

Function: Program Management

Family: Enterprise Business Intelligence

Job Title: Data Scientist 5

Job Abbreviation: [To be Completed by Compensation Team]

Hierarchy Group: Professional

Hierarchy Level: Professional 5

FLSA Status: [To be completed by Compensation Team]

Summary:

Responsible for leveraging internal and external data to provide insights and information which supports a facts-based decision making process. Provides input into strategy, analysis methods, and tool selection. May work independently or as part of a team on more complex projects. Provides mentoring and guidance to more junior team members. May be responsible for leading a team, but does not directly manage people.

Core Responsibilities:

- Lead complex interdepartmental data science programs that designs solutions across one or more technologies.
- Develop data mining, machine learning, statistical and graph-based algorithms designed to analyze massive data sets for business insights and partner with the data engineering team to ensure proper implementation and usage of algorithms.
- Review and evaluate data scientist programs enterprise level to determine appropriate use of algorithm-driven products and solutions.
- Educate other departments on data science methodologies, concepts and algorithmic advancements.
- Lead a small group of less experienced team members on analytical projects or on cross-functional teams. Frequently serves as team lead on multiple projects, mentor and train junior team members.
- Lead development and implementation of scalable big-data driven solutions for accurate targeting of users with relevant business treatments and efficient algorithmic inventory. Manage challenges associated with investigating and understanding large datasets, and building models based on Big Data solutions.
- Define enterprise data strategy and data monetization processes through analysis of rich streams of unstructured data to find correlations between events and identify opportunities to optimize defined desired outcomes

Education Level: Master's degree

Field of Study: Quantitative fields such as Economics, Statistics, Mathematics, Decision Science, Operational Research, Computer Science or Engineering.

Certifications:

Years of Experience: Generally requires 6-8 years related experience.

Skills:

- PhD preferred
- Advanced level proficiency with statistical probabilistic modeling techniques such as regression, decision trees, neural networks, support vector machines, supervised/unsupervised clustering techniques, etc.
- Advanced skills in developing statistical targeting models using at least 2 of the following tools; SAS, R, KNIME, SPSS, Python, Rapid Miner, KXEN, Bayesia, MATLAB, Statistica, Weka etc.
- Expert working within enterprise data warehouse environments platforms (Teradata, Netezza, Oracle, etc.) and working within distributed computing platforms such as Hadoop and associated technologies such as SQL, HQL, MapReduce, Spark (MLib, SQL, R Py), Storm, Yarn, Kafka, Sqoop and Hive
- Expert in at least 1 scripting and/or programming language such as Scala, Julia, C#, Python, Perl, Java, C++
- Can look at the big picture and handle multiple projects at one time. Has good listening skills and demonstrates flexibility
- Ability to explain complex statistical problems and solutions to laymen.
- Has a good understanding of overall business, including financial acumen, ability to convert complex data into insights and action plans, demonstrated in-depth understanding of predictive modeling life cycle and architects projects through implementation

Compliance: Bluewhale is an EEO/AA/Drug Free Workplace.

Disclaimer: The above information has been designed to indicate the general nature and level of work performed by employees in this role. It is not designed to contain or be interpreted as a comprehensive inventory of all duties, responsibilities and qualifications.