



CENTRAL ML AND AI FACTORY ENABLED DATA SCIENCE TEAMS TO SHARE USE CASES ACROSS THE CLOUD

IMPACT

- Created self-servicing data and compute provisioning with CML to drive cost efficiencies
- SDX provided security and governance on premises and in the public cloud
- Data science teams can be onboarded quickly and more efficiently

The Insurance industry is a large factor in the economy with the amount of money it collects, investments it makes, and the role it plays in covering risks for individuals and companies. This large multinational insurance company in Europe has a vast portfolio of products. They serve one quarter of employees in the private sector, and are also the second biggest health insurer (excluding banking groups). The company has over 10 million customers including more than a half million business customers.

Challenge

Managing all customer data in the cloud has transformed the company, bringing ease of deployment, resourcefulness, and flexibility to the financial services and insurance sector. Finding ways to be more efficient for both customers and insurance advisors to connect and collaborate is paramount. Security and governance remain critical and preventing breaches is equally vital. The challenge was managing data on premises and in the public cloud while protecting the company from data breaches. There was a strong need for a central machine learning (ML) and artificial intelligence (AI) factory across data science teams to collaborate and to access use cases. In addition, onboarding data science teams took time, and the company wanted to move more efficiently and securely.

The traditional data center architecture for the company was limiting data science usage; it offered no graphic processing units (GPU) for asset liabilities, limited resources, and scalability. Furthermore, there were no data science industrialization methodologies. The company needed a robust and scalable ML platform for enabling more AI use cases in their business at scale to embrace a hybrid cloud strategy.

Solution

The multinational insurance company turned to the Cloudera Data Platform (CDP) running on Microsoft Azure to deliver Cloudera Machine Learning (CML). By implementing this new platform they created self-service data and compute provisioning with CML. Teams now had the same user experience between on premises and in the cloud enabling consistent and secure experiences no matter when and where data scientists need to work. This approach not only provided them with the speed and agility that comes with CDP Public Cloud, but also helped drive cost efficiencies through variable compute capabilities.

Machine learning model training often uses lots of compute power for a short space of time, but the development process requires continuous access to compute resources for the data science teams. Having access to resources on premises and in the cloud with a consistent user experience lets the company choose the environment that makes the most economic sense for what they are doing, but does not impose any limitation on how big the environment can be.

About Cloudera

At Cloudera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cloudera delivers an enterprise data cloud for any data, anywhere, from the Edge to AI. Powered by the relentless innovation of the open source community, Cloudera advances digital transformation for the world's largest enterprises.
Learn more at cloudera.com

Cloudera's machine learning operations (MLOps) capabilities added significant value through monitoring, serving, and governance for model deployment. Through this they were able to automate the deployment, monitoring, and management of ML models into production in a scalable and governed way providing measurable business value. The foundation allowed data development as well as production teams to work collaboratively and leverage automation for monitoring, and govern machine learning services within the company.

Now with Cloudera, they have the flexibility with CML to deliver auto-scaling elastic resources. The agility allowed the company to plan the design, develop and test before deployment. Shared Data Experience (SDX) provided security and data governance both on premises and in the public cloud.

Results

Previously, the insurance company had limitations with workloads and was heavily dependent on their IT infrastructure team to build new environments. Now with CML, the team can leverage infinite, diverse computing resources; whether CPUs or GPUs, and can now scale to any type of workload. CML has the ability to autoscale based on the resources requested, both up and down, providing the compute for the workload but keeping costs down when the jobs have completed.

Now with this new modern data architecture with CDP, the company has been able to build a central AI factory across all teams. What previously took a long time for IT to provision the environment can now be done in a couple of minutes. They have a secure hybrid approach to take advantage of the cloud, all while controlling cloud usage. Having a secure machine learning production pipeline with governance and monitoring enables them to easily onboard new data scientist teams quickly and more efficiently.