**Data Scientist Prospect**

Please, can tell who you are.

My name is Abdulrahman Yusuf, a senior data scientist, and data analyst at Arewa Data Science Software Company. I have stated working with them last two years as an assistant data scientist, with at least five years of experience in the field.

What can you say about the data science profession?

Data science is one of the most lucrative professional jobs these days and has secured the highest pay in America in the last three consecutive years. The profession appears to be very attractive. However, data scientists need to be versatile, meaning they should have basic knowledge from various backgrounds, such as statistics, mathematics, computer technology, database technology, computer science, and data science.

Can you describe your job as a scientist at the company

As mentioned earlier, data scientists should have broader knowledge that involve different background. It is well known that AI, ML, and DL require organized and well-fitted data to perform well. How those data are collected and prepared to be trainable and testable is part of data scientiests’ duties. First, I need to know or have a business question to answer; before I indulge in data organization.

In trying to develop a ML system inorder to answer a business question, the knowledge of mathematics and statistics is explored and applied. A data scientiests, should have basic skills to clearly understand the types of data and the relationships among its variables. Thereafter, different models or algorithms from mathematics and statistics will be looked into with the business question in mind while the data availabilty is conssidered too. Some of the models are Bayes, regression, nearest neighbor, Markov chain, backpropagation neural network, and many others. This is to say, a data scientist without basic knowledge of mathematics and statistics is nearly an empty cage. Is from those models you can find best fit and build.

The application of ethical issues should be put into practice to make sure the data is selected and handled in such a way that it does not result in negative harm to the group of people involved or the data as well. This is what is called It is my responsibility to make a thorough assessment for identifying harms and affected groups for delineation. Harms related to allocation, quality of service, stereotyping, denigration, and over and underrepresentation should be thoroughly assessed and mitigated appropriately using some tools like Fairlearn.

Assuming the model is built after all those formalities. The next action is to train the model and later evaluate it. In the model training, this is when I should come up with model-fit, which is to send in the feature variable as an array of values (usually 'X') and a target variable (usually 'y'). The model is further evaluated. This ensures that its quality has not degraded by avoiding under- and over-fitting. Finally, the prediction is carried out using completely new data to ensure model accuracy.

It is obvious, there are several common questions to answer in evaluating data scientists. Many are based on how they carry out their duties and how they deal with some real world problems that may arise in a particular case. So, always expect the unexpected in an interview as a data scientist.