**Job Description**

**Role/Job Title: Senior Data Scientist**

**Business: Data & Analytics**

**Function/ Department: New Frontiers**

**Place of work: Mumbai**

**Skills & Experience Required**

* Strong understanding of advanced data mining techniques, curating, processing, and transforming data to produce sound datasets.
* Strong understanding of the Machine Learning lifecycle - feature engineering, training, validation, scaling, deployment, scoring, monitoring, and feedback loop.
* Experience in analysing complex problems and translating it into an analytical approach.
* Experience in Supervised and Unsupervised Machine Learning including Classification, Forecasting, Anomaly Detection, Pattern Detection, Text Mining, using variety of techniques such as Decision trees, Time Series Analysis, Bagging and Boosting algorithms, Neural Networks, Deep Learning.
* Experience with analytical programming languages, tools, and libraries (Python ecosystem preferred, but R will be considered).
* Has used frameworks/libraries such as Scikit-learn, PyTorch, Keras, NLTK
* Deep understanding of data structures and algorithms
* Excellent programming skills in Python and PySpark
* Hands-on experience using Spark (in Python) - to develop high volume ETL pipelines.
* Have implemented machine learning projects from initiation through completion with particular focus on automated deployment and ensuring optimized performance.
* Have maintained and optimized the machine learning/deep learning models developed by the Data Scientist and ensured seamless deployment in different environments while enabling model tracking, model experimentation and model automation.
* Experience in SQL and relational databases, Big Data technologies e.g., Spark/Hadoop and Cloud technologies.
* Hands-on experience with workflow management tools such as Airflow
* Strong understanding of RESTful APIs and / or data streaming a big plus
* Required experience of modern version control (GitHub, Bitbucket)
* Hands-on experience with containerization (Docker, Kubernetes, etc.).

**Roles & Responsibilities**

* Perform Data cleansing activities such as data cleaning, merging and enrichment etc.
* Perform feature engineering through extracting meaningful features from measured and/or derived data
* Perform exploratory and targeted data analyses to get key insights
* Build Stochastic and Machine learning algorithms that potentially address business problems
* Lead and implement Machine Learning projects from initiation through completion with particular focus on automated deployment and ensuring optimized performance.
* Maintain and optimize the machine learning/deep learning models developed by the Data Scientist and ensure seamless deployment in different environments (Dev/QA/Prod) while enabling model tracking, model experimentation and model automation
* Collaborate with the data engineers and data scientists on model development to containerize and build out the deployment pipelines for new models
* Collaborate on MLOPS life cycle & experience with MLOPS workflows & traceability and versioning of datasets
* Ensure traceability and versioning of datasets, models & evaluation pipelines
* Design, prototype, build and maintain APIs for consumption of machine learning models at scale.
* Facilitate the development and deployment of POC machine learning systems
* Using standard methodologies framework to ensure data quality and reconciliation checks are in place and are transparent to everyone

**Professional Qualifications:**

* At least 5+ years professional experience as a data scientist or ML Engineer

**Educational Qualifications:**

Advanced degree in an analytical field (e.g., Data Science, Computer Science, Engineering, Applied Mathematics, Statistics, Data Analysis, Operations Research, Artificial Intelligence, or Applied Math)

**Minimum Experience: 5 years**

**Maximum Experience: 8 years**