Lachlan Sinclair

OSU CS 419

Assignment 2 Reflections


A)

IBM Watson Studio Cloud is a reasonable application for performing data analytics and machine learning on arbitrarily large datasets. Its designed to integrate with a multitude of cloud storage systems and provides data management and shaping capabilities. And unlike SPSS it doesn't need to load all the data into memory, it has systems that are built to handle arbitrarily sized datasets. However, one drawback would be the cost it requires to run a system on this application. I suspect continually using this application to perform data analytics and machine learning on large datasets would cost a lot of money.


B)

Benefits:

One of IBM Watson Studio Cloud's biggest benefits for a real data science project is that it has deployment support. This makes implementing a solution developed on this application practical to use in a real scenario. It also has features that allow for data management and data shaping. It also organizes work in to projects similar to other coding platforms. Also, to my surprise it integrates with other cloud services like AWS which shows IBM is willing to relinquish some control of features and projects developed on this platform.

Challenges:

To use this in the real world you would likely need to purchase the full IBM Watson service which means you are beholden to IBM and will need to pay for their cloud space and other features continually.

Efficiently using CUH would become a challenge, all development time spent creating models will be eating up resources which creates a more stressful development lifecycle.

New features would probably need to be requested through IBM which would take time and may never be developed if it isn't deemed worth their developer's time. I don't think other standard users are able to produce their own features, however I wasn't able to find much information on this.