Watson Machine Learning Overview

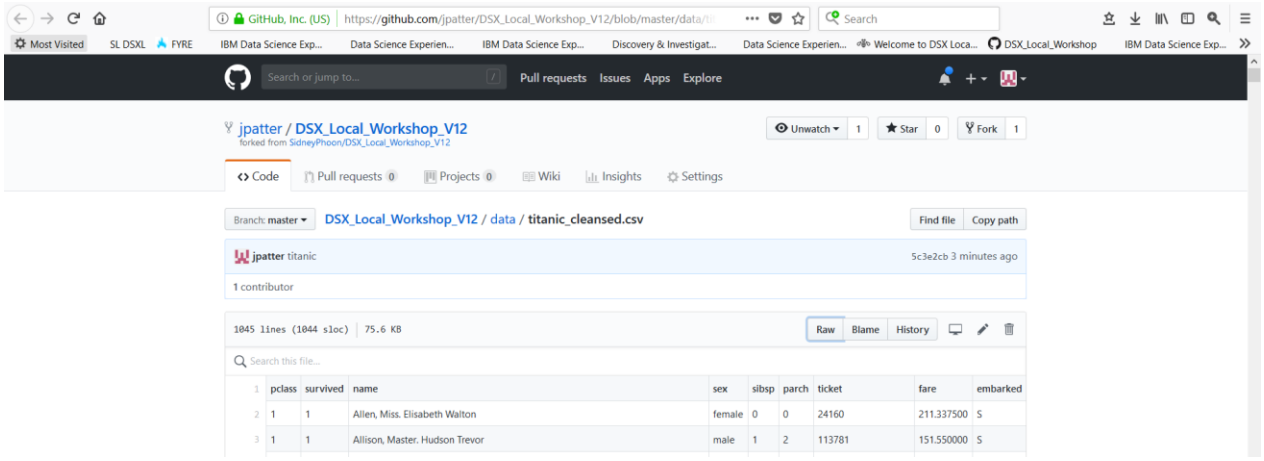
This lab will introduce the Watson Machine Learning capability using the Titanic dataset. The lab will consist of the following steps:

1. Adding a data asset to the DSXL project
2. Creating a Model to predict whether a person would survive
3. Testing the Model

Step 1: Adding a Data Asset to the project

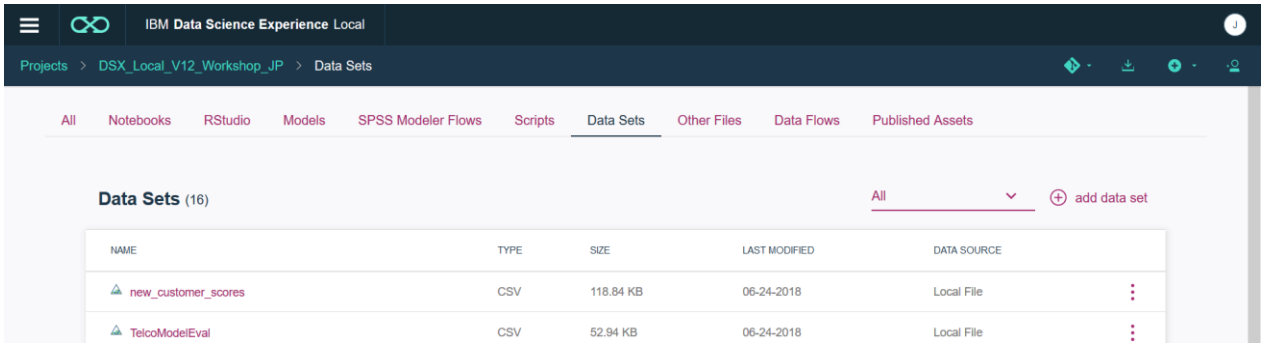
1. Download the Titanic data file from the following location by clicking on the link Cleansed Titanic Data Set and following the instructions below.

2. Right-click on **Raw** and select **Save link as…**

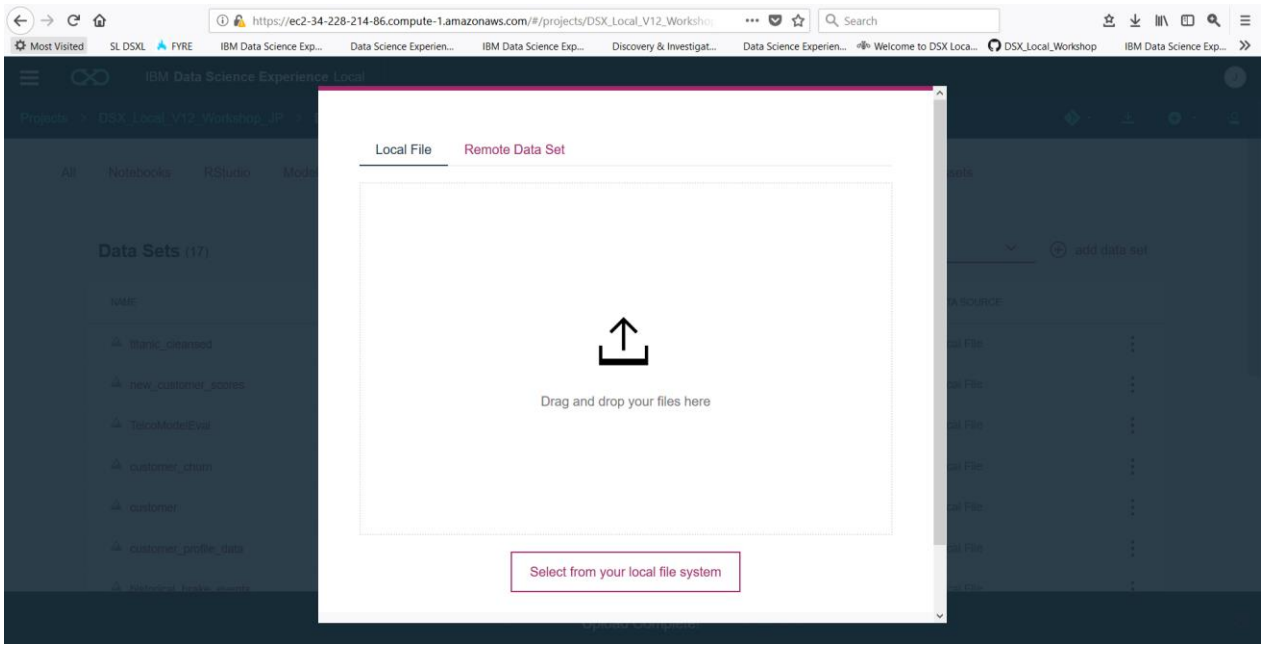


3. Save the file in your local filesystem

4. In your DSXL project go to **Data Sets** and select **add data set**

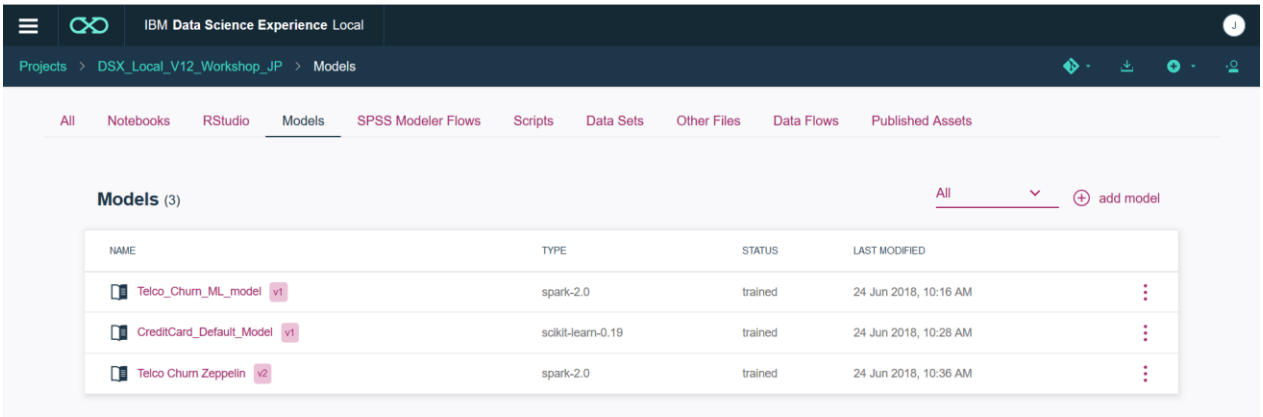


5. **Browse** or drag the **titanic_cleansed.csv** file

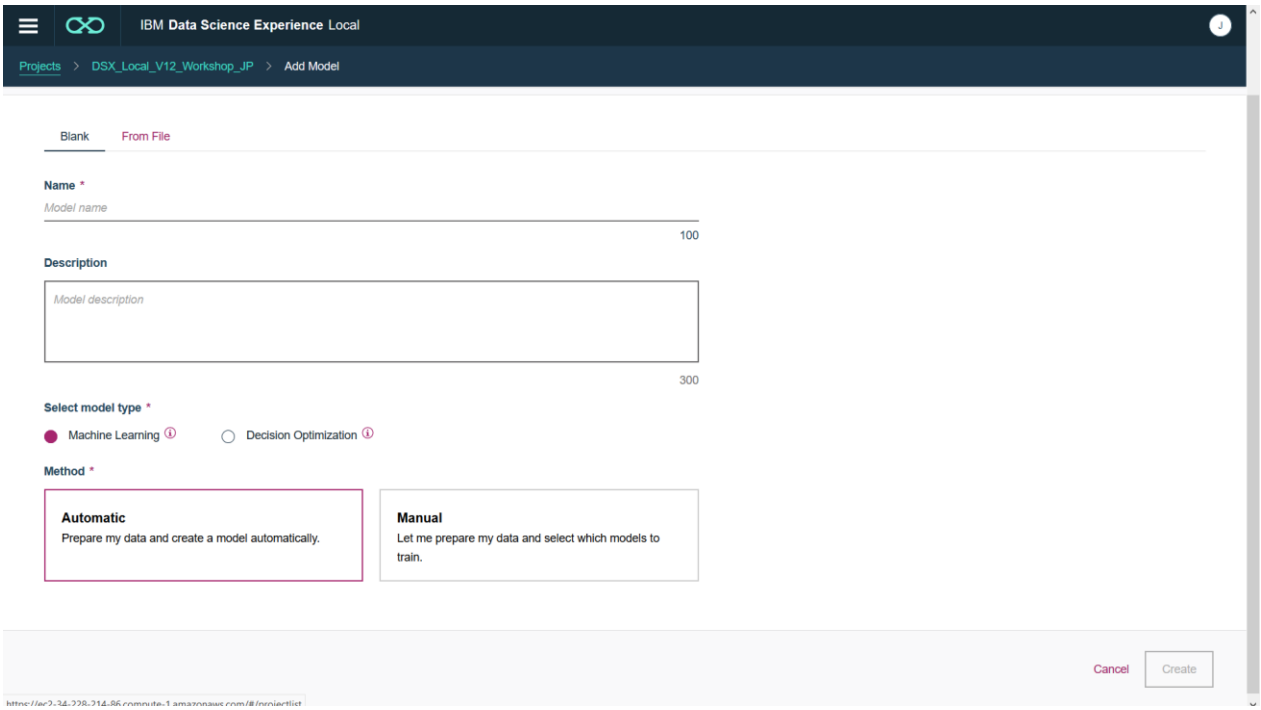


Step 2: Create a Model to predict survival

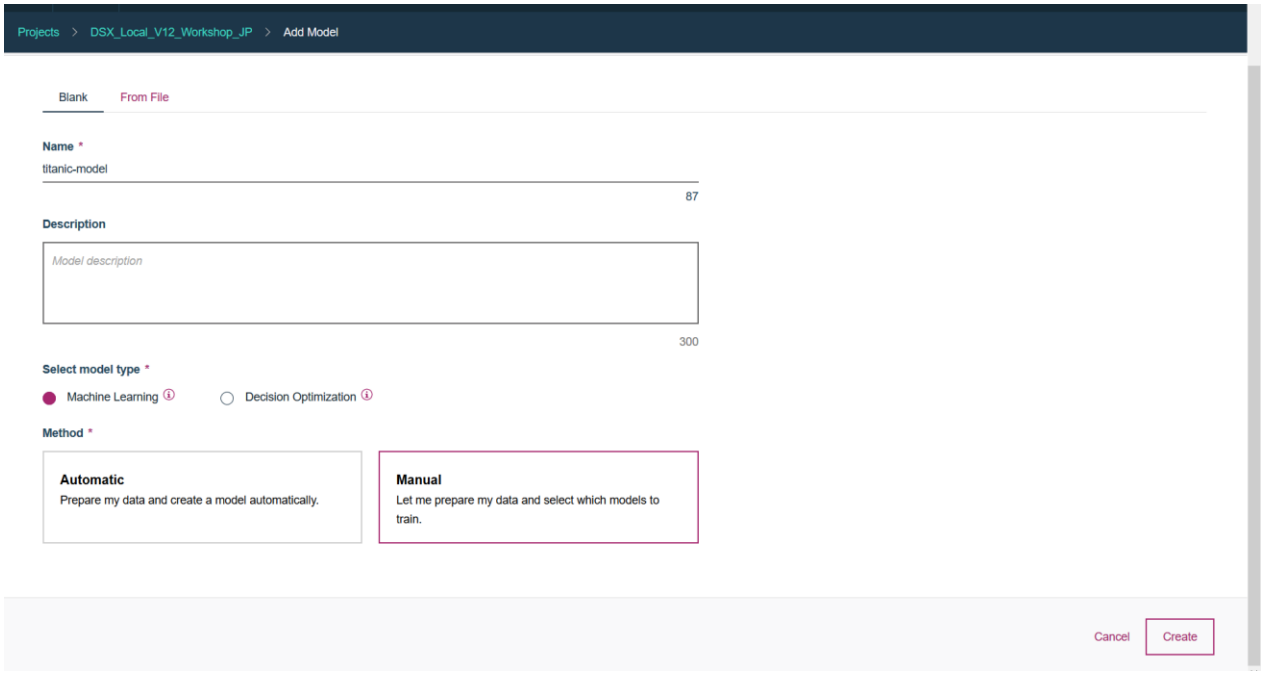
1. Select the **Models** tab



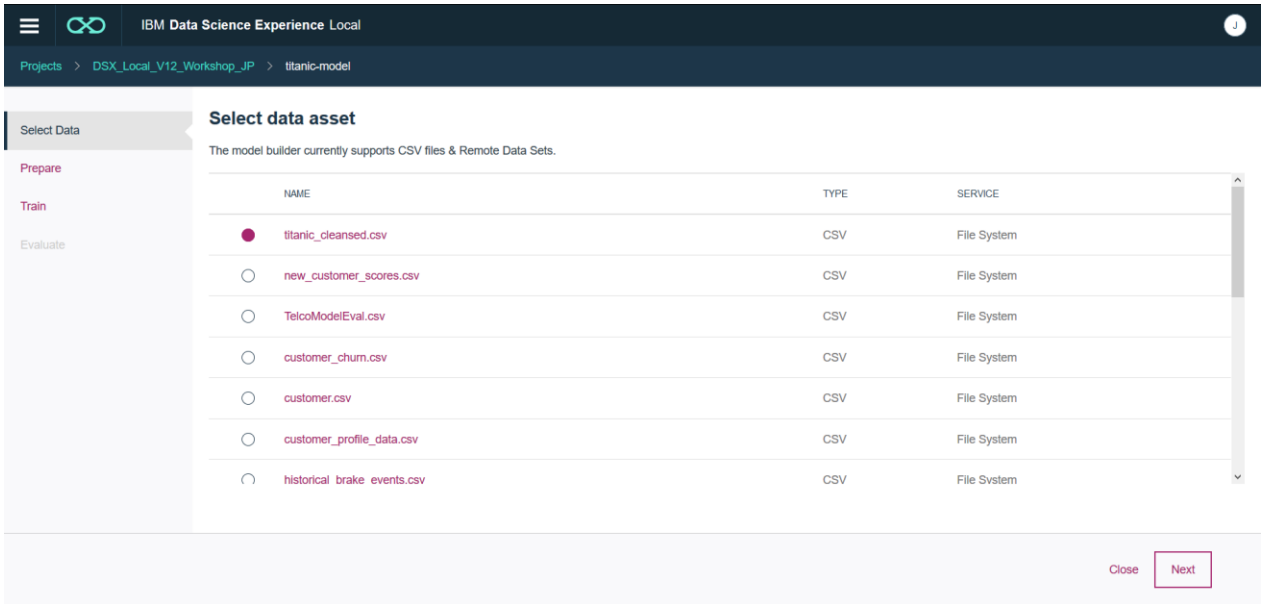
2. Select **add model**



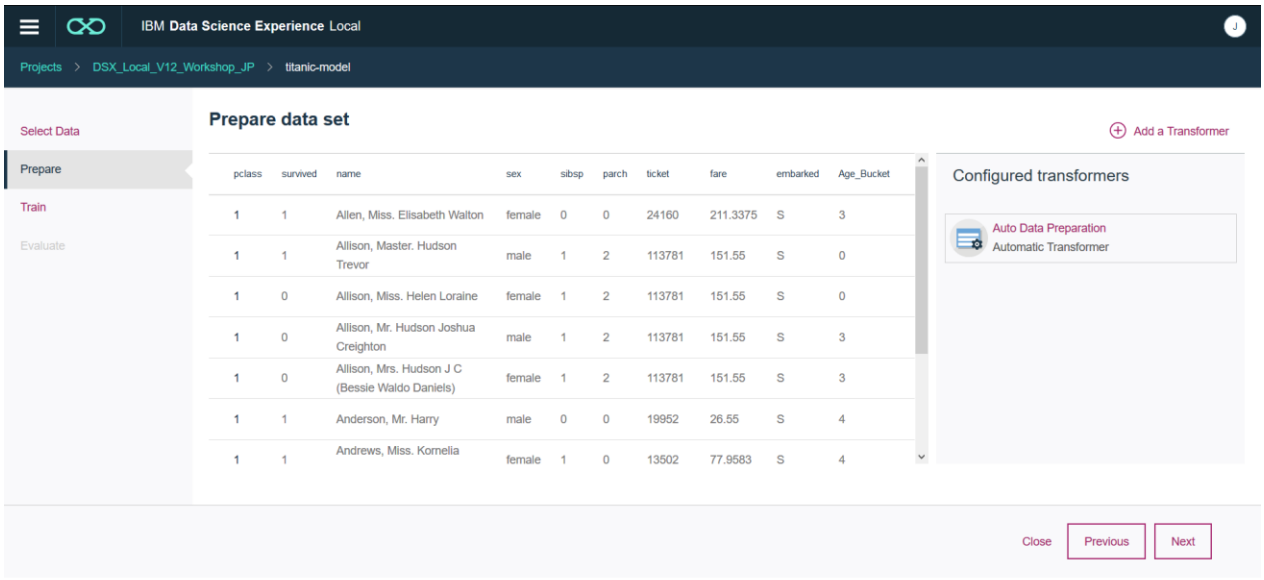
3. Enter a model **Name** (eg Titanic), optionally a **Description**, select **model type** of **Machine Learning** and select **Method** of **Manual**. Click on **Create**.



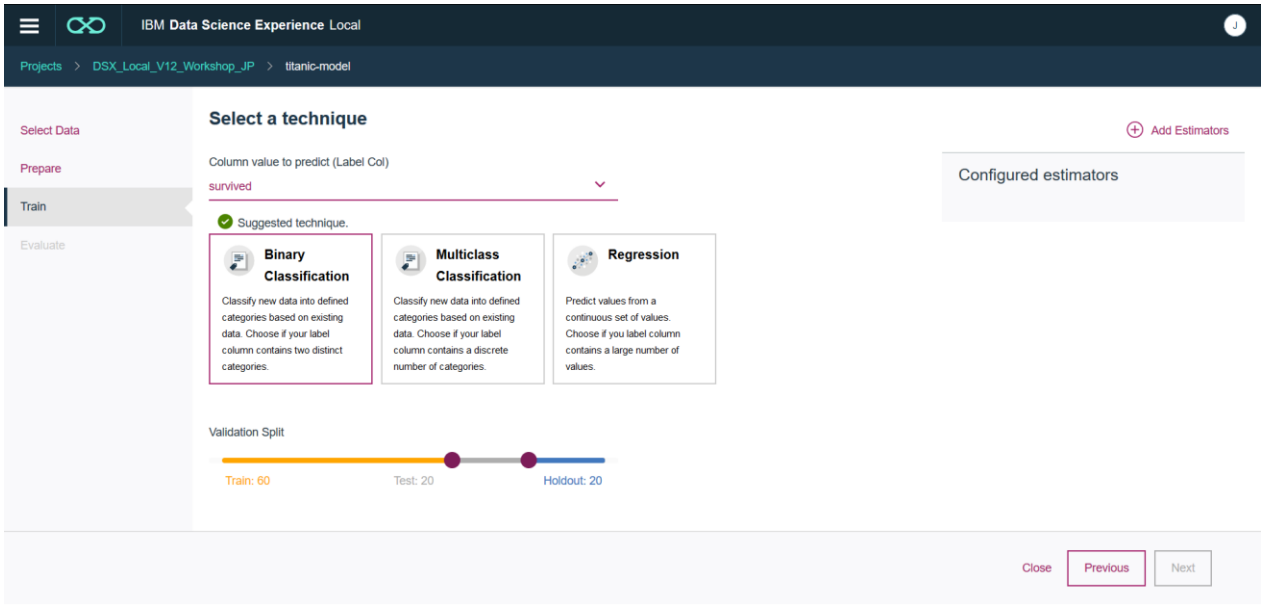
4. Click on the **titanic_cleansed.csv** and click on **Next**



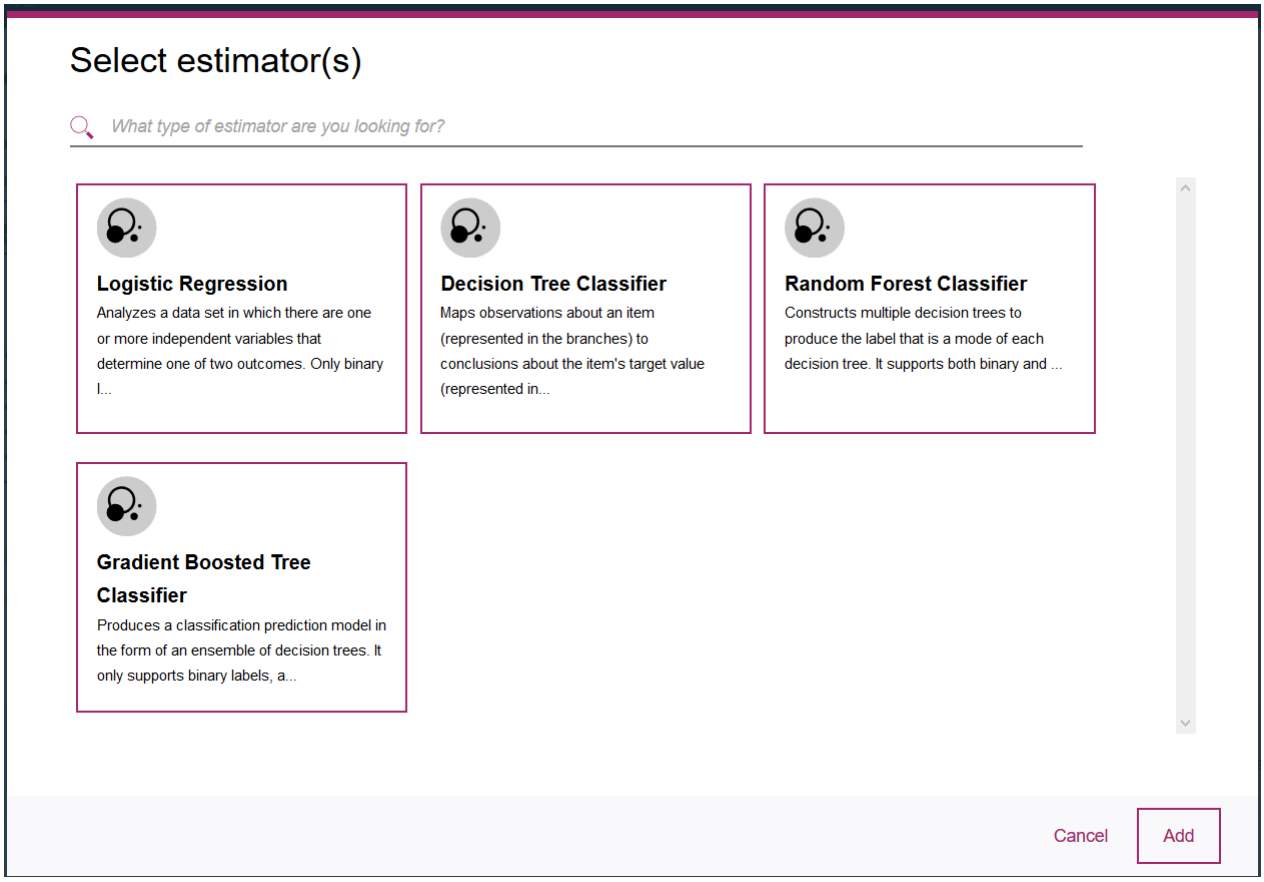
5. Select **Add a transformer** to see all available transformers. **Cancel** and use the configured **Auto Data Preparation** transformer. Select **Next.**



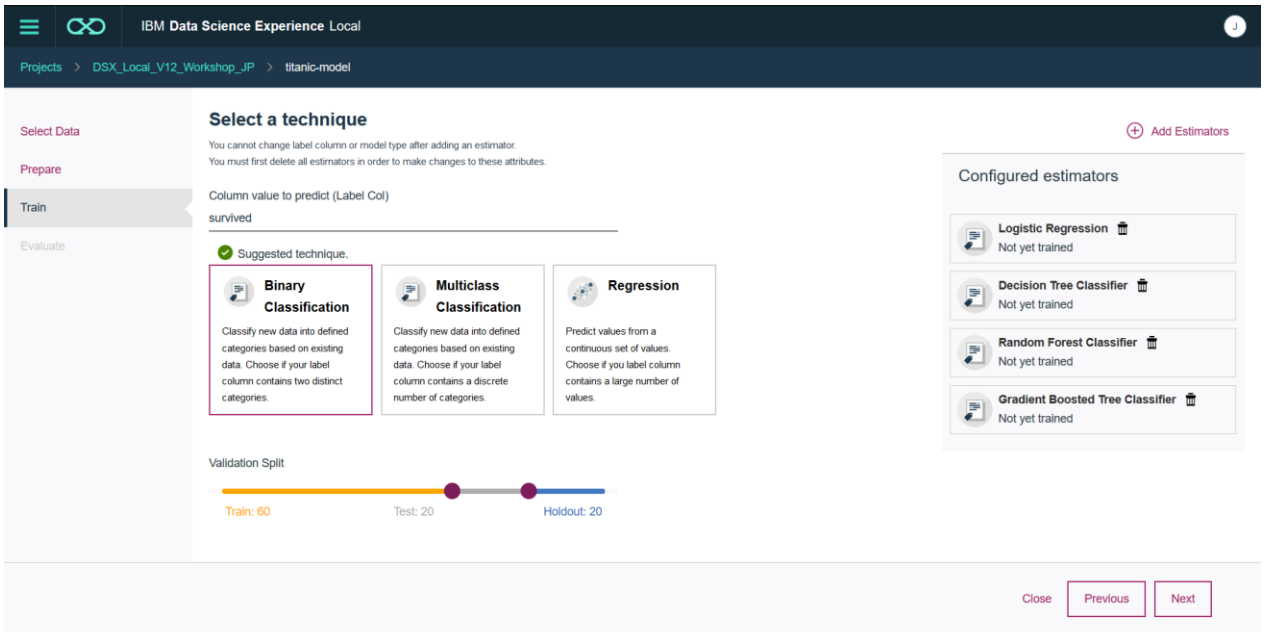
6. Select **Label Column** to **survived**. This will automatically set **Suggested technique** to Binary Classification.



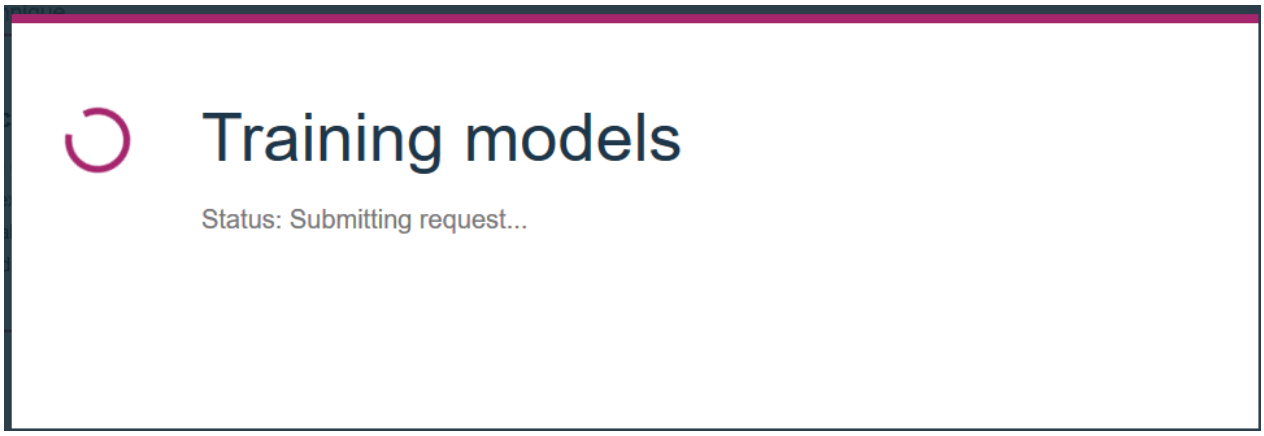
7. Select **Add Estimators.** Select all estimators and select **Add.**



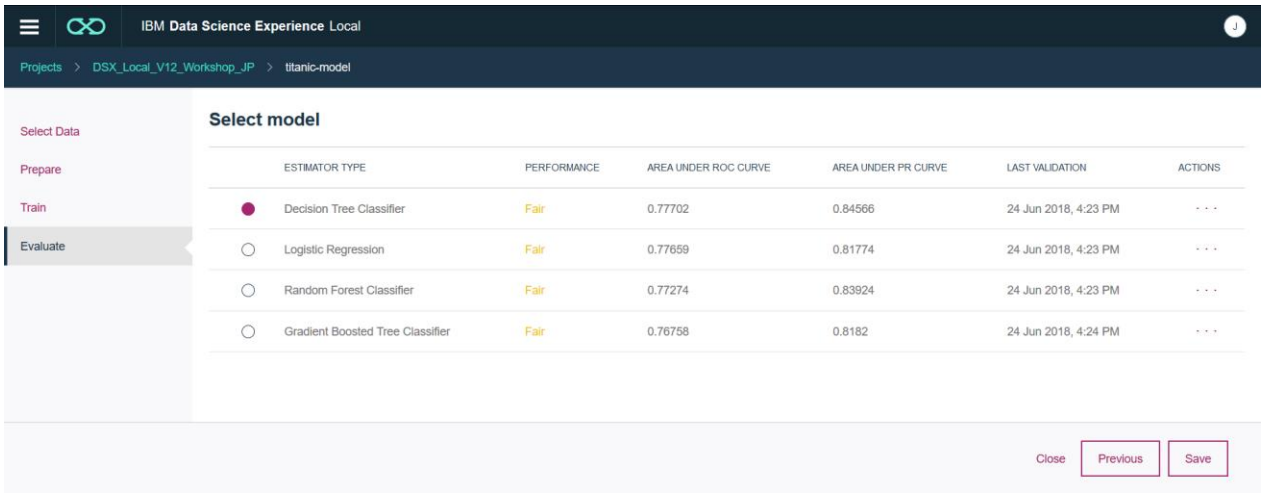
8. Select **Next.**



9. Wait for all models to be trained



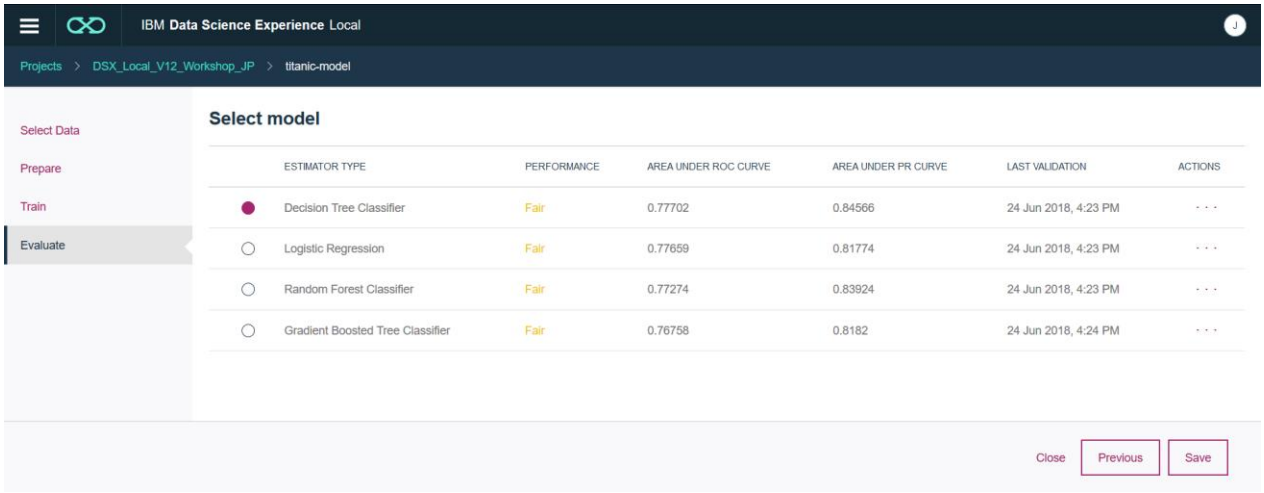
10. Review model performance. Models are ranked from best to worst performing.



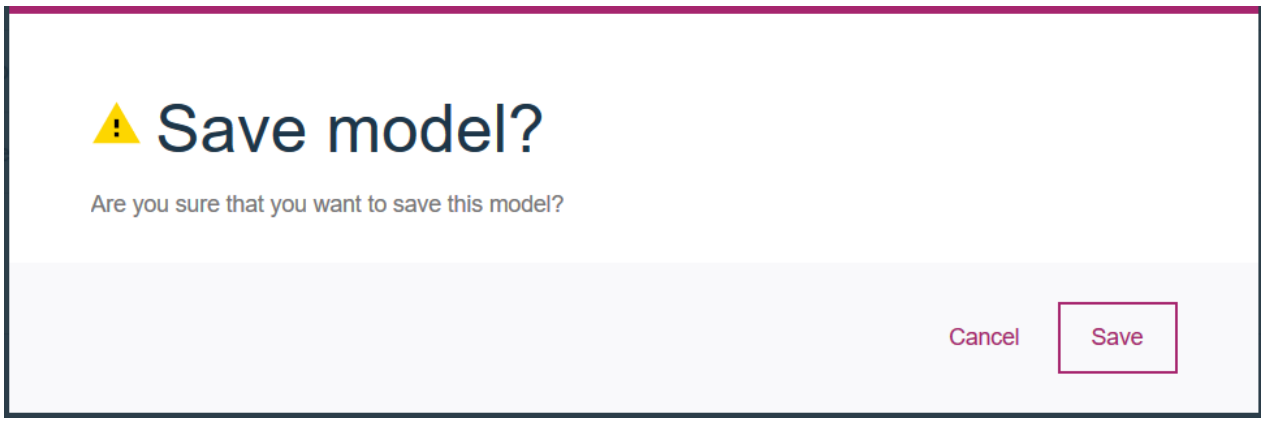
Step 3: Saving and Testing a Model

We can deploy the model to enable applications to invoke it via an API call. This is a Web Service deployment or Online deployment.

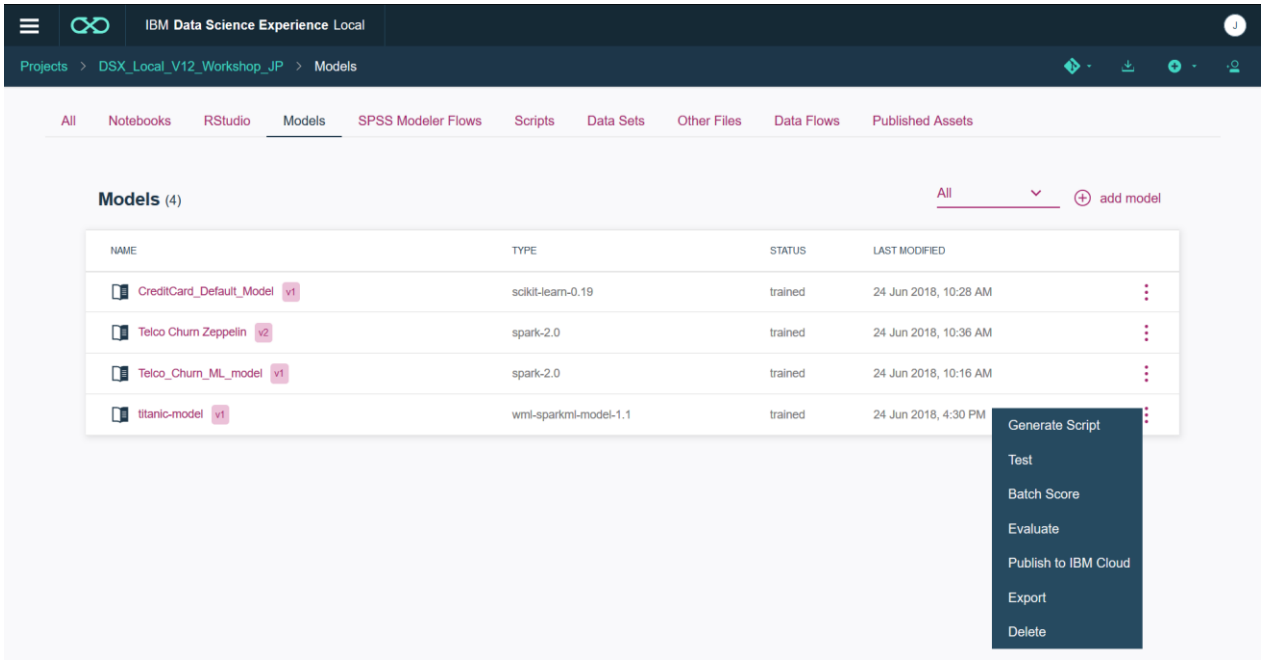
1. Select the **Save** button for the model you wish to deploy



2. Confirm the save.



3. The model now exists inside the **Models** tab of the project


IBM Data Science Experience Local
Projects > DSX_Local_V12_Workshop_JP > Models
All
Notebooks
RStudio
Models
SPSS Modeler Flows
Scripts
Data Sets
Other Files
Data Flows
Published Assets
Models (4)
All
add model
NAME
TYPE
STATUS
LAST MODIFIED
CreditCard_Default_Model v1
scikit-learn-0.19
trained
24 Jun 2018, 10:28 AM
Telco Churn Zeppelin v2
spark-2.0
trained
24 Jun 2018, 10:36 AM
Telco_Churn_ML_model v1
spark-2.0
trained
24 Jun 2018, 10:16 AM
titanic-model v1
wml-sparkml-model-1.1
trained
24 Jun 2018, 4:30 PM
Generate Script
Test
Batch Score
Evaluate
Publish to IBM Cloud
Export
Delete