Put the power of AI and data to work for your business

IBM Watson Studio: Accelerating value for AI in the enterprise

Highlights

- IBM Watson® Studio provides a rich set of tools on the cloud that simplify working with data, training and managing models, and deploying AI-powered applications

- It supports complete AI lifecycle, from curating training data to training and deploying machine learning models. It can be highly customized and has seamless integration with various Watson services.

Data may be the new natural resource, but without artificial intelligence (AI) applications, the insights and competitive advantage that can be derived from data will remain untapped. With the right tools to build AI-powered applications, including machine learning and deep learning, businesses can tackle difficult problems, ranging from predicting consumer demand and controlling costs to developing new products and even new businesses.

The key to operationalizing AI across the enterprise is a single environment that integrates the AI and modeling tools used by data scientists, developers, business analysts and line-of-business experts with the structured and unstructured data that is used to train the models.

Whether your organization is just getting started with AI or looking to expand early projects into production environments, IBM Watson Studio can provide that platform for success. A SaaS solution delivered on the IBM Cloud™, Watson Studio is an integrated environment designed to make it easy for builders who work with data to develop, train and manage models and deploy AI-powered applications.

With Watson Studio, data scientists, developers and analysts get the tools to do their jobs—drawn from across open source and IBM technologies—integrated into a single platform with a consistent experience. These tools are preconfigured, which means builders don't have to spend time installing, setting up and maintaining them. The built-in catalog function enables knowledge sharing and retention. This seamless collaboration leads to big productivity gains that save both time and money in building AI applications. On demand work and model training environments tap the scalability of cloud to provide compute resources as they are needed, with automatic deallocation when training runs are complete.

Work with a wide variety of open source data science tools

Watson Studio integrates Apache Spark, Jupyter Notebooks and RStudio®. Data scientists can use Notebooks or RStudio for their analysis within a project. Jupyter Notebooks enable data scientists to create and collaborate on Python, R and Scala programs that contain code and visualizations. Watson Studio also provides access to RStudio—a popular open source integrated development environment—that allows for fast development of R scripts.

These tools run in highly configurable environments with varying levels of compute resources based on your teams' needs. Open source tools can run in elastic single node compute environments or scale with Apache Spark powered by services from IBM Cloud or remote clusters. Watson Studio environments make it easier to work with open source data science tools by enabling easier reproducible research by using shared software configurations. These environments are also elastic, so you can easily scale to additional compute resources with a single click.

Experiment-centric deep learning and advanced modeling with neural nets

Go from machine learning to deep learning within Watson Studio, a cloud-based, experiment-centric model training environment offering a powerful set of tools and pay-as-you-use resources. It includes Neural Network Modeler, an intuitive drag-and-drop interface for designing neural architectures using the most popular deep learning frameworks: TensorFlow, Caffe, PyTorch and Keras. Users can quickly capture network designs and then export them for experimental optimization.

Experiment Assistant is a suite of tools that guide and manage the experimental process, simplifying the hundreds to thousands of training runs it takes to find the right combination of network layer configurations and hyperparameters. Each training run is automatically started, monitored and stopped upon completion. Training progress and cross-model performance can be viewed in real time and revisited later. Training history and assets are tracked, then automatically transferred to object storage.

Distributed training capabilities allow training of a single neural network across dozens of GPUs to reduce compute time from days to hours or minutes, speeding time to market with new AI solutions. Support for distributed training is provided using native TensorFlow, IBM Distributed Deep Leaning framework and Uber's Horovod.

Enrich apps with integrated Watson AI Services

Watson Studio is designed for tight integration with IBM Watson services, enabling users to create a service instance within Studio and associate projects with services to enable collaboration and use within notebooks and other tools. One such service is Watson Visual Recognition. Use this machine learning service to quickly and accurately tag, classify and train visual content. Visual Recognition analyzes images for scenes, objects, faces, colors, food, text, explicit content and other attributes that you select and train to provide insights into visual content.

Give business analysts a drag-and-drop solution for insights without coding

IBM® SPSS® Modeler, incorporated into Watson Studio, offers business analysts and data scientists an easy-to-use, interactive way to develop predictive models without the need for programming. It provides automated modeling with out-of-the-box, industry-leading algorithms as well as a range of advanced analytics, including text analytics, geospatial analysis and optimization.

Built-in data preparation and profiling with Data Refinery

Speed time to productivity with easy-to-use data prep. Before the real work of data science can begin, data must typically be cleansed, refined and enriched. In fact, it's estimated that up to 80 percent of a data scientist's time is spent on this kind of data preparation work, leaving only 20 percent for real analysis. Data Refinery, included with Watson Studio, can help data scientists and data analysts push that ratio significantly toward greater productivity.
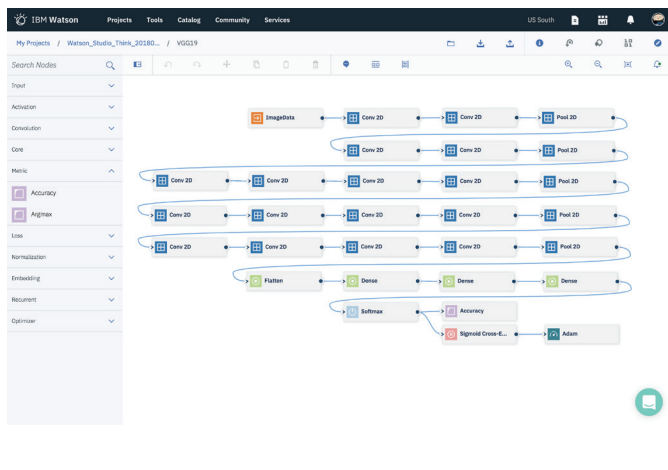
Figure 1. Drag-and-drop interface of Neural Network Modeler

It provides a powerful set of self-service capabilities that can be used to explore datasets, iteratively cleanse and refine them, and visualize the results to make sure it is what was intended. The solution also automatically tracks and documents every step in the refinement process to provide end-to-end data flow and makes it easy to save and share the output with other knowledge workers across the organization.

Data Refinery integrates with a wide range of cloud and on-premises data stores using a secure gateway, enabling users to load data from source systems in minutes. The availability of so many out-of-the-box connectors means that data scientists and other knowledge workers are no longer reliant on support from the IT team. If a suitable connector exists, and their access credentials are valid, they can set up a new connection for themselves.

Act on insights in real time with Streams Designer

Streams Designer is another integrated tool within Watson Studio, which provides a simple approach to defining the intent of an app through visual composition. Apps can then be deployed to the IBM Streaming Analytics service to analyze data in motion. This enables real-time scoring of models and analysis using Python code created in other parts of Watson Studio, allowing users to act on new insights in real time.
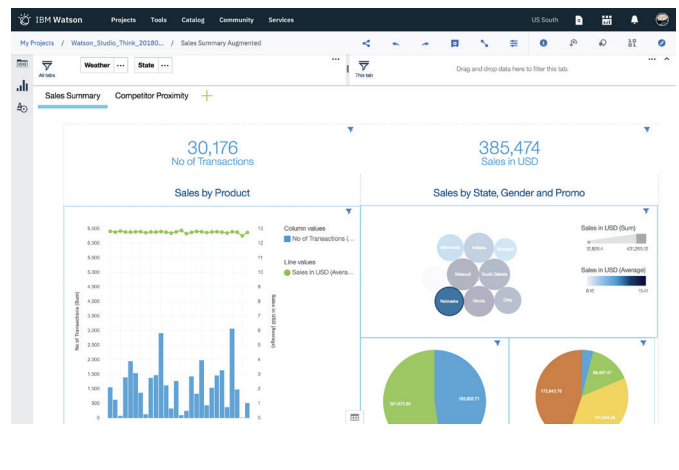

Figure 2. Visualize analytics with built-in dynamic dashboarding

Visualize the insights with integrated dashboards

Integrated analytic dashboards can enable analysts to discover insights from data and turn the analytical results into user-friendly visualizations for sharing with a broad audience of business users.

Dashboards can be created from within a project using drag-and-drop tools accessing available data assets. The system provides automatic recommendations for effective visualizations based on the type of data selected, while built-in templates and styles make it fast and easy to format reports as needed.

Deliver self-service access to data and other knowledge assets

Integrated with Watson Studio is the Watson Knowledge Catalog, which includes tools to automate and simplify data discovery, facilitate curation and provide active governance. Intelligent AI-powered search capabilities help users find the structured and unstructured data, notebooks and other knowledge assets they need, while metadata such as tags, comments, and quality metrics help them decide whether a data set will be useful to them and how best to extract value from it. Lineage of assets, including models, is automatically captured to give consumers the ability to understand where an asset came from, where it was used and what the inputs were.

Integrated active governance capabilities give users confidence that they are permitted to use a given data set while automatically masking sensitive data so they are not able to see it. This helps ensure that the assets in the catalog are used responsibly by others in the organization.

Accelerate analytics development with IBM Analytics Engine

IBM Analytics Engine is another key component of IBM Watson Studio. It is the next generation of IBM's Apache Spark and Apache Hadoop cloud-based service that enables data scientists to rapidly provision, manage, run and retire Apache Hadoop and Apache Spark clusters. It increases flexibility by keeping compute and storage infrastructure separate, so each can scale independently to prevent loss of data if a compute cluster fails. The data is stored in IBM's Cloud Object Storage service, and the Hadoop and Spark clusters connect to the object storage repository when they need to access it. It simplifies the analytics infrastructure and streamlines workflow.

Choose the right pricing plan for your organization

Whether you are an individual user looking to learn about data science, an AI professional working alone or with a small team, or an enterprise with department-level AI programs, IBM has a SaaS plan sized to meet your needs. They include on-demand, scalable compute and pay-as-you-use GPU resources for model training.

For more information

To learn more about how IBM Watson Studio can help you develop new business solutions based on AI and deep learning, contact your IBM Business Partner or visit our website: **ibm.com**/cloud/watson-studio