Watson Machine Learning is now Generally Available

Today we are excited to announce the general availability of the IBM Watson Machine Learning service. Over the past 12 months we've got feedback from hundreds of beta users of the [Watson Machine Learning](#) (WML) service. During the beta period, we've been actively collecting feedback provided via email, Slack, and targeted surveys. The WML product team has been actively engaged in those conversations and wherever possible we've worked to incorporate your feedback in to the service. With today's announcement, we are now opening this service to the general public and rolling out a number of new features. Read on to learn more…

What is WML and why are we building it?

WML is a Bluemix service that enables users to perform two fundamental operations of machine learning.

- **Training**: this is the process of refining an algorithm so that it can 'learn' from a dataset. The output of this operation is called a model. A model encompasses the learned coefficients of mathematical expressions.

- **Scoring**: the operation of predicting an outcome using a trained model. The output of the scoring operation is another dataset containing predicted values.

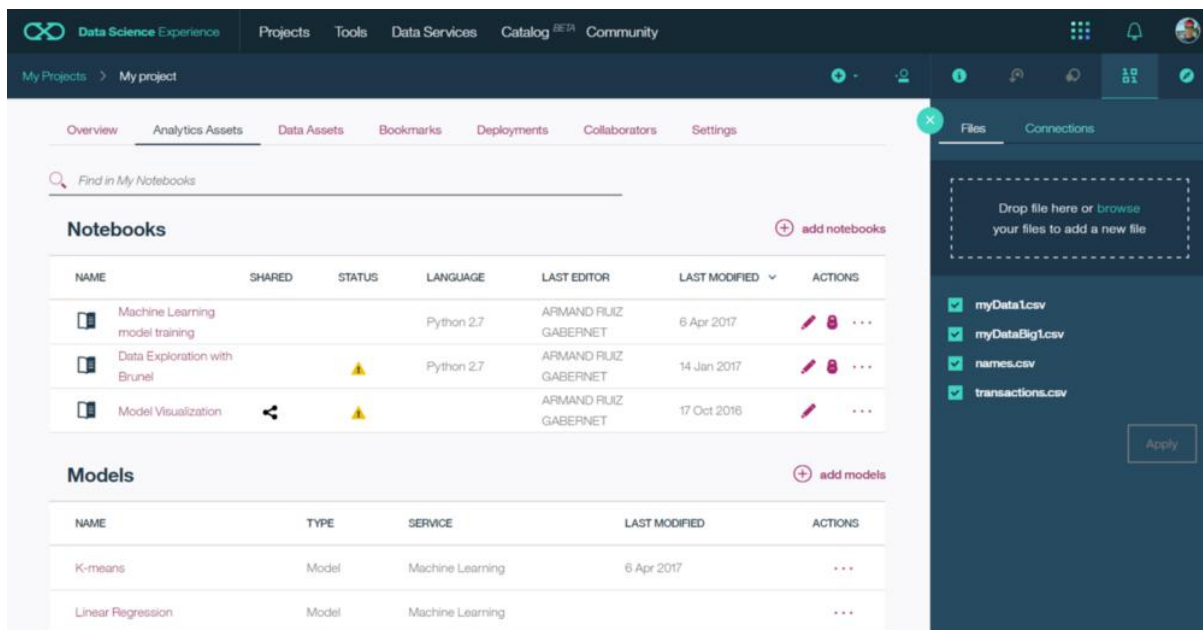WML is designed to address the needs of two primary personas:

- **Data Scientists**: create machine learning pipelines that leverage data transformations and machine learning algorithms. They typically use notebooks or external tooling to train and evaluate their models. Data scientists often collaborate with Data engineers to explore and understand the data.

- **Developers**: build intelligent applications that leverage the predictions output by machine learning models.

Although training is a critical step in the machine learning process, it is not the primary value proposition of the WML service. Over the past decade there have been an explosion of open source projects that enable the training of various types of models. Data scientists already have a significant set of quality tools they can use in order to perform training operations. The real challenge facing data scientists is how to operationalize those models. How can data scientists deploy models in production and derive actual business value? Once those models are in production, how do they adapt and evolve over time? These are the challenges that WML is intended to address.

What's new in the service?

Here are some highlights of the new features:

- **Models as First-Class Entities**: In Watson Data Platform and [Data Science Experience](#)we've made "Models" a first-class entity. Models are now associated with "Projects". By associating Models with Projects, users are able to easily share those models and collaborate with others. Over time, we intend to add additional collaborative features to Models including comments, version control, and the ability to import Models created outside of DSX such as those conforming to the PMML interchange format. The deployment of Models created in the IBM SPSS offering are already supported.

Data Science Experience · Projects · Tools · Data Services · Catalog BETA · Community

My Projects > My project

Overview · Analytics Assets · Data Assets · Bookmarks · Deployments · Collaborators · Settings

Files · Connections

Find in My Notebooks

Notebooks

(+) add notebooks

NAME	SHARED	STATUS	LANGUAGE	LAST EDITOR	LAST MODIFIED	ACTIONS
Machine Learning model training			Python 2.7	ARMAND RUIZ GABERNET	6 Apr 2017	✏ 🔒 …
Data Exploration with Brunel		⚠	Python 2.7	ARMAND RUIZ GABERNET	14 Jan 2017	✏ 🔒 …
Model Visualization	◁	⚠		ARMAND RUIZ GABERNET	17 Oct 2016	✏ …

Drop file here or browse your files to add a new file

☑ myData1.csv
☑ myDataBig1.csv
☑ names.csv
☑ transactions.csv

Apply

Models

(+) add models

NAME	TYPE	SERVICE	LAST MODIFIED	ACTIONS
K-means	Model	Machine Learning	6 Apr 2017	…
Linear Regression	Model	Machine Learning		…

- **Model Builder**: During the beta we introduced a new user interface called "Model Builder". The intent of this user interface was to simplify the creation of Machine Learning models using a more intuitive visual builder experience. Feedback from our Beta participants, it quickly became apparent that the interface simply was not intuitive enough and therefore not very easy to use. Based on this feedback, our design team went back to the drawing board and came back with a new design for a simpler version of the flow that presents the user with two options in the Model Builder: Automatic and Manual. The Automatic Path will automatically prepare data for training and will present the user with recommendations on the algorithm and technique to use based on the characteristics of the data. In this automatic path, the user quickly prepare the data, train a model and deploy that model in only a few clicks. Gartner recently shared a report stating that by 2020 more than 40% of Data Science Tasks will be automated. Leveraging work from IBM Research, this automatic path in the Model Builder is our first step in that direction.

Automatic
Prepare my data and create a model automatically

Manual
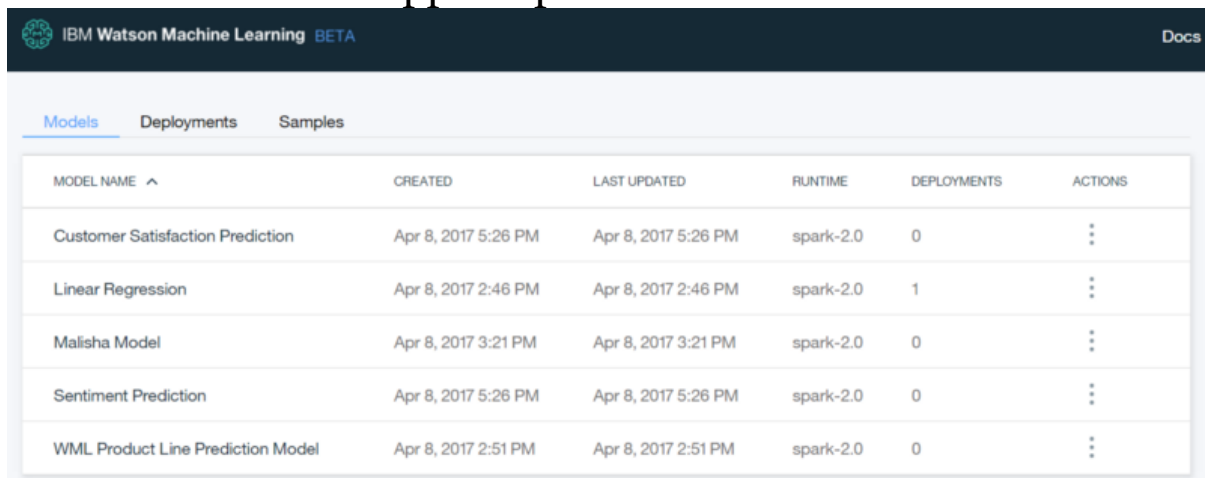Let me prepare my data and select which models to train

- **Notebook experience**: data scientists love notebooks and are already training models in Scala, R and Python using that interface. In order to accommodate these users, we are making the WML APIs available within Jupyter notebooks in the IBM Data Science Experience. Now it is possible to cover the end-to-end flow without leaving the notebook! Train the model, save the model to a project and deploy that model simply by calling our intuitive APIs!

- **Associate Watson Machine Learning service with Projects**: a recent article in Forbes about the best practices for collaboration between data scientists stated that one of those practices is to share the computational environment. For this reason, we've enabled users in DSX to not only share data and analytics assets within the project, but also to share underlying services such as Spark. Data scientists waste huge amounts of time because computational environments aren't currently shared by default. We believe that data scientists shouldn't need to make sure that they are running the same version of a Python library as their colleague in order to execute the same code so we're making it possible to share the underlying service in the project.

Associate Services
➕ add associate service ⌄

NAME	SERVICE TYPE	PLAN	ACTIONS
Spark-DEMO	Spark	Personal	...
ML-DEMO	Machine Learning		...

- **Collaboration between the App Developer and Data Scientist**: Watson Data Platform enables collaboration between different personas: data scientists, app developers, data engineers and analysts. App developers will now have access in the Bluemix Dashboard to all of the models created in DSX, as well as the ability to create to easily integrate the ML APIs in to their applications. For app developers who are

new to the WML service, we've made available a number of resources to help them get started, including 3 sample ML models and new app templates.

MODEL NAME ∧	CREATED	LAST UPDATED	RUNTIME	DEPLOYMENTS	ACTIONS
Customer Satisfaction Prediction	Apr 8, 2017 5:26 PM	Apr 8, 2017 5:26 PM	spark-2.0	0	⋮
Linear Regression	Apr 8, 2017 2:46 PM	Apr 8, 2017 2:46 PM	spark-2.0	1	⋮
Malisha Model	Apr 8, 2017 3:21 PM	Apr 8, 2017 3:21 PM	spark-2.0	0	⋮
Sentiment Prediction	Apr 8, 2017 5:26 PM	Apr 8, 2017 5:26 PM	spark-2.0	0	⋮
WML Product Line Prediction Model	Apr 8, 2017 2:51 PM	Apr 8, 2017 2:51 PM	spark-2.0	0	⋮

- **New WML APIs**: WML provides a powerful set of REST APIs that can be called from any programming language. These APIs are fully documented here: http://watson-ml-api.mybluemix.net/

- **Understand your Models**: One of the hardest things in Machine Learning is to understand how a trained model will perform in production. With the Model Visualization component you can immediately see which features have the most impact on your predictions with variable importance. The automated model visualization provides very detailed statistics on the model as well as additional information provided in the PMML and StatXML generated by the model. Using this data, our Model Output team created an interactive model visualization tool that can be invoked with one single line of code
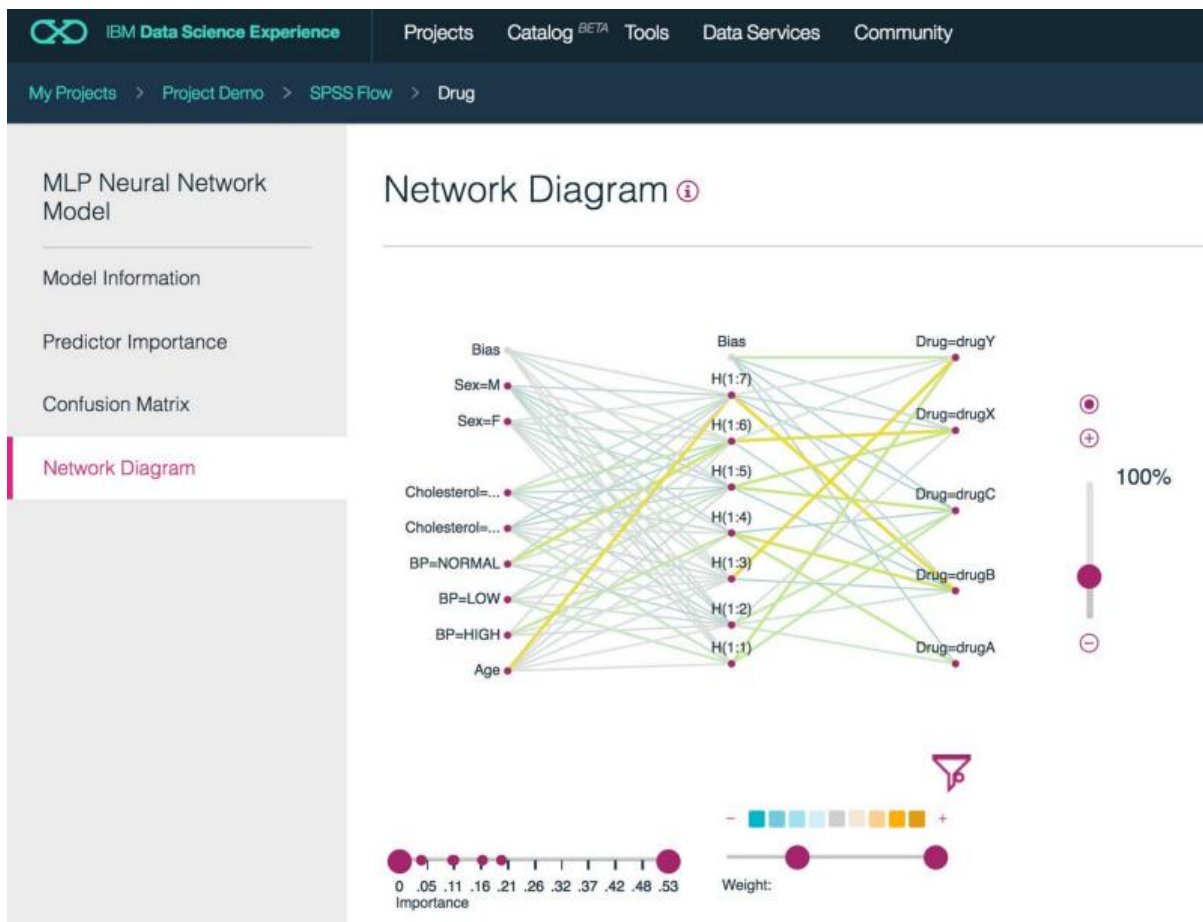
- **Performance**: the product team has really focused on improving the performance of the service. The goal from the start has been to provide a scalable, secure and high performing model training and deployment service.

- **Visual Modeling**: this interface is easy for everyone to learn and use—from business users to data scientists. Uncover valuable insights quickly for rapid time-to-value. Then, deploy your machine learning models into production to create intelligent applications. Ready to see how it works? Watch the following video to see how to bring data in, clean that data up, and create a Neural Network, all in a matter of seconds.

What is next?

- We are just getting started, watch out for more exciting updates this year. We are working hard on a new set of

Deep Learning capabilities and we will soon invite Close Beta users. To subscribe to the waitlist please fill this survey: Join the waitlist

- For more information about IBM Watson Machine Learning you can find here the announcement blog post in World of Watson 2016: http://datascience.ibm.com/blog/machine-learning-for-everyone/