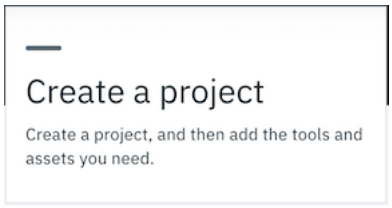
IBM® Watson Studio - Speed up ML/DL development with Modeler Flows

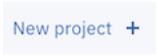
Creating a project

Before we can dig into the modeler flows tooling, we need to create a project where we'll put our "assets".

1. Click on **New project**.



or



2. Click on "Create an empty project"
3. In the "New project" screen:

- Enter the project name: **Tutorial**
- Optionally enter a description
- Click the **Create** button in the lower right corner.
- Click on the **Assets** tab:

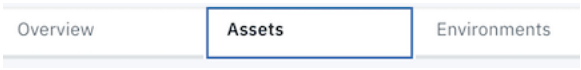
We can now move to the next task.

Setting up the environment

This lab requires the use of a few services and some data. We'll add more data specifically for the neural network section when we get to it.

1. Add the churn data to train the model

- The data file, **`customer_churn.csv`** is located HERE [https://ibm.box.com/s/5b00roqif0fw2t4lsbgaoc9sd8um90k4]. Download the file to your workstation.
- Click on the Assets tab at the upper left of the screen:



- The data window on the right side of your project is ready to load data. Drop the **`customer_churn.csv`** file in the window or use the browse option to locate the file on your machine.

2. Add services to the project

- Select the **`Settings`** tab at the top of the project screen
- Scroll down to the **`Associated services`** section
- Click on **`Add services`** and select **`Watson`** then find **`Machine learning`** in the list.
  - Scroll down and select the **lite** plan then click on **`Create`**

3. Retrieve the Machine learning service credentials

- Open a new tab and go to `https://cloud.ibm.com`
  Use the same username/password that you used in Watson Studio.
- Click on your machine learning service name
- Select **`Service credentials`** from the left menu
- Click on view credentials We'll be using those credentials later when scoring records from a notebook so make a copy of them for later.
  Return to the Watson Studio tab for the next task. Make sure to go under the **`Assets`** tab.

Modeler flows with the SPSS runtime

In this section, we use the Watson Studio Modeler Flows with the SPSS runtime to create and use a machine learning model for churn analysis.

1. Create a new flow
2. Click in the blue button **Add to project** at the upper right of the screen

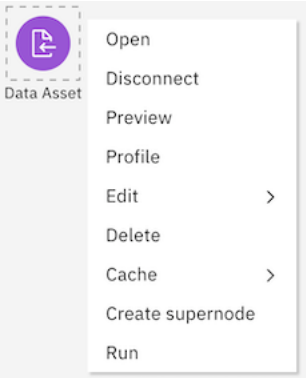
- Click on **`Modeler flow`** on the right side of the screen
- Type in a name, something like: **SPSS Churn Flow**
  Note that the default flow type is **Modeler Flow** ad the default runtime is **IBM SPSS Modeler**
- Click **`Create`**

2. Create the SPSS Churn flow

- Open the import section of the palette on the left side of the screen

- Drag the Data Asset operator onto the canvas

- Right-click on the data asset icon and select **`Open`**



Open DataAsset section icon

- In the right side of the screen, click on **`Change data asset`**
- On the left side of the screen, click on **`customer_churn.csv`**
- Click **Ok** in the bottom right, then **Save**

- Open the **Field Operations** section
- Drag the **Type** operator onto the canvas
- Connect **Data Asset** to **Type**

- Right-click on the **Type** icon and then click on **Open**
- Click on **`Read Values`**
- After it completes, click **Save**
- Click **`Save`**
- Close the **Field Operations** section
- Open the **Modeling** section
- Drag the **Auto Classifier** operator onto the canvas
- Connect the **Type** icon to the **Auto Classifier** icon
- Right-click on the **Auto Classifier** icon and then **Open**
- Check the **Use custom field roles** box
- In the drop down menu under target, select **CHURN**
- Click on **Add Columns**
- Check the box to the left of **Field name**
- Click **`OK`**
- Click **`Save`**

3. Generate the model
   - Click the **Run** icon

You will get a new **CHURN** icon on the canvas.

- Right-click on it and open **View Model**
- In the use column, only select the top model (should be C5.0)
- If you want, explore the models by clicking on their names

You can come back to the flow by clicking on **SPSS Churn Flow** at the top of the screen.

- Open the palette
- Open the **Outputs** section
- Drag a **Table** operator onto the canvas
- Connect the new **CHURN** icon to the **Table** icon
- Right-click on the **Table** icon and select **Save branch as model**
- Enter the model name such as: **SPSS Churn Model**
- Click **`Save`**
- Click **`Close`**
- Click on the the project name, **Tutorial** to return to the project list of assets.

4. Deploy the model

- Find the Watson Machine Learning section in the project asset page.
- Find the SPSS model: SPSS Churn Model
- Click on the three action dots and select **Deploy**
- Click on **Add Deployment**
- Enter a deployment name such as: **SPSS Churn Deployed**
- Click **`Save`**

1. Retrieve the scoring endpoint

   In this section, we retrieve the URL that is required to score data.

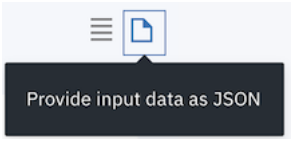
   - Once the status indicated **`Deploy success`**, click on **SPSS Churn Deployed** (you may have to refresh the screen for the status to update)
   - Click the **Implementation** tab

   In this section, you can see the scoring end-point and example code for cURL, Java, JavaScript, Python and Scala.

2. Score a record through Watson Studio

   - Download the following file [https://ibm.box.com/s/p6tpduxygbernetgrcs01cp4cm853e3v]
   - Click on the **Test** tab

- Select JSON input



- Copy the content of the file you just downloaded and paste it into the payload window
- Click on Predict
  You should see the prediction appear shortly

3. Creating a notebook

- Download the following notebook [https://ibm.box.com/s/53844nihp6qngr7e5sfmfes7lv438tz3]
- In the Tutorial project, click **Add to project**, then click **notebook**
- Click **From file**
- Click **Choose File**
- Drag the notebook file into the appropriate section
- Click on **Create Notebook**

7. Scoring records in a notebook

Here, instead of using the scoring end-point we saw earlier, we find the information programmatically. If we were to use the end-point, we could skip to the last three cells of the notebook but we would be less flexible.

- In the first cell, replace the wml_credentials with the one you retrieved in task 2 (step 3) from the cloud.ibm.com environment.
- Execute all the cells in order and look at the results

Note that the scoring response is a JSON document. The last cell of the notebook simply extracts specific fields from that response.

You can return to the **Tutorial** project by clicking on the project name at the top of the screen.

Modeler flows with the Spark runtime

In this section, we use the Watson Studio Modeler Flows with the Spark runtime to create and use a machine learning model for churn analysis.

1. Create a new flow

- Scroll down to the **Modeler Flows** section
- Click on **New flow** on the right side of the screen
- Type in a name, something like: **Spark Churn Flow**

Note that the default flow type is **Modeler Flow** ad the default runtime is **IBM SPSS Modeler**

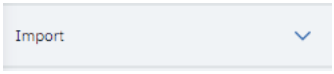
- Select the **Spark** runtime
- Select the **Scala Spark 2.1** runtime

The machine learning service and spark service that were added to the project in task 2 will automatically be selected for this flow.

- Click **`Create`**

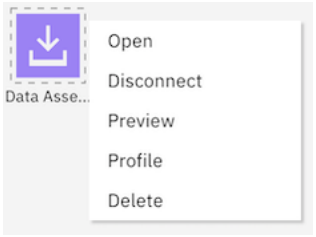
2. Create the Spark Churn flow

- Open the import section of the palette on the left side of the screen



- Drag the Data Asset operator onto the canvas

- Right-click on the data asset icon and select **`Open`**



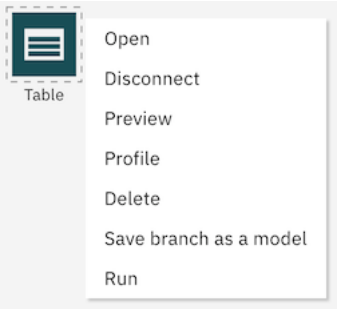
- In the right side of the screen, click on **`Change data asset`**
- On the left side of the screen, click on **Data assets** and then on **`customer_churn.csv`**
- Click **Ok** in the bottom right, then **Save**
- Open the **Modeling** section
- Drag the **Decision Tree Classifier** operator onto the canvas
- Connect the **Data Asset** icon to the **Decision Tree Classifier** icon

- Right-click on the **Decision Tree Classifier** icon and then **Open**
- In the `Target column` drop down menu under target, select **CHURN**
- Click on **Add Columns**
- Check the box to the left of **Field name**
- Click **`OK`**
- Click **`Save`**

3. Generate the model
   - Right-click on the **Decision Tree Classifier** icon and then **Run**

This generates an additional node with a diamond like mage in it.

- Add a **Table** output to the canvas and connect the **Diamond** to it
- Right-click on the **Table** icon and then **Save branch as a model**



- Enter the model name such as: **Spark Churn Model**
- Click **`Save`**
- Click **`Close`**
- Click on the the project name, **Tutorial** to return to the project list of assets.

4. Deploy the model

- Find the Watson Machine Learning section in the project asset page.
- Find the Spark model: Spark Churn Model
- Click on the three action dots and select **Deploy**
- Click on **Add Deployment**
- Enter a deployment name such as: **Spark Churn Deployed**
- Click **`Save`**

5. Retrieve the scoring endpoint

In this section, we retrieve the URL that is required to score data.

- Once the status indicated **`Deploy success`**, click on **Spark Churn Deployed** (you may have to refresh the screen)
- Click the **Implementation** tab

In this section, you can see the scoring end-point and example code for cURL, Java, JavaScript, Python and Scala.
It is more expedient to use the scoring end-point but we'll instead retrieve the information programmatically in the next step.

- Click on the **Tutorial** project to return to the asset page.

6. Creating a notebook

- Download the following notebook [https://ibm.box.com/s/3ncjlksakjqu9l2cloakvdcht4u9c0vv]
- In the Tutorial project, click **`New notebook`**
- Click **`From file`**
- Click **`Choose File`** and select the downloaded notebook
- Click on **`Create Notebook`**

7. Scoring records in a notebook

Here, instead of using the scoring end-point we saw earlier, we find the information programmatically. If we were to use the end-point, we could skip to the last three cells of the notebook but we would be less flexible.

- In the first cell, replace the wml_credentials with the one you retrieved in task 2 (step 3) from the cloud.ibm.com environment.
- Execute all the cells in order and look at the results

Note that the scoring response is a JSON document. The last cell of the notebook simply extracts specific fields from that response.

You can return to the **Tutorial** project by clicking on the project name at the top of the screen.

Congratulation

You completed the tutorial. Congratulations!

You completed the IBM® Watson Studio Modeler Flows: IBM® Watson Studio - Speed up ML/DL development with Modeler Flows. Throughout the tutorial, you saw:

- How efficient model building is using a drag and drop editor
- How easy it is to train a model
- How quickly a model can be made available through the deployment feature
- How a deployed model can be accessed programmatically.