Citi - AVP, Data Scientist (via WayUp)

**Posting ID:** EM1861825E

**Company:** Citi

**Position Type:** Full-Time

**College Major(s):** Computer Science (CS)

**College Level(s):** Undergraduate-Freshman, Undergraduate-Sophomore, Undergraduate-Junior, Undergraduate-Senior

**Company Website:** https://jobs.citi.com/search-jobs

**Work Location:** New York, NY

**Salary:** DOE

**OVERVIEW**

Citi enables growth and progress for individuals, communities, institutions and governments around the world and has been doing so for more than 200 years. They recognize that their ability to best serve their clients and customers comes from the diverse and engaged employee base and they work tirelessly to deliver opportunities for growth and development to the most talented people. This is a place where the best come to get better.

**Roles and Responsibilities**

Citi Compliance Technology Services is looking for additions to their team to advance strategic innovative programs that are based on machine learning and behavior analytic principles with a goal of replacing legacy rules-based detection approaches in the areas of Global E-Communication and Voice Surveillance/Review. The program needs to onboard data scientists who will develop machine learning models and algorithmic solutions that will identify and rank the bank’s actors and other entities with respect to their resemblance to certain risky or undesired behaviors.

As a data scientist you would have the opportunity to take a significant part in driving data-intelligence projects with high business impact and visibility within Citi organizations. The key expectation of the candidates is that they will be able to deliver comprehensive solutions using machine learning systems that analyze huge volumes of electronic communication such as emails and chats.

**Education and Qualifications**

- The ideal candidate for this Innovation Program should possess the following knowledge and skill-sets.
- Ability to build risk or forecasting models with a hands-on appreciation of data complexity, volume, and domain (Structured and Unstructured) such as that found in large financial and banking institutions.
Self sufficient in generating machine learning models with a custom tool set, such as creating their own java code, leveraging open source libraries or academic papers, or using Matlab
- Familiarity with Hadoop (preferably Cloudera) and its tools, such as Spark, Impala, UDF for Hive etc.
- Be capable to match Data labels to the participant information and any derived behavioral features, as these will be sourced from different systems without clear match keys
- Be inquisitive and research-oriented, and try several machine learning approaches for this project, such as decision trees, factorization models, and deep learning at a minimum, and other techniques as desired
- Define Key Performance Indicators to measure model accuracies and project success based on business goals. Be accountable to the overall quality of models developed
- Must be able to write everything up, including detailed comparisons of model performance against the success criteria, and functional requirement documents that will be used to scale this project to a global firm-wide implementation
- Have excellent communication skills and be comfortable working in a large matrix environment
- Be prepared to work fully within Citi’s infrastructure due to data privacy and security concerns and could not export the data or any documentation to an external cloud service or their own personal devices.
- Must to be accessible during regular business hours for meetings, workshops, and collaboration; even though working from University site and remote is expected.
- Bachelors Degree in Computer Science or a related field of study

How to Apply