

Advanced Data Scientist (Level 2)

About Liveline

Liveline enables dramatic improvements in manufacturing performance through a unique application of artificial intelligence to provide real-time process control and predictive assistants for plant personnel. Our focus is on automating complex processes, not simply providing dashboards for managers and operators.

Our team combines experts in AI with world-class process engineers who can focus on the “last mile” with customers: Extracting data from the process and implementing controls on the shop floor. We speak the language of AI but also industrial controllers.

Our hardware and software offerings are scalable and cost-effective whether customers have one production line or hundreds, delivering an ROI that’s attractive to small and medium-sized enterprises.

We are passionate about democratizing the power of analytics and advanced automation for manufacturers of almost any size. Through our approach, producers can de-mystify complex processes and free up valuable technicians to focus on more advanced tasks instead of constantly monitoring and adjusting equipment parameters.

Role Description

Data Scientist Level 2 is responsible for developing advanced models of physical production processes and process controls using machine learning techniques. These models optimize real-time process controls and to provide predictive analytics for process operators. Models are developed with the intention of re-using and updating with data from new but similar processes. This role will implement advanced modeling techniques, create the data science platform and mentor junior data scientists.

Primary Responsibilities

- Design, develop, and iterate on world class algorithms and predictive models, leveraging statistical, machine learning, and AI methods, particularly in the areas of time series forecasting and anomaly detection
- Deploy models for inference within mission critical manufacturing environments, optimizing for real-time and near time response
- Define model success, provide evidence regarding the performance of your models
- Present and communicate the data insights/findings (preferably actionable insights) to both specialists as well as a non-technical audience
- Provide technical direction and supervision to data science team
- Assess the effectiveness and accuracy of new data sources and data gathering techniques
- Collaborate with business subject matter experts to select the relevant sources of information

- Make strategic recommendations on data collection, integration and retention requirements incorporating business requirements
- Collaborate with process domain experts on what data to collect and quality of data
- Define Metrics for ML model and hyperparameter optimization

Experience:

- Master's Degree/PhD in engineering, science, computer science required
- 3+ years of industrial experience using predictive models, data science,
- Demonstrated skills in time series data analysis
- Demonstrated experience in DL including LSTM, RNN and CNN
- Experience with cloud technologies; AWS, Azure, Databricks, or Hadoop/Spark
- Experience in interacting with data ingestion pipelines or data warehouse
- Excellent statistical intuition and knowledge of various analytical approaches, Bayesian techniques

Hard Skills:

- Strong knowledge of statistics and experience using statistical packages for analyzing large datasets
- Expertise with popular data analysis libraries like Numpy, Pandas, Scikit-learn, SciPy, Matplotlib, PyTorch, Tensorflow, Pyro, Spark ML, StatsModel
- Working on Linux ML libraries and compilation
- Expertise in one or more of the following languages: MatLab, Python, R, C, C++, java
- Expertise with data visualization
- Proficiency in advanced machine learning methods, including comprehension of theory, modeling/identification strategies and limitations and pitfalls
- Experience with Code development tools like GitHub

Soft Skills:

- Excellent interpersonal, analytical, computer & software skills
- Experience deploying models in local or cloud environments
- Ability to explain moderately complex information to others in concise manner
- Good written and presentation skills
- Intellectual curiosity, entrepreneurial drive, innovative thinking and problem-solving skills
- Willingness and ability to travel, as necessary

Interview Questions:

Can you give me an example of a failed data science project? How did you recover? What were the keys learning?

How do you approach a sparse dataset with dense data?

How do you handle missing data or outliers with step changes?

How do you correlate data sets with delays, a process with different time shifts? How do you align the data?

Give an example of how you have written code that was eventually used in production?

Give examples of your code optimization process?

Give examples of machine learning problems that you have worked on?

What libraries, what languages, how did you measure success

Can you use a laymen term explain Bayesian network?

What visualization tools have you used?

Give an example where you have used feature extraction for analysis?

White Board Question:

Do we want highly technical or do we want architecture?

Deploying a model to a manufacturing environment

Start with the known model and work through the deployment,

Keys to success

- How to validate the model
- Validate that the control variables are responding based on the models predictions
- When would retrain
- How do you maintain data integrity
- What happens when the model breaks

How to retrieve time from a JSON file