Advances in Big Data Research
Lessons Learned and How to Replicate Success

Overview

Almost every pharmaceutical company uses Big Data and data science in some capacity to unlock the value of its rich data sets. Combining structured and unstructured data in a data lake for a 360-degree view of compounds, diseases, treatments, etc. is one key to generating new and valuable insights.

But creating the 360-degree view is just the first step to realizing the full value of this data. The data in the data lake must also be made accessible for analysis by researchers using industry-leading tools.

The “Journey of Big Data in Research” is at a crossroads. Data lakes are becoming more commonplace. They reduce the cost and reliance on the traditional data warehouses while at the same time opening up the world of unstructured data. But where do we go from here?

Pivotal packs a “One – Two Punch” with our Data Scientists (technologists at heart) working side by side with our clients to build models that run on top of Pivotal’s Big Data Suite, effectively turbo charging your analytics capabilities by creating or enabling existing models to run over the complete data set (NO sampling or aggregation) in far less time than today’s traditional approach.
BUILD HIGH-RESOLUTION MODELS, RUN INTEGRATIVE STUDIES, ITERATE RAPIDLY, INTERROGATE MODELS FOR INSIGHTS

Real Life Results
With our clients we demonstrate drastic improvement in processing time both in deriving features and building models yielding various benefits, (e.g. operating over granular data, integrating various data modalities, iterating over models.)

For example, Pivotal helped one client build analytic models by using single-cell resolution, integrating datasets (e.g. image, structural and gene expression data) and leveraging all samples in record time. This enables scientists to test more hypotheses, interrogate models to understand the underlying biology, and more. Pivotal’s Big Data Suite is capable of parallelizing data transformation and running models where data resides to avoid data movement. The models are sent to the data for parallel execution!

What does this mean? You can test millions of hypotheses to determine what signals are real. Build models in our Big Data Suite on billions of training examples in minutes. Dig into and transform trillions of rows of data to build more features.

You will be more agile, ask more questions, integrate all data, and derive more value as the data grows!

Pivotal’s Cloud Native platform drives software innovation for many of the world’s most admired brands. With millions of developers in communities around the world, Pivotal technology touches billions of users every day. After shaping the software development culture of Silicon Valley's most valuable companies for over a decade, today Pivotal leads a global technology movement transforming how the world builds software.

Pivotal, Pivotal Cloud Foundry, and Cloud Foundry are trademarks and/or registered trademarks of Pivotal Software, Inc. in the United States and/or other Countries. All other trademarks used herein are the property of their respective owners. © Copyright 2016 Pivotal Software, Inc. All rights reserved. Published in the USA. PVTL-DS-9/16