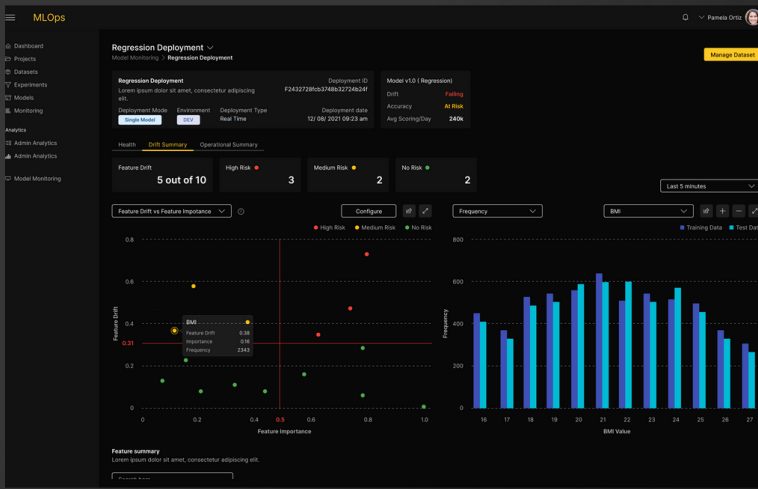


H2O.ai

Product Brief

H2O MLOps





H2O MLOps provides a collaborative environment for Data Scientists and IT teams to manage, deploy, govern, and monitor machine learning models.

Why H2O MLOps?

As enterprises mature their artificial intelligence initiatives, a new set of challenges emerge. Deploying a growing number of models in a highly available and scalable infrastructure is a challenge that many businesses face. Collaboration between teams is necessary for enterprise-wide data science efforts that aim to build models at scale. Additionally, model monitoring is needed to ensure that drift or performance degradation is addressed throughout a model's lifecycle. Finally, maintaining reproducibility, traceability, explainability, and verifiability of machine learning models allows enterprises to meet compliance standards.

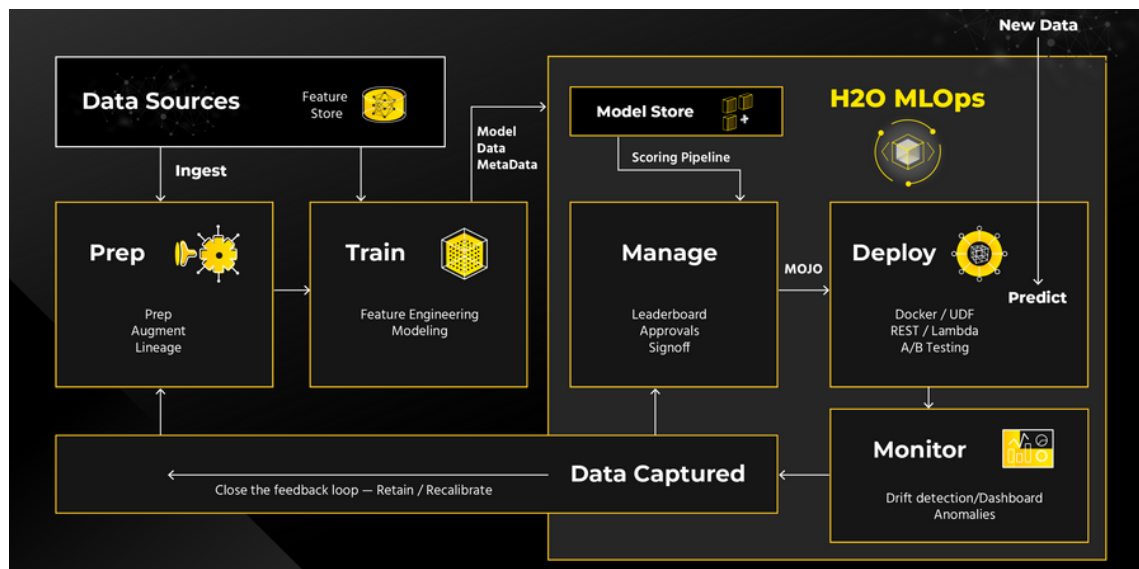
The Data Science Workflow

The data science workflow requires Data Scientists to iterate and collaborate on a number of steps that include training and optimizing models. Once the models have been trained and tuned, the next step is to deploy ML models into production.

Managing the model lifecycle after its development is quintessential in delivering continuous business value. This is where H2O MLOps comes in.

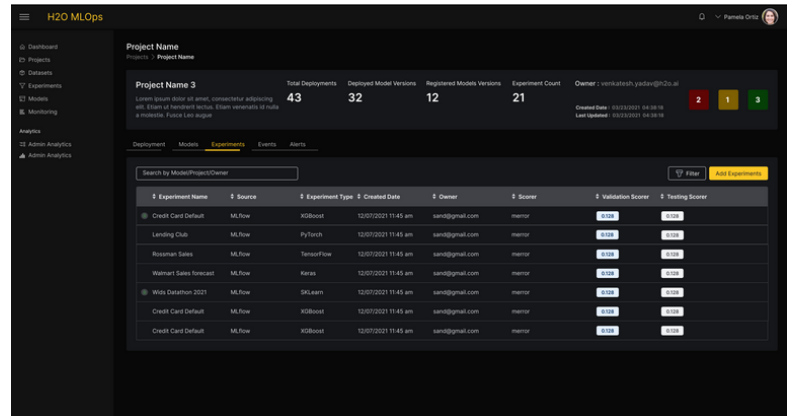
MLOps consists of four major modules:

- 1) Model Management
- 2) Model Deployment
- 3) Model Governance
- 4) Model Monitoring



Model Management

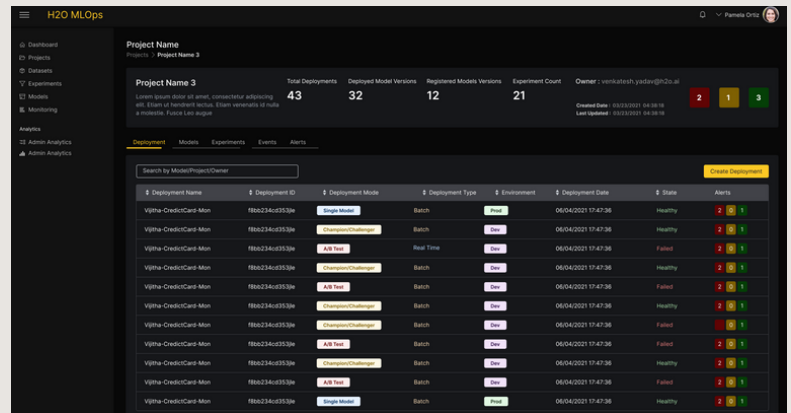
Data Scientists often work in teams to build and train models for a particular use case. Data Scientists can choose across a wide variety of ML frameworks for optimal experimentation. H2O MLOps allows Data Science teams to collaborate and maintain a central repository of models, irrespective of the ML framework used to train it. Teams are able to compare experiments against each other using a Leaderboard, across many evaluation metrics. Teams can also view associated experiment summaries and metadata, which informs their decision to promote certain experiments. Once Data Scientists have selected their experiment of choice, they can register the experiment using Model Registry, to create a new Model Version.



Model Deployment

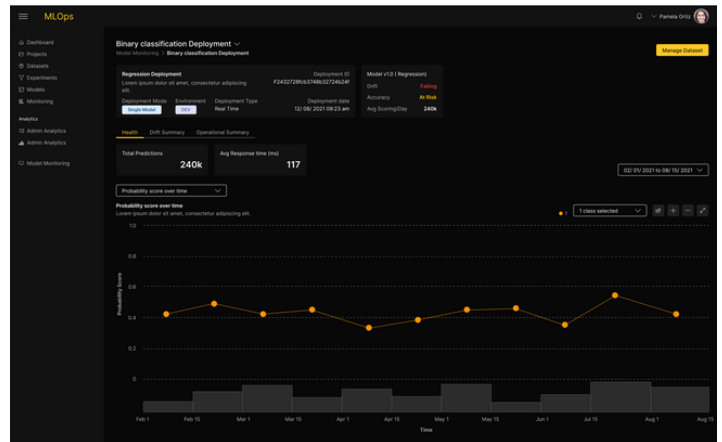
Model deployment comes with its own set of challenges and often involves a team effort across Data Scientists, infrastructure (IT), and operations experts (DevOps). This can become even more complex as teams scale the number of models and the frequency of retraining.

H2O MLOps simplifies this process, allowing for models to be deployed on a server and made available as a REST endpoint (real-time or asynchronous processing). They can also be deployed to score in batch from a data source on a regular basis. Within each of these deployment types, businesses have the ability to choose from the following deployment modes: Single Model (simple deployment), A/B Test (route partial traffic to different models), and Champion/Challenger (compare how a new challenger model performs against incumbent champion model). Models can also be deployed to specific environments, which can be configured by teams. MLOps includes a complete history of a deployment, and has the ability to roll back to a previous version, if required. All of this can be done within a few clicks of a button on H2O MLOps.



Model Governance

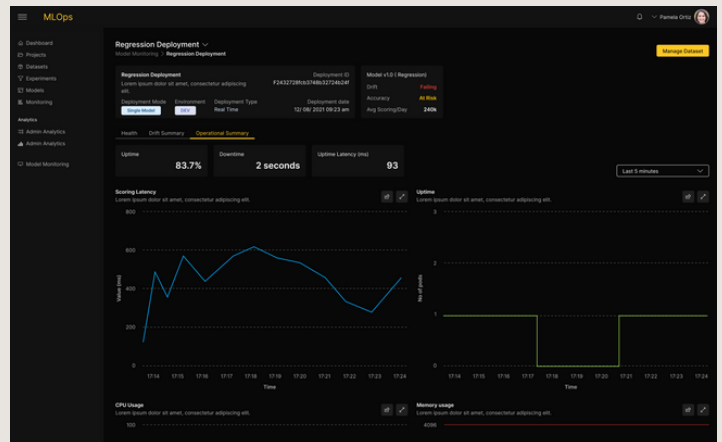
H2O MLOps stores and manages data, artifacts, experiments, models, and deployments, and draws lineage across each of them, increasing overall transparency for Data Scientists and Compliance teams. It also provides explanations and feature importance, along with scoring results, to make regulatory compliance simpler. There are also permissions at a User and Group level, ensuring only the authorized people are able to make changes to the environment. There is also an Admin Dashboard, which provides visibility of the entire organization to admin users, on users, experiments, deployments, and audit logs.



Model Monitoring

In general, model performance is known to degrade over time. Businesses wishing to maximize the performance of their applications need to monitor the performance of the models to detect the optimal moment for swapping in new models. H2O MLOps includes real-time monitoring of models for detecting anomalies, drift, accuracy, and fairness degradation. IT teams are also able to monitor their operational metrics. Metrics and alerts are presented in a real-time dashboard, allowing Data Scientists to dive in deeper to gain a complete understanding.

The alerting capabilities of H2O MLOps allow teams to be notified when discoveries are made. With built-in integrations, teams are kept in the know with real-time alerts. When the threshold and anomalies are triggered alerts will be sent to the dashboard. The Data Scientist can configure automatic model retraining, in case a particular metric falls below a threshold.



MLOps for the Data Science Workflow

Today, as organizations increase their usage of machine learning models, many organizations are facing challenges that limit the impact and scale of consumption of these models because of the lack of automation of the process. H2O MLOps operationalizes, scales, and manages production deployments. H2O.ai customers can expect a full range of support, training, and expertise to assist them with their AI journey.

Request a Demo of the **H2O AI Cloud**

Request Demo

H2O.ai