

AutoAI Lab

Introduction

This lab will demonstrate the award-winning AutoAI capability to build and deploy an optimized model based on the Titanic data set.

AutoAI in Cloud Pak for Data automatically analyzes your data and generates candidate model pipelines customized for your predictive modeling problem. AutoAI algorithms analyze your dataset to discover data transformations, estimator algorithms, and parameter settings that work best for your problem setting. Results are displayed on a leaderboard, showing the automatically generated model pipelines ranked according to your optimization objective.

Using AutoAI, you can build and deploy a machine learning model with sophisticated training features and no coding. The tool does most of the work for you.

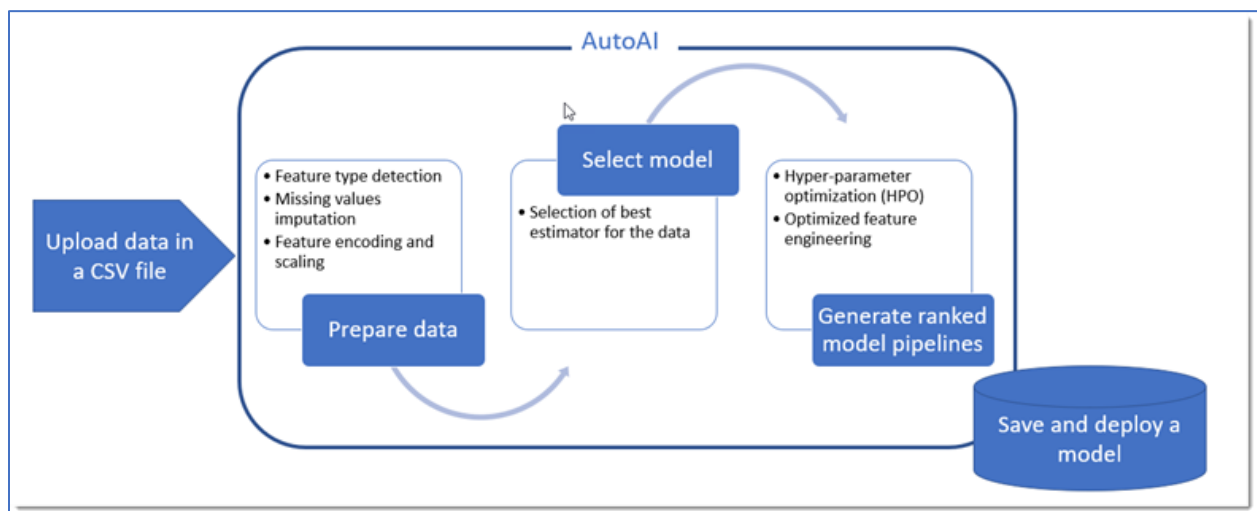


Figure 1- Auto AI Flow

The AutoAI process follows this sequence to build candidate pipelines:

- **Data pre-processing** - Most data sets contain different data formats and missing values, but standard machine learning algorithms work with numbers and no missing values. AutoAI applies various algorithms, or estimators, to analyze, clean, and prepare your raw data for machine learning. It automatically detects and categorizes features based on data type, such as categorical or numerical. Depending on the categorization, it uses hyper-parameter optimization to determine the best combination of strategies for missing value imputation, feature encoding, and feature scaling for your data.
- **Automated model selection** - The next step is automated model selection that matches your data. AutoAI uses a novel approach that enables testing and ranking candidate algorithms against small subsets of the data, gradually increasing the size of the subset

for the most promising algorithms to arrive at the best match. This approach saves time without sacrificing performance. It enables ranking a large number of candidate algorithms and selecting the best match for the data.

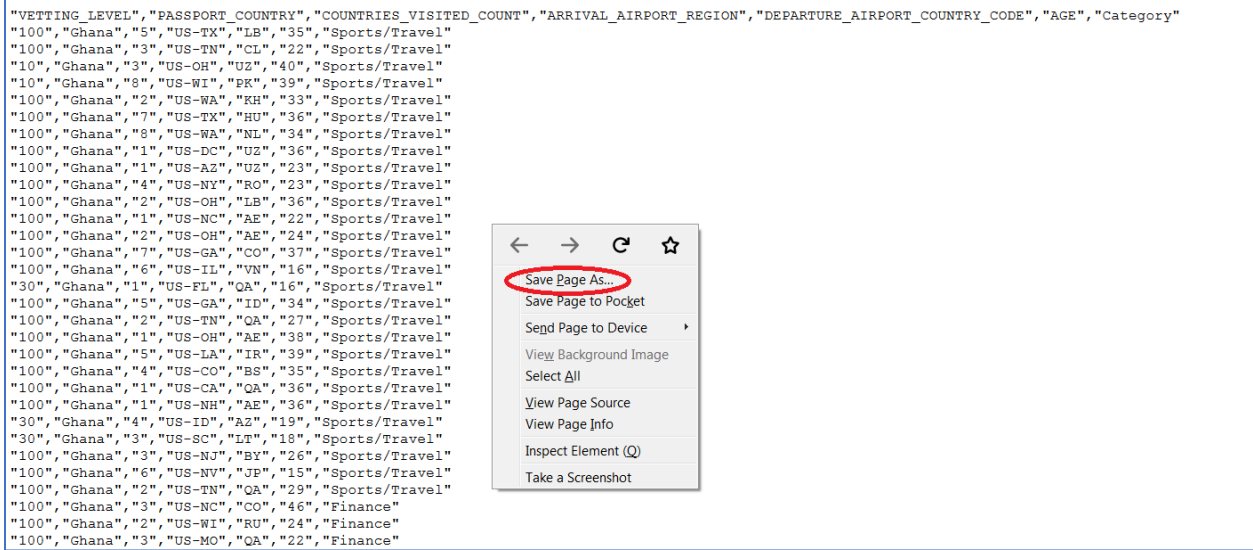
- **Hyperparameter optimization** - Hyper-parameter optimization refines the best performing model pipelines. AutoAI uses a novel hyper-parameter optimization algorithm optimized for costly function evaluations such as model training and scoring that are typical in machine learning. This approach enables fast convergence to a good solution despite long evaluation times of each iteration.
- **Automated feature engineering** - Feature engineering attempts to transform the raw data into the combination of features that best represents the problem to achieve the most accurate prediction. AutoAI uses a unique approach that explores various feature construction choices in a structured, non-exhaustive manner, while progressively maximizing model accuracy using reinforcement learning. This results in an optimized sequence of transformations for the data that best match the algorithms of the model selection step.
- **Repeat Hyperparameter optimization** – The Hyperparameter optimization step is repeated including the derived features from the feature engineering step.

We will perform the following steps in this lab:

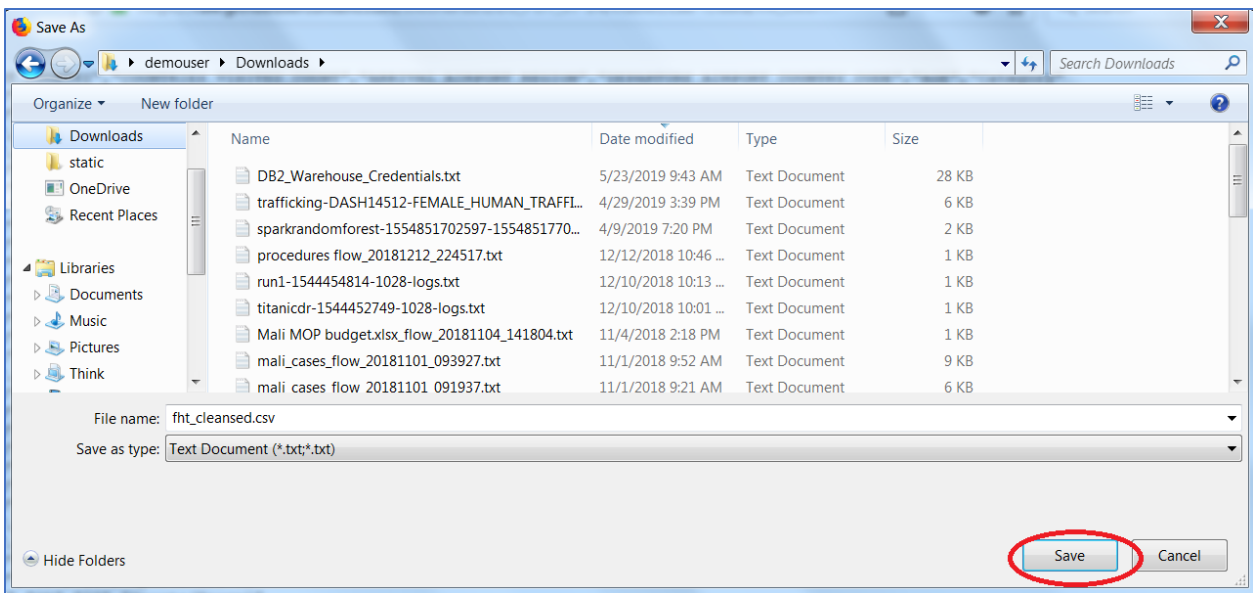
1. Download a female human trafficking cleansed dataset from the github repo
2. Add an Auto AI Experiment to create a model to predict the trafficking risk
3. Save and Deploy the model
4. Test the model

Step 1: Download a female human trafficking cleansed dataset

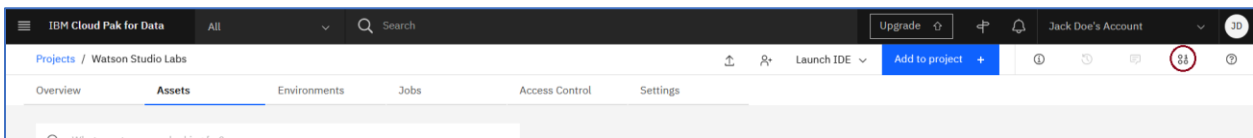
1. Download the fht_cleansed.csv data file from the following location by clicking [here](#)
2. Right-click on the window, and click **Save Page As...**



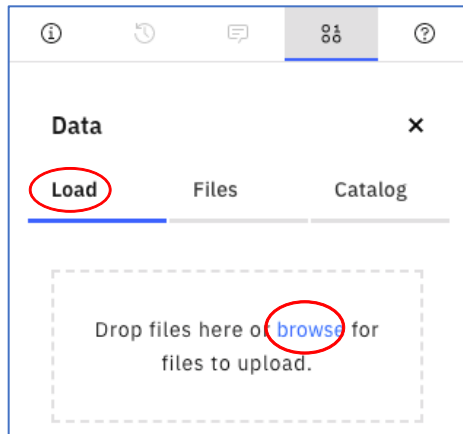
3. Click on Save.



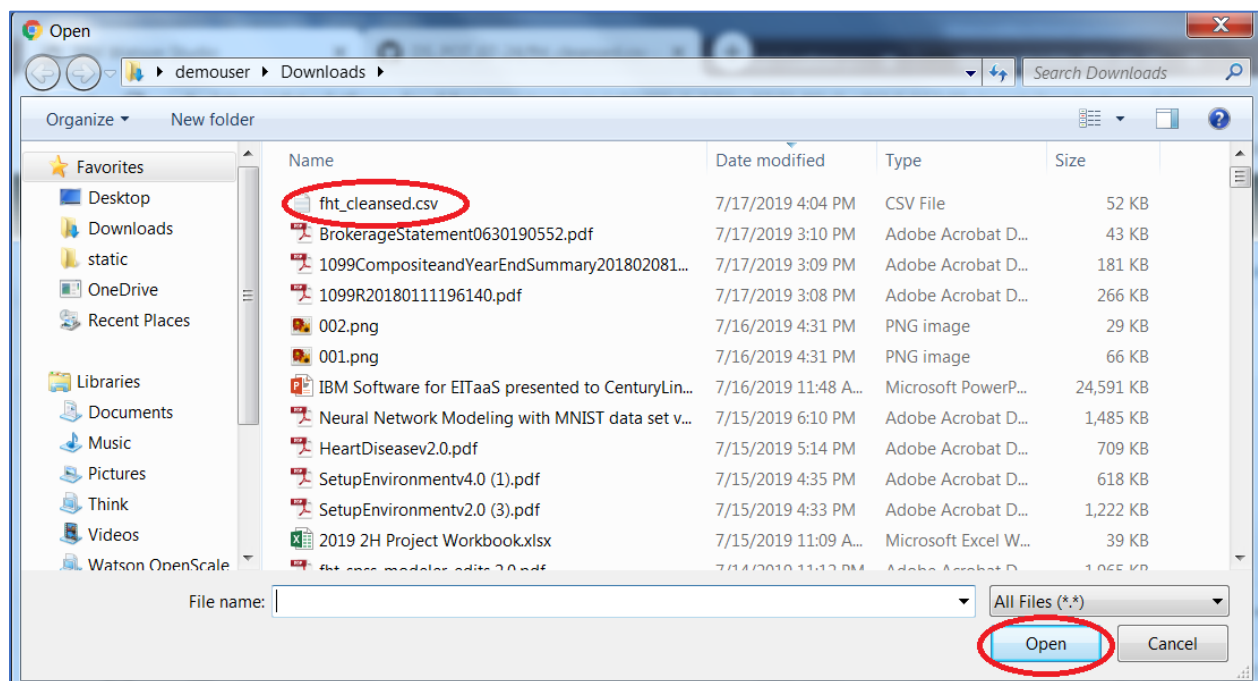
4. Go back to your project. Click on the icon.



5. Click on the **Load** tab and then click on **browse**. If you don't see the **Load** tab, click on the  icon again.



6. Go to the folder where the fht_cleansed.csv file is stored. Select the fht_cleansed.csv file and then click **Open**.

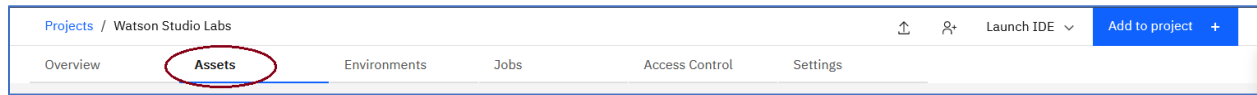


7. The fht_cleansed.csv file is now added as a Data Asset.

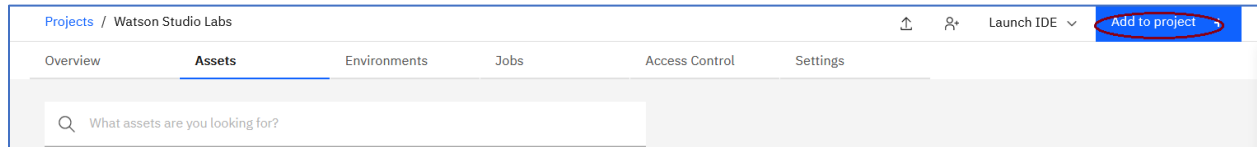
Data assets				
0 assets selected.				
<input type="checkbox"/>	Name	Type	Created by	Last modified
<input type="checkbox"/>	CSV fht_cleansed.csv	Data Asset	FCOT Labs	Aug 03, 2020, 10:18 AM

Step 2: Add an AutoAI Experiment.

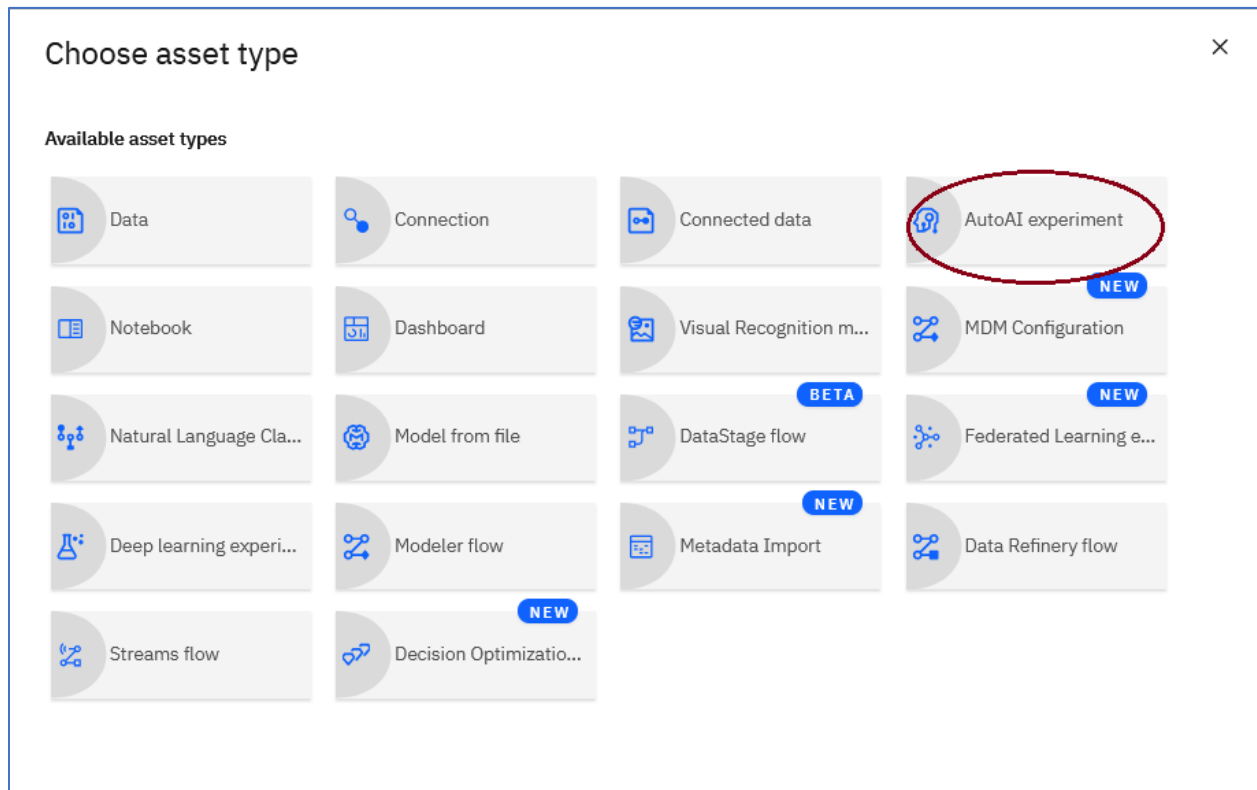
1. If not on the **Assets** page, click on the **Assets** Tab



2. Click on **Add to project**.



3. Click on **AutoAI Experiment**.



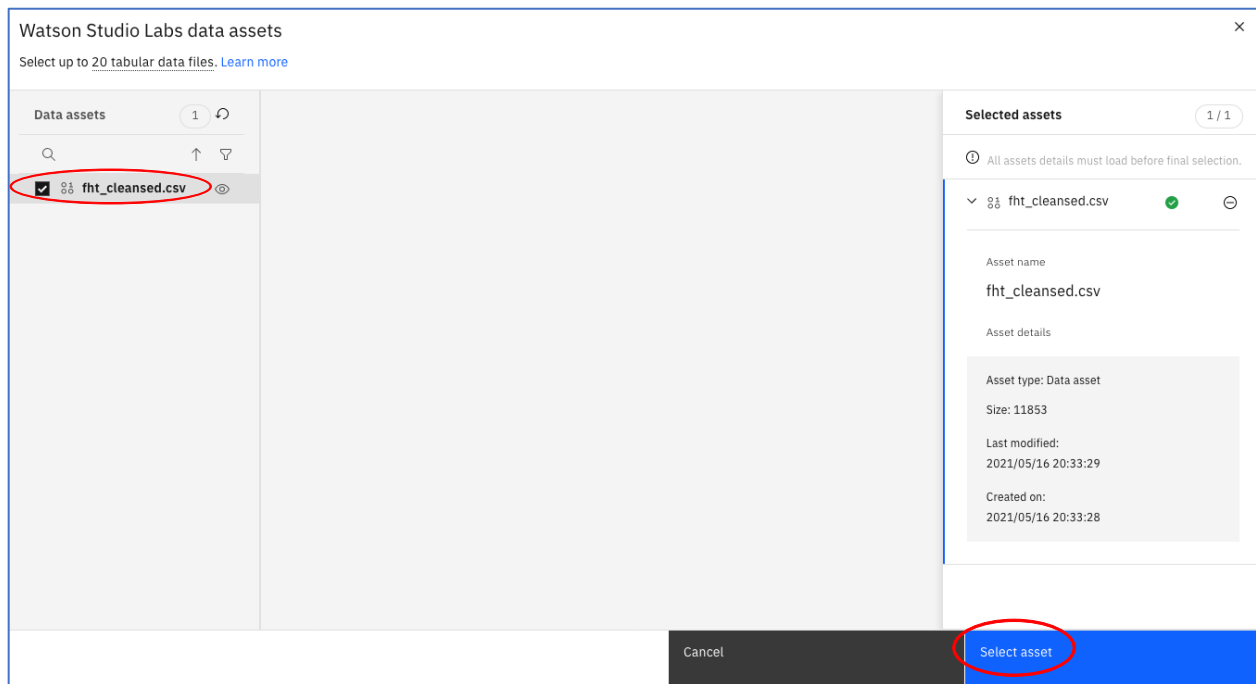
4. Enter an **Asset name**, leave the defaults for the **Watson Machine Learning** and **Compute configuration** and click on **Create**.

The screenshot shows the 'Define details' section of the Watson Machine Learning console. On the left, a sidebar contains a '+ New' button and a 'Gallery sample' link. The main form has two columns. The left column, titled 'Define details', contains a 'Name *' field with the text 'FHT AutoAI' (circled in red) and a 'Description' field with the placeholder text 'Description of AutoAI experiment'. The right column, titled 'Associate services', contains a 'Watson Machine Learning Service Instance *' dropdown menu with 'WatsonMachineLearning' selected, and a 'Compute configuration *' dropdown menu with '8 vCPU and 32 GB RAM' selected. Below these, a note states: 'This compute configuration consumes 20 capacity units per hour. [Learn more](#) about capacity unit hours and Watson Machine Learning pricing plans.' At the bottom right of the form are two buttons: 'Cancel' and 'Create' (circled in red).

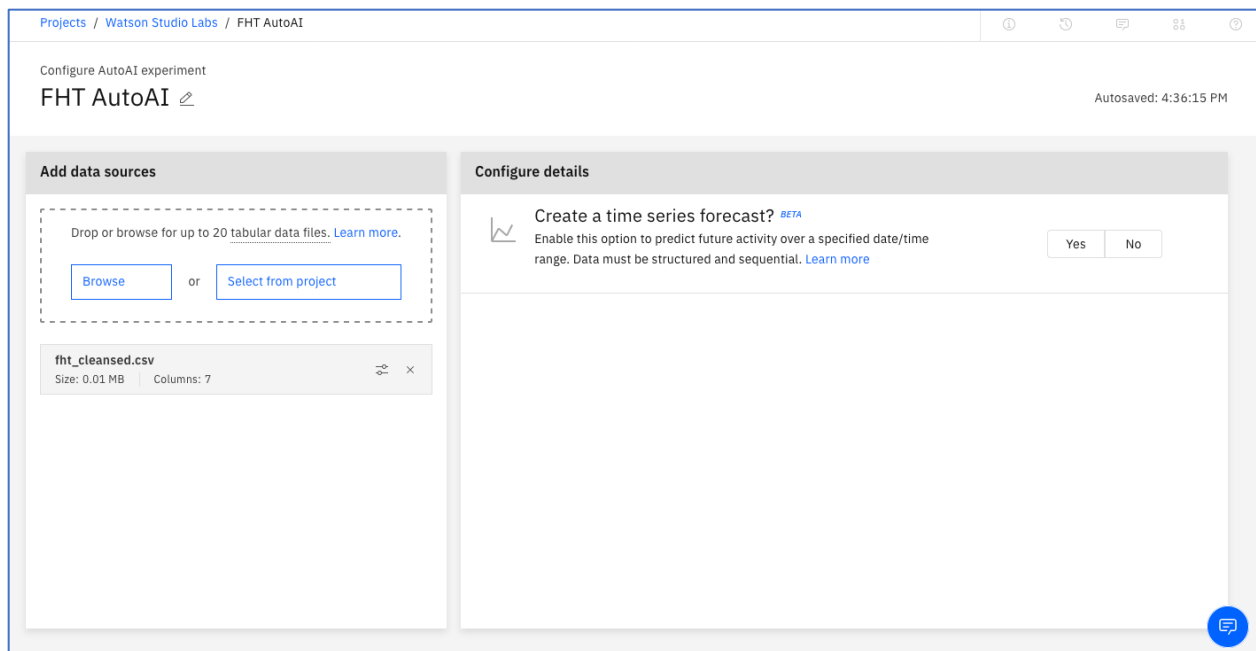
5. Click on **Select from project**.

The screenshot shows the 'Add data source' dialog in the Watson Machine Learning console. The title bar at the top says 'FHT AutoAI' with an edit icon. Below the title bar is a section titled 'Add data source'. The main area of the dialog is a large dashed rectangle. At the bottom center, there is a circular icon containing a document with a plus sign and a hand cursor. To the right of this icon, the text 'Add data source' is displayed, followed by the instruction 'Drop or browse for a csv file.' Below this instruction are two buttons: 'Browse' and 'Select from project' (circled in red).

6. Select the **fht_cleansed.csv** dataset and click on **Select asset**.



7. Click **No** next to Create a time series forecast?



8. Select **VETTING_LEVEL** as the column to predict, leave the default for **OPTIMIZED METRIC** and click on **Run Experiment**. Note the system scanned the **VETTING_LEVEL** values to determine that a **Multiclass Classification** was the **PREDICTION TYPE**.

Configure details

Create a time series forecast? [BETA](#)
Enable this option to predict future activity over a specified date/time range. Data must be structured and sequential. [Learn more](#)

What do you want to predict?
Prediction column ⓘ
VETTING_LEVEL × ▾

Prediction column: VETTING_LEVEL CUH remaining: 20 CUH

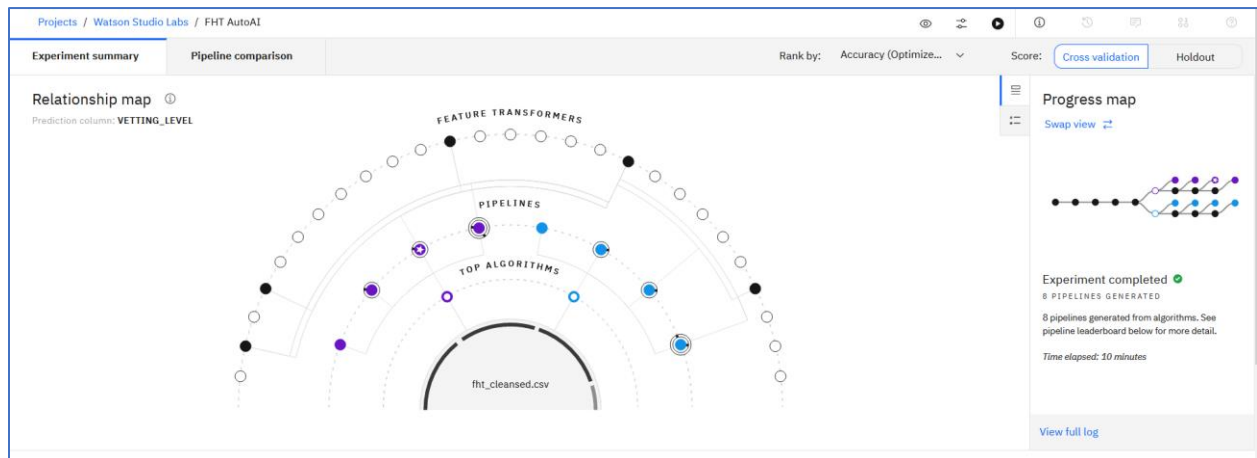
PREDICTION TYPE
Multiclass Classification



OPTIMIZED FOR
Accuracy & run time

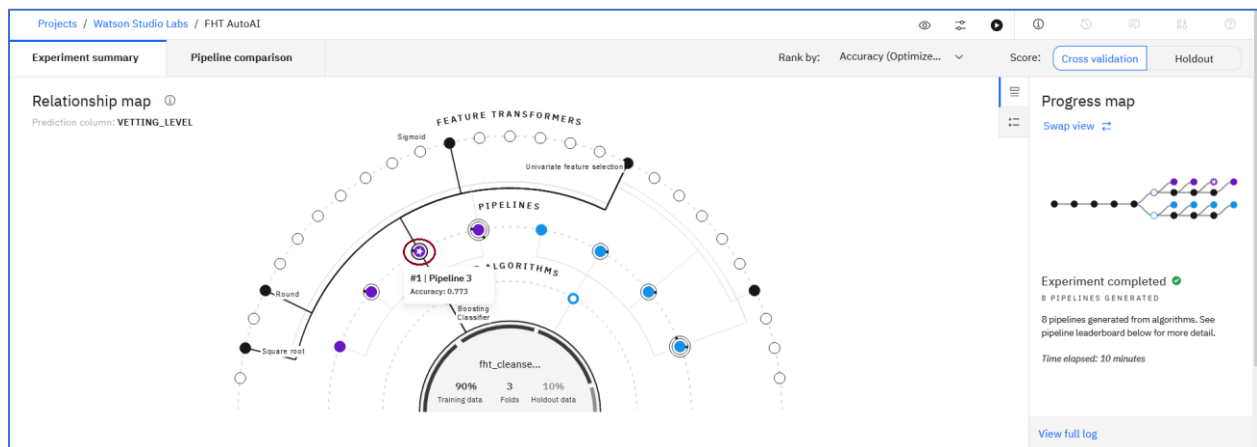
Experiment settings **Run experiment**

9. It will take several minutes for the eight alternative pipelines to be analyzed. The first pipeline picks the best algorithm. In this case, it is a **Gradient Boosting** classifier. The second pipeline performs a hyperparameter optimization to see if tuning the algorithm parameters will improve the performance metric. The third pipeline will derive new features (i.e. feature engineering) to try to improve the performance metric. The fourth pipeline will do another hyperparameter optimization including the newly derived features. The next 4 pipelines do the same thing for the second-best algorithm. Note, you can move ahead at any point after the first pipeline has been created. We are going to proceed as if Pipeline 3 is the best ranked pipeline. Yours could be different.

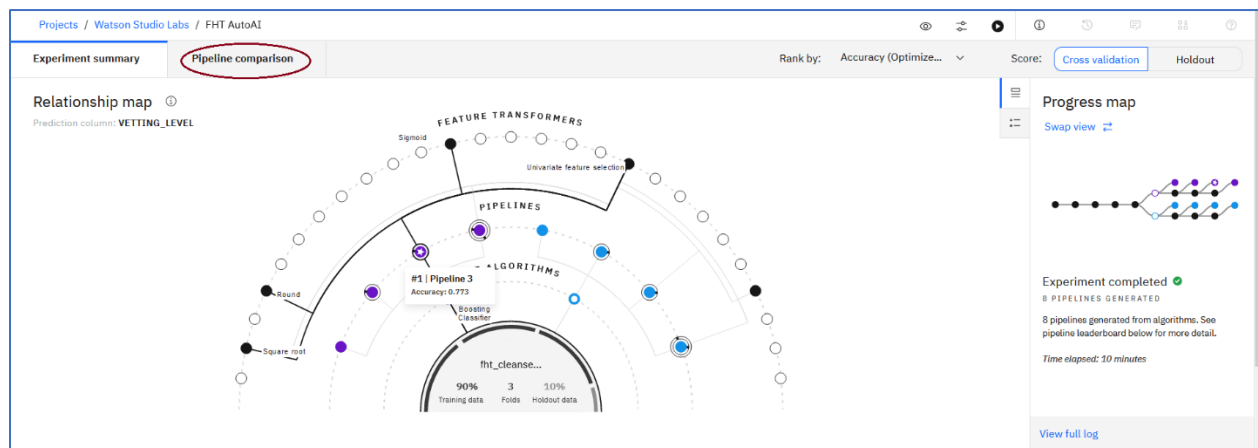
10. When the AutoAI run is completed you will see a display similar to below.



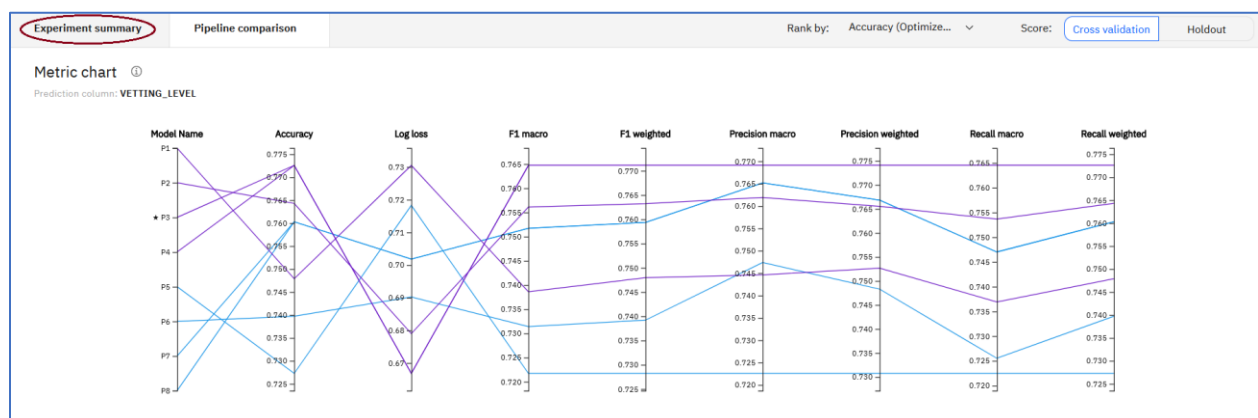
11. Hover over the  icon in the Pipelines semi-circle. This is the top ranked pipeline based on the accuracy metric. It is shown to be Pipeline 3 with a Gradient Boosted classifier and an accuracy of .77. Note, if you click on the  icon, you will be taken to the metrics for Pipeline 3 (don't do this yet).




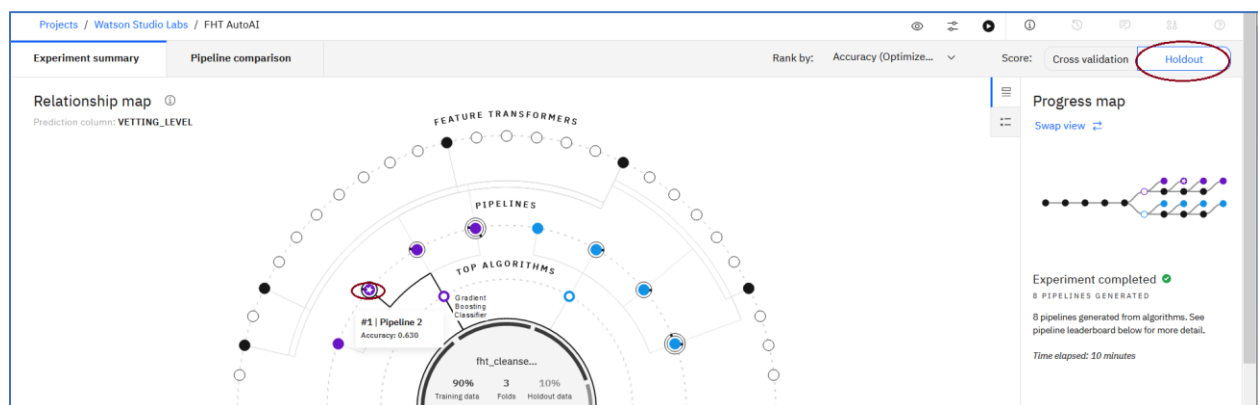
12. Click on the **Pipeline comparison**.



13. This will display a diagram showing the metrics associated with each pipeline on a single chart. Click on **Experiment summary** to return to the prior panel.



14. Click on **Holdout**. Note that the  icon may have shifted.



15. Click on the **Cross validation** tab and then scroll down. The Pipeline leaderboard is displayed with the pipelines ranked by accuracy.

Experiment summary

Pipeline comparison

Rank by: Accuracy (Optimize...)

Score:


Cross validation

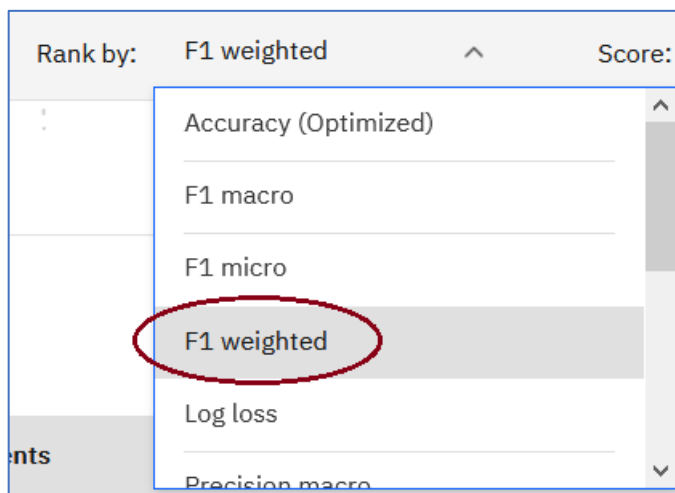
Holdout

View full log

Pipeline leaderboard

Rank	↑	Name	Algorithm	Accuracy (Optimized)	Enhancements	Build time
★ 1		Pipeline 3	<div></div> Gradient Boosting Classifier	0.773	<div>HPO-1FE</div>	00:01:12
2		Pipeline 4	<div></div> Gradient Boosting Classifier	0.773	<div>HPO-1FEHPO-2</div>	00:00:43
3		Pipeline 2	<div></div> Gradient Boosting Classifier	0.764	<div>HPO-1</div>	00:00:16
4		Pipeline 7	<div></div> XGB Classifier	0.760	<div>HPO-1FE</div>	00:02:08
5		Pipeline 8	<div></div> XGB Classifier	0.760	<div>HPO-1FEHPO-2</div>	00:01:29
6		Pipeline 1	<div></div> Gradient Boosting Classifier	0.748	None	00:00:03
7		Pipeline 6	<div></div> XGB Classifier	0.740	<div>HPO-1</div>	00:00:24
8		Pipeline 5	<div></div> XGB Classifier	0.727	None	00:00:01

16. Click on the  icon adjacent to Accuracy and click F1 Weighted



17. The list is now ranked by the F1 Weighted metric.

Experiment summary

Pipeline comparison

Rank by: F1 weighted

Score:

Cross validation

</

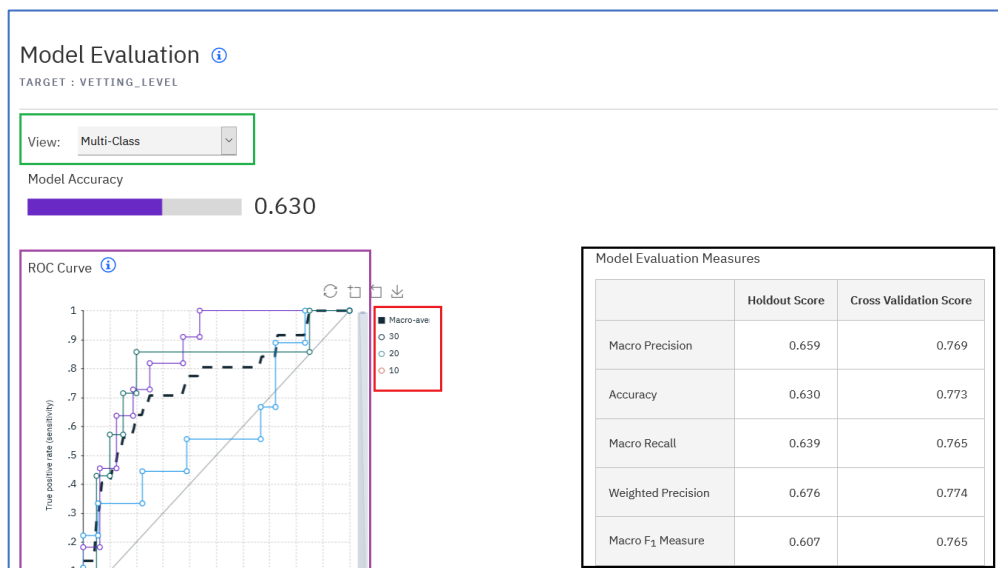
18. Return to the Accuracy ranked list by clicking on the  and selecting **Accuracy (Optimized)**. Click on the highest ranked pipeline (**pipeline 3 in the picture**).

Experiment summary

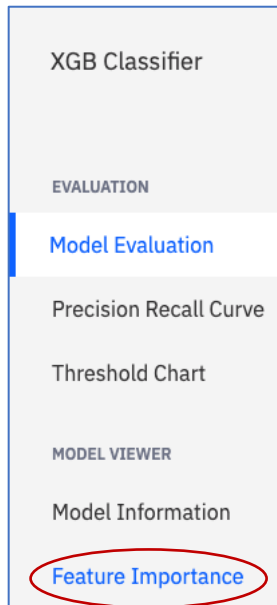
Pipeline comparison

</

19. The **Model Evaluation** panel is displayed. The model evaluation metric for the holdout sample is displayed. Highlighted in blue rectangle below are links to additional model information and metrics. Highlighted in the green rectangle is a drop down containing the options multi-class, and then options for 3 binary metrics (30 versus rest, 20 versus rest, and 10 versus rest). Highlighted in the purple rectangle is the Receiver Operating Characteristic (ROC) curve. This plots the true positive rate versus the false positive rate at varying thresholds. Highlighted in the red rectangle is a clickable legend. Clicking on the circles will toggle the corresponding display on/off. Finally, highlighted in the black rectangle are model metrics for both the holdout sample, and the cross validation sample.



20. Click on **Feature Importance**.



21. According to the **Feature Importance**, the **COUNTRIES_VISITED_COUNT** feature is the most important.

Step 3: Save and Deploy the Model

In this section, we will save and then deploy the model. You create a deployment for a machine learning model so you can submit new data to a model and get a score, or prediction back. To deploy the model, you must have a *deployment space* where you can organize the assets you need to create and monitor deployments. A space contains an overview of deployment status, the deployable assets, deployments, associated input and output data, and the associated environments.

To deploy a model, we first **promote** the model to a deployment space. When you promote a model to a space, components required for a successful deployment, such as a training library, model definition, other dependent assets, or environment definition are automatically promoted as well. After promoting a model and the data assets needed to a deployment space, you can create a deployment in the space.

You can create a deployment space, promote the model, and create a deployment through the user interface or programmatically.

1. **Click on Save as.**



- AutoAI can generate both a machine learning model as well as a Jupyter notebook. The Jupyter notebook can contain the source code to build the model, or the source code plus the deployment code. Feel free to experiment with saving the notebooks. For this lab, we will save the model. Make sure the Model is checked, optionally change the **Name**, and click **Create**.

Select asset type

Model

Create a Watson Machine Learning model asset that you can test with new data, deploy to generate predictions, and trace lineage activity.

Notebook Beta

Create a notebook if you want to view the code that created this model pipeline or interact with the model programmatically.

Define details

Name

FHT AutoAI - P3 GradientBoostingClassifierEstimator

Description (optional)

Enter description here

Tags

Add tags to make assets easier to find.

Add a tag

Cancel Create

- A message is displayed showing the model was successfully saved. Click on **View in project**.

Rank 1 Pipeline 3

Holdout Accuracy (Optimized) 0.630

Algorithm Gradient Boosting Classifier

Enhancements HPO-1 FE

Build time 00:01:19

Save as

Gradient Boosting Classifier

Feature Transformations

TARGET : VETTING_LEVEL

Saved model successfully.
FHT AutoAI - P3
GradientBoostingClassifierEstimator
was successfully saved to Watson
Studio Labs.

View in project

- As mentioned above, we will need to promote the model to a deployment space. Click on **Promote to deployment space**.

FHT AutoAI - P3 GradientBoostingClassifierEstimator

Promote to deployment space

Overview Activities

Input Schema

Input

Column	Type
AGE	"integer"
ARRIVAL_STATE	"other"
COUNTRIES_VISITED_COUNT	"integer"
Category	"other"
DEPARTURE_AIRPORT_COUNTRY_CODE	"other"

5. A deployment space already exists due to the Watson OpenScale set up in Lab-1. We will create a new space for our model deployment. Click on **New space**.

Promote to space

Target space

Select or create a space ^

Create a new deployment space +

openscale-express-path-15c8dbf8-a3e6-4da4-ac8d-d7b4180b6a12

6. Enter **AutoAI Deployment Space** for the Name, scroll down and select **WatsonMachineLearning** for the **Select machine learning service**, and click **Create**.

Create a deployment space

Use a space to collect assets in one place to create, run, and manage deployments

Name

AutoAI Deployment Space

Description (Optional)

Deployment space description

Deployment space tags (optional) ⓘ

Add a tag

Select storage service ⓘ

Cloud Object Storage-zj

Select machine learning service (optional) ⓘ

WatsonMachineLearning

Cancel Create

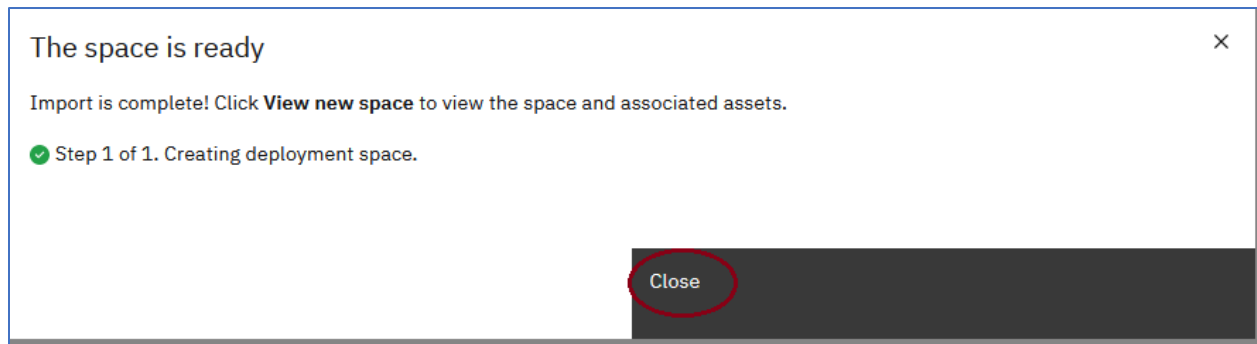
7. Two message boxes are displayed. Click **Close** in the second message box.

The space is being prepared... X

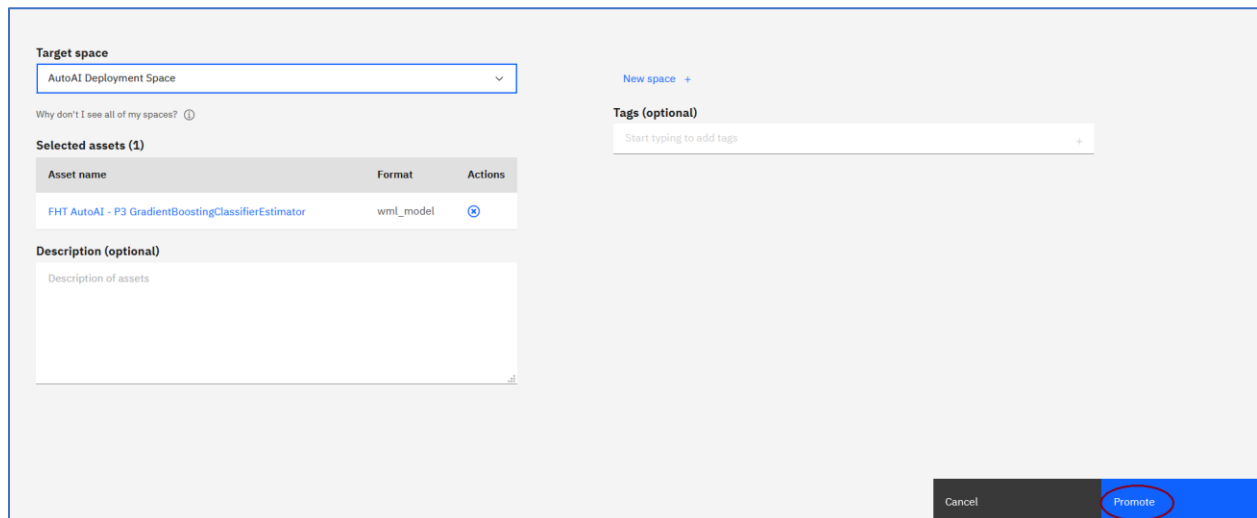
The space "AutoAI Deployment Space" is being created.

Step 1 of 1. Creating deployment space.

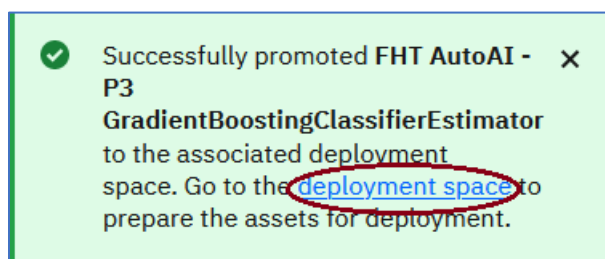
Close





8. Click **Promote**.



9. A message is displayed. Click **Deployment Space** in the message.






10. Hover over the promoted model until you see the deploy icon  appear on the right. Click on the  icon.

AutoAI Deployment Space

Assets Deployments Jobs Access control Settings

What assets are you looking for?

Models (1) [Import model +](#)

Name	Type	Software specification	Asset tags	Last modified	
 FHT AutoAI - P3 GradientBoostingClassifierEstimato...	wml-hybrid_0.1	hybrid_0.1		Jan 12, 2021 9:22 PM	 

Deploy

11. Click **Online** for an **Deployment type**, enter **FHT AutoAI Deployment** for the **Name**, and click **Create**.

Associated asset
FHT AutoAI - P3 GradientBoostingClassifierEstimator

Deployment type

Online
Run the model on data in real-time, as data is received by a web service.

Batch
Run the model against data as a batch process.

Name
FHT AutoAI Deployment

Description
Deployment description

Software specification
hybrid_0.1
The software specification is predefined for the asset type. You can update or customize the software specification programmatically. [Learn more](#)

Cancel Create

12. Click the Deployments tab. Once the asset has “Deployed” under Status, click on the deployment name.



Deployments / AutoAI Deployment Space [Add to space +](#)

AutoAI Deployment Space

Assets **Deployments** Jobs Manage

What deployments are you looking for?

Deployments (1)

Name	Type	Status	Asset	Tags	Last modified
 FHT AutoAI Deployment	Online	 Deployed	FHT AutoAI - P2 XGBClassifier		May 16, 2021 6:42 PM

13. The **API Reference** panel provides information for the application developers to invoke the deployed model. It includes sample code in various programming languages and the scoring endpoint to be used when invoking the web service.

The screenshot shows the 'FHT AutoAI Deployment' interface. At the top right, there are status indicators: a green checkmark for 'Deployed' and a grey circle for 'Online'. Below the title, there are two tabs: 'API reference' and 'Test', with 'Test' being the active tab and circled in red. The 'Direct link' section displays an 'Endpoint' URL: `https://us-south.ml.cloud.ibm.com/ml/v4/deployments/75a93d69-f6ac-4b0a-b4d8-43b79d38e411/predictions`. To the right of the URL is a 'Bearer <token>' field with a dropdown menu currently set to 'IAM'. Below this, the 'Code snippets' section has tabs for 'cURL', 'Java', 'JavaScript', 'Python', and 'Scala', with 'cURL' selected. The cURL tab contains the following text:
NOTE: you must set \$API_KEY below using information retrieved from your IBM Cloud account.
`curl --insecure -X POST --header "Content-Type: application/x-www-form-urlencoded" --header "Accept: application/json" --data-urlencode "grant_type=urn:ibm:params:oauth:grant-type-token-request"`
the above CURL request will return an auth token that you will use as \$IAM_TOKEN in the scoring request below
TODO: manually define and pass values to be scored below
`curl -X POST --header "Content-Type: application/json" --header "Accept: application/json" --header "Authorization: Bearer $IAM_TOKEN" -d '{"input_data": [{"fields": [{"name": "PASSPORT_COUNTRY", "value": "Ghana"}, {"name": "COUNTRIES_VISITED_COUNT", "value": 3}, {"name": "ARRIVAL_STATE", "value": "OH"}, {"name": "DEPARTURE_AIRPORT_COUNTRY_CODE", "value": "RU"}, {"name": "AGE", "value": 20}, {"name": "Category", "value": "Sports/Travel"}]}]}`

14. Enter the following values:

PASSPORT_COUNTRY – **Ghana**

COUNTRIES_VISITED_COUNT – **3**

ARRIVAL_STATE – **OH**

DEPARTURE_AIRPORT_COUNTRY_CODE – **RU**

AGE – **20**

Category – **Sports/Travel**

Click **Add to list**.

Deployments / AutoAI Deployment Space / FHT AutoAI - P2 XGBClassifier / FHT AutoAI Deployment

FHT AutoAI Deployment Deployed Online

API reference **Test**

Enter input data

PASSPORT_COUNTRY

Ghana

COUNTRIES_VISITED_COUNT

3

ARRIVAL_STATE

OH

DEPARTURE_AIRPORT_COUNTRY_CODE

RU

AGE

20

Category

Sports/Travel

Add to list

+

Input list (0)

The prediction list is empty.
Add one or more items to the list to request a prediction.

Predict

Result

Submit a prediction to see scoring results.

15. Click Predict (1)

Deployments / AutoAI Deployment Space / FHT AutoAI - P2 XGBClassifier / FHT AutoAI Deployment

FHT AutoAI Deployment Deployed Online

API reference **Test**

Enter input data

PASSPORT_COUNTRY

other

COUNTRIES_VISITED_COUNT

Integer

ARRIVAL_STATE

other

DEPARTURE_AIRPORT_COUNTRY_CODE

other

AGE

Integer

Category

other

Add to list

+

Input list (1)

[Ghana, 3, OH, RU, 20, Sports/Travel]

Predict (1)

Result

Submit a prediction to see scoring results.

16. The prediction for VETTING_LEVEL is **30** with 44.9 % confidence.

The screenshot displays the 'FHT AutoAI Deployment' interface. The top navigation bar shows the path: Deployments / AutoAI Deployment Space / FHT AutoAI - P2 XGBClassifier / FHT AutoAI Deployment. The main header indicates the deployment is 'Deployed' and 'Online'. Below this, there are tabs for 'API reference' and 'Test', with 'Test' being the active tab.

The 'Test' tab is divided into three main sections:

- Enter input data:** This section contains several input fields for testing the model:
 - PASSPORT_COUNTRY: other
 - COUNTRIES_VISITED_COUNT: Integer
 - ARRIVAL_STATE: other
 - DEPARTURE_AIRPORT_COUNTRY_CODE: other
 - AGE: Integer
 - Category: other
- Input list (1):** This section shows a single input example: [Ghana, 3, OH, RU, 20, Sports/Travel].
- Result:** This section displays the model's output in JSON format. The prediction is 30, and the probability is 0.4497275650501251. The prediction value '30' is circled in red.

You have successfully completed the Lab!!!

- ✓ Downloaded a female human trafficking cleansed dataset from the github repo
- ✓ Added an Auto AI Experiment to create a model to predict the trafficking risk
- ✓ Saved and Deployed the model
- ✓ Tested the model