HPE Data Scientist

Job Overview

HPE is looking for a Data Scientist who will support our Supply Chain team with insights gained from analyzing data from multiple sources and systems.

The ideal candidate will be adept at using large data sets to find opportunities for process optimization and using models to test the effectiveness of different courses of action. They must have strong experience using a variety of data mining/data analysis methods, using a variety of data tools, building and implementing models, using/creating algorithms and creating/running simulations. Candidates must be comfortable working with a wide range of stakeholders and functional teams, and will have a passion for discovering solutions hidden in large data sets and working with stakeholders to improve business outcomes. Advanced written, verbal, and presentation skills will be required.

Responsibilities for Data Scientist

- Work with stakeholders throughout the organization to identify opportunities for leveraging company data to drive business solutions.
- Mine and analyze data from various company systems to drive optimization and improvement of business processes and strategies.
- Assess the effectiveness and accuracy of new data sources and data gathering techniques.
- Develop custom data models and algorithms to apply to data sets.
- Use predictive modeling to increase and optimize customer experiences, revenue generation, and business outcomes.
- Coordinate with different functional teams to implement models and monitor outcomes.
- Develop processes and tools to monitor and analyze model performance and data accuracy.

Qualifications for Data Scientist

- Strong problem solving skills with an emphasis on data discovery and business process correlations.
- Experience using statistical computer languages (R, Python, SQL, etc.) to manipulate data and draw insights from large data sets.
- Experience working with and creating data architectures.
• Knowledge of a variety of machine learning techniques (clustering, decision tree learning, artificial neural networks, etc.) and their real-world advantages/drawbacks.

• Knowledge of advanced statistical techniques and concepts (regression, properties of distributions, statistical tests and proper usage, etc.) and experience with applications.

• Excellent written and verbal communication skills for coordinating across teams.

• A drive to learn and master new technologies and techniques.

• We’re looking for a recent college graduate who desires to demonstrate their capabilities to manipulate data sets and build statistical models, who is familiar with the following software/tools:

  • Coding knowledge and experience with several languages: C, C++, Java, JavaScript, etc.

  • Knowledge and experience in statistical and data mining techniques: GLM/Regression, Random Forest, Boosting, Trees, etc.

  • Experience querying databases and using statistical computer languages: R, Python, SQL, etc.

  • Experience creating and using advanced machine learning algorithms and statistics: regression, simulation, scenario analysis, modeling, clustering, decision trees, neural networks, etc.

  • Experience with distributed data/computing tools: Map/Reduce, Hadoop, Hive, Spark, Gurobi, MySQL, etc.

  • Experience visualizing/presenting data for stakeholders using various presentation mediums.