modeling using a combination of machine learning and deep learning techniques
- Design and development of the AI components of the data augmented physics-based model training from batches of operational data across multiple heterogeneous physics-based models
- Design and development of the AI-components involved in model analytics, selection and servicing for delivering optimal design & operational intelligence
- Enhancing data collection procedures and data quality management from both numerically simulated and operational data sources
- Processing, cleansing, and verifying the integrity of data used for analysis
- Creating automated anomaly detection systems and constant tracking of its performance

### **About you**

The successful candidate will have experience in working in innovative projects with fast-paced delivery schedules in startups & large enterprises:

- Proven track record of analyzing large-scale complex data sets, modeling and machine learning algorithms
- Excellent understanding of machine learning techniques and algorithms, such as k-NN, Naive Bayes, SVM, Decision Forests, etc.
- Experience with common data science toolkits, such as R, Python (NumPy, SciPy, Pandas), Matlab etc.
- Experience in deep learning frameworks (e.g., Tensorflow, MxNet), and Large-scale optimization preferred.



720 University Avenue Suite 200  
Palo Alto CA94301 U.S.A.  
Ph: (650) 234-0400

<https://geminus.ai/>  
Twitter: @Geminus\_AI  
Email: [info@geminus.ai](mailto:info@geminus.ai)

## GEMINUS

- Proficiency in using query languages such as SQL
- Experience with big data infrastructure like HDFS, Spark data-frames, MongoDB etc.
- Good applied statistics skills, such as distributions, statistical testing, regression, etc.
- 5 - 10 years of experience in Applied Machine learning
- Educational background or prior experience in engineering physics and/or simulation is desirable

Please send your resume to [jobs@geminus.ai](mailto:jobs@geminus.ai).