

Delhivery

Associate Data Scientist

Description

If you want to use Artificial Intelligence to solve the most pressing problems in the Indian logistics industry, or optimize the distribution network of India's largest third party logistics provider, our data team is the place to be. We are the country's premium data science team featured widely in press and news; see [Factordaily](#), [LeMonde](#), [Economic Times 1](#), [Economic Times 2](#).

The team comprises of leading experts in the field of Operations Research and Particle Physics, along with budding engineers and scientists in the field of Machine Learning and Artificial Intelligence.

In the recent past, the data science team at Delhivery has built key capabilities in the following areas:

• **Maps:**

Our business relies heavily on understanding where our end customers live and how to get there. This is a huge India specific challenge, primarily because of the unstandardized way of writing addresses and poor quality of road/pincode data in smaller towns/villages. We process a large amount of location data from our devices on the ground to learn locality boundaries, routes taken by ground staff, etc.

• **Machine Learning:**

We use a range of machine learning techniques to automate decision making at the ground level, e.g., identifying whether a shipment is safe to fly based on its product description, identifying which shipments have a high probability of being returned, etc.

• **Discrete Optimization:**

We rely on a range of optimization methods to ensure our distribution network is designed for cost efficiencies and scale, e.g., Vehicle Routing Problem to optimize the shipment collection process from clients, Facility Location Problem to ensure our distribution centres are suitably located, etc.

• **Simulation:**

The scale of our distribution network often makes many problems intractable due to the existence of millions of variables and how they interact with each other over time. We are investing to build an in-house Simulator to enable us to measure the impact of changes in the network over time and ultimately design a system that can be "self aware".

We are looking to expand our team to further our capabilities in these areas.

Responsibilities:

- Contribute to the development/ deployment of machine learning algorithms, operational research, semantic analysis, and statistical methods for finding structure in large data sets
- Use advanced statistics and machine learning on large scale multidimensional data and generate actionable insights that will be leveraged to drive operations and develop strategies for Delhivery.
- Breakdown business problems by working closely with Product Team and Leadership Management Team, analyze requirements and look into data to recommend products and solutions
- Optimize Delhivery operations by applying data science in solving complex business problems
- Collaborate with cross-functional teams including but not limited to Logistics, BD, Sales, Marketing, Security, Customer Service, etc.
- Disseminate original research in peer reviewed journals and conferences

Qualifications

A suitable candidate should satisfy as many of the following requirements:

- B.Tech/PhD/M.Tech/M.S in Computer Science, Operational Research, Statistics, Applied Mathematics or Natural Sciences with a very clear understanding of probability and statistics, analytical approach to problem solving, and capability to think critically on a diverse array of problems. Publications in peer-reviewed

journals will count in your favour.

- Supervised Machine Learning Algorithms: Logistic Regression, Bayesian Approach, Decision Trees, Support Vector Machines, Neural Networks, Ensemble Methods, Feature selection techniques etc. Understanding of advanced algorithms (i.e. Deep Learning) will be good to have.
- Good understanding of Unsupervised and Semi-Supervised Machine Learning Algorithms.
- Optimization Algorithms: Mathematical Programming (Linear/Non-linear techniques), Convex Optimization, Transportation Problem, Vehicle Routing Problem, Facility Location Problem, Queuing Theory, Inventory Management, Forecasting Techniques, etc. Knowledge about meta-heuristics like Genetic Algorithm, Tabu Search, Simulated Annealing, etc would be beneficial.
- Strong programming skills in Python or Java. Good understanding of C++ and R will be appreciated.
- Significant experience with SQL databases, especially experience in handling geographic data like PostGIS on PostgreSQL will be appreciated.
- Experience with NoSQL databases like MongoDB, Elastic Search, Redis or any graph database.
- Some experience with big data tools like Spark, Hadoop is a plus
- Most importantly, an inquisitive mind, an ability for self-learning and a risk appetite for experimentation and failure