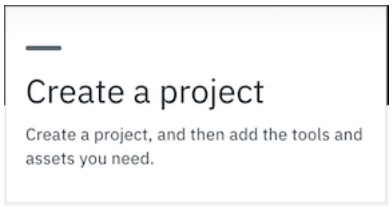
AutoAI Deep Dive Workshop

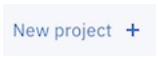
Creating a project

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



1. Click on "Create an empty project"
2. 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.

1. 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.
     - You should have an available service in a drop-down list
     - If not, click on the new tab
     - 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 **Services**
   - Click on your machine learning service name
   - Select **`Service credentials`** from the left menu
   - Click on the carret (v) at the left of the credential name to view the content. 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.

Creating an AutoAI Model

In this section, we use the Watson Studio AutoAI to create and use a machine learning model for churn analysis.

1. Click in the blue button **Add to project** at the upper right of the screen
   - Click on **`AutoAI experiment`**
   - Type in a name, something like: **AutoAI Churn Experiment**
   - Click **`Create`**

2. Select the dataset
   - Click **Select from project**
   - Select **customer_churn.csv**
   - Click **Select asset**

3. Select the target value

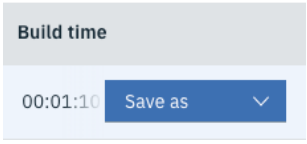
- Click on **CHURN** in the dropdown of "Select prediction column"
  This value is a string that provides a binary choice. For this reason, AutoAI selects Binary classification.

4. Look at the experiment settings
   - Click on **Experiment settings**
   - Look at the Data source settings
     Look at the training data split and columns included
   - Click on **Prediction**
     . Scroll down to see the Optimized metric choices.
     . Look at the available algorithms
     . Note that the number of algorithms that can be selected
   - Click on **Runtime**
     Take a look at the settings
   - Click on **Cancel**
5. Run the Experiment
   - Click on **Run experiment**
     At this point, AutoAI is creating and evaluating algorithms pipelines. We told AutoAI to use two algorithms. After looking at the data, it selects the two most promising algorithms.
     The creation process includes hyperparameters tuning and feature engineering.

6. Explore the run results
   - Take a look at:
     . Progress map **swap view**
     . Pipeline comparison tab
   - Scroll down to look at the resulting pipelines
     . Click on the ">" symbol at the left of the first pipeline
     You can collapse the information by clicking on "v"
     . Click on the pipeline name
     Go through the different tabs from **Model Evaluation** to **Feature Importance**
     . Click on **Back to AutoAI Churn Experiment** at the upper left of the screen

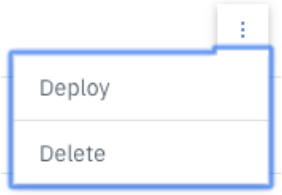
7. Save the model
   - Put your cursor to the right of the top pipeline. You should see:

   - Click on **Save as** and then on **Model**
   - Change the name to: **AutoAI Churn Model**
   - Click **Save**

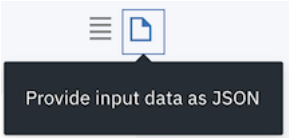
- Click **Save**
- Click on the project name at the upper left of the screen

8. Deploy the model
   - Under the asset tab, scroll down to the Models section
   - Click on the three vertical dots at the right of AutoAI Churn Model and click on **Deploy**



   - Click on Add Deployment
   - Give it the name **AutoAI Churn Deployed** and click Save
     The deployment may take a minute to complete.

9. Score data with the deployed model
   - Once the status indicated **`Deploy success`**, click on the deployment name, **AutoAI 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
   - Download the following file: ChurnTestData.txt
     [https://ibm.box.com/s/uxjd89q8dxk8sh0xsg8t58jx3svm2gw6]
   - Click on the **Test** tab
   - Select the 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

10. Create a notebook
    - Download the following notebook: AutoAI Model Access.ipynb
      [https://ibm.box.com/s/bffffhyk00nkuoyf7dm7ocbkx7wygya7]
    - In the Tutorial project, click **Add to project**, then click **`Notebook`**
    - Click **`From file`**
    - Drag the notebook file into the appropriate section

- Click on **`Create Notebook`**

11. 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 second cell, replace the wml_credentials content with the one you retrieved (in step 3 in setting up the environment) 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.