

CUSTOMER SPOTLIGHT

How TruDiagnostic trained AI and statistical models 10x faster, saving a year of development time while cutting infrastructure costs by 60%

SUMMARY :

TruDiagnostic faced an infrastructure challenge, managing the world's largest epigenetic dataset with fragmented tools. By adopting Posit Workbench with Amazon SageMaker, they empowered their scientists to control their computational resources with greater efficiency, unifying their R and Python environments. This transformation saved them a full year of product development time, cut their cloud infrastructure costs by 60%, and increased both their AI and statistical model training performance tenfold.



ABOUT :

TruDiagnostic is the leading epigenetic testing company focused on biological aging, nutritional deficiencies, and cellular systems health—all based on DNA methylation analysis. They provide testing to clinical practitioners and their patients, researchers, and consumers, with the goal of helping people live longer, healthier lives.

INDUSTRY :

Biotechnology

DATA CLOUD PARTNER INTEGRATION :

Amazon SageMaker

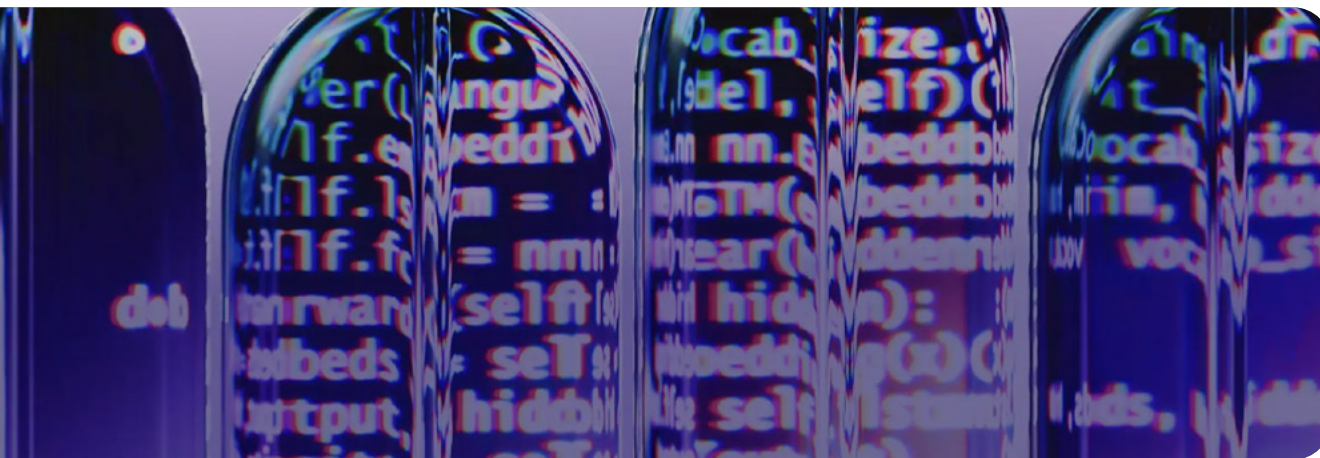
TECHNOLOGY USED :

Posit Workbench,
Amazon SageMaker,
Amazon EFS, Amazon S3,
Amazon EC2

The Challenge:

Infrastructure scaling & fragmentation

- **ENGINEERING & COMPUTE BOTTLENECKS:** Scientists were originally limited by local hardware and a single EC2 instance. They had to wait on an engineer to manually scale compute resources up or down, causing research to lag and model training (5,000 models) to take months to complete serially.
- **LANGUAGE AND TOOL SILOS:** The team was fragmented by “preferred tools”—some used Python and Jupyter for MLOps, while others used R and RStudio for bioinformatics. This lack of a unified environment prevented seamless collaboration and created “language silos” that hampered product delivery.



The Solution:

A unified, cloud-based platform

- **INTEGRATION OF POSIT WORKBENCH WITH SAGEMAKER:** By combining these tools, TruDiagnostic created a single environment that supports both R and Python. This unified platform allowed scientists to keep their familiar coding environments (like RStudio or Jupyter) while collaborating on a shared infrastructure.
- **ON-DEMAND PARALLEL PROCESSING:** The transition to a cloud-based architecture (using Kubernetes and SageMaker) enabled the team to parallelize their workloads. This shifted them from a “MacBook Air” limitation to massive cloud power, allowing them to train 1,600 model iterations in one hour—a process that previously took months.

Learn more at posit.co > | [Read the full story](#) >