

Lab overview -

This document is designed to give you a walkthrough for performing lab using IBM Data Science experience, Watson Machine learning and Object Storage on IBM Cloud.

The exercise makes use of financial data made available from the open source database available related to credit card customers.

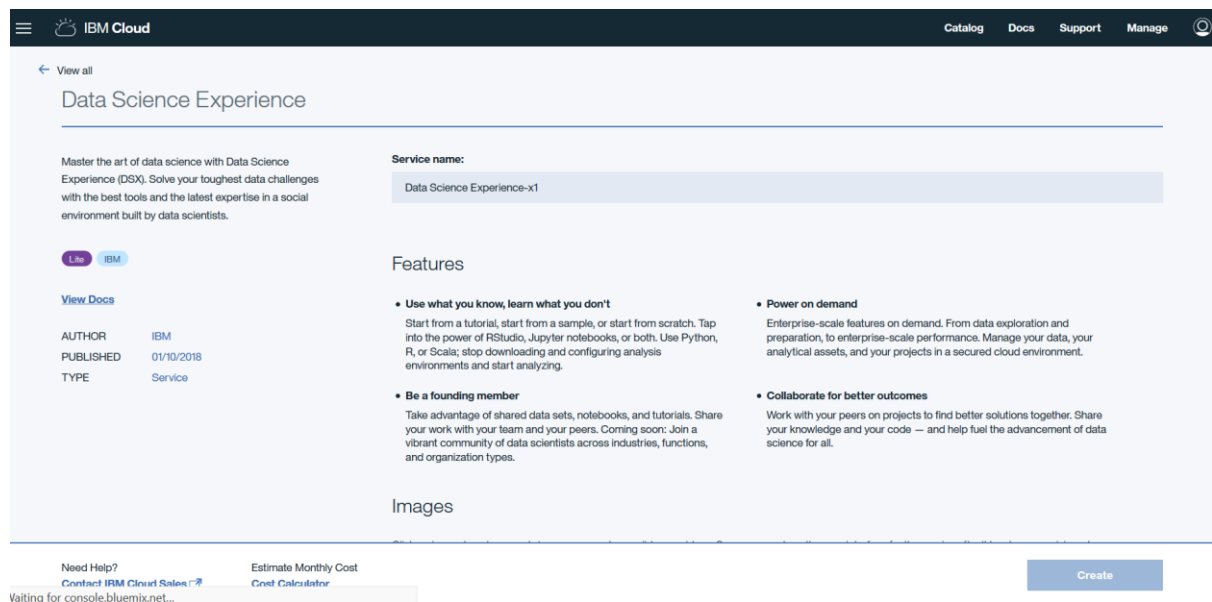
The dataset can be downloaded from - <https://github.com/IBMDevConnect/IBMCodeDay-2018>

Prerequisite -

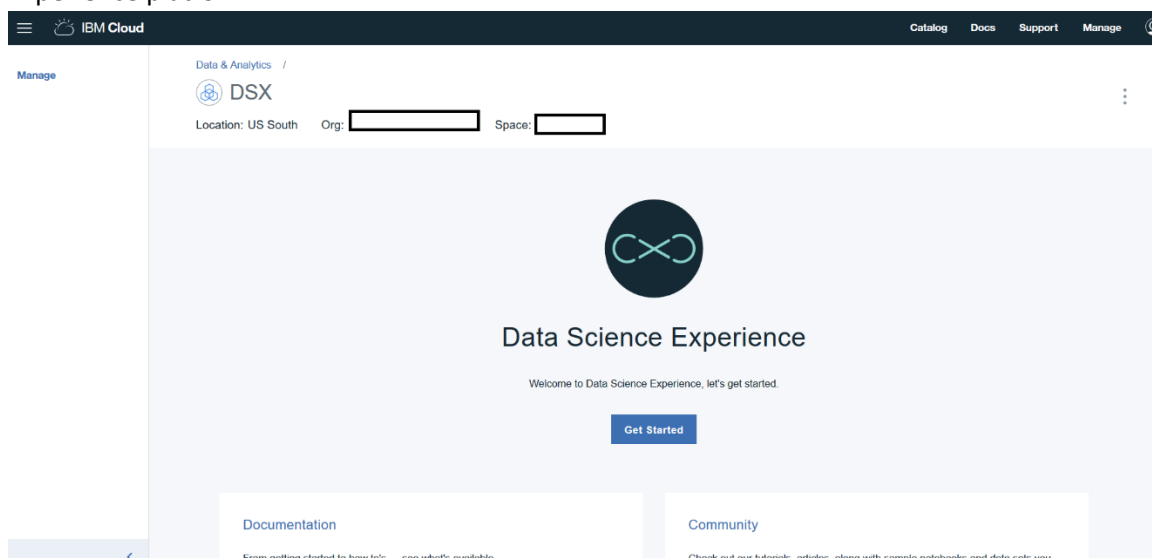
In order to complete the lab, you need to have an active account on IBM Cloud with access to US South region. (Access can be obtained by using a promo code or applying credit card details)

Steps to perform the lab-

1. Create an instance of Data Science experience service from the Catalog of IBM Cloud.



2. Once the DSX instance is created, click on Get Started to launch the IBM Data Science Experience platform.



3. As a part of configuring the platform, the DSX will assist in creating following services on IBM Cloud required by DSX –
 - a. Cloud Object Storage
 - b. Spark
 - c. Machine learning service

Please ensure that appropriate service is created with associated lite plan before proceeding further.

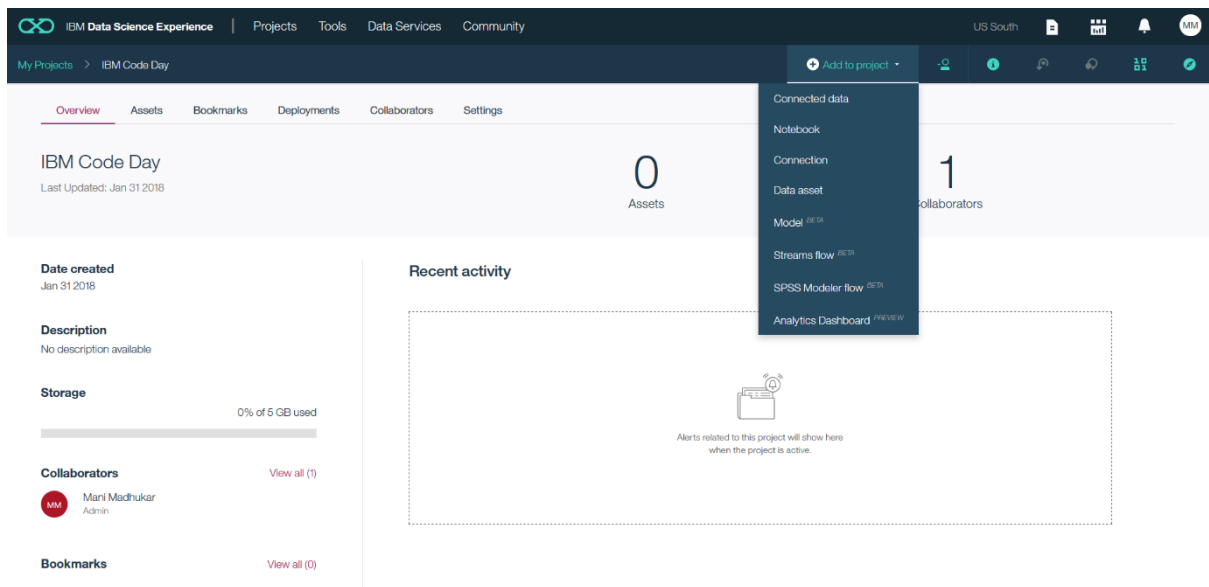
4. Create a new project to get started.

The screenshot shows the 'New project' form in the IBM Data Science Experience interface. The form is divided into two main sections: 'Define project details' and 'Define storage'. The 'Define project details' section includes a 'Name' field with the value 'IBM Code Day' and a 'Description' field with the placeholder text 'Project description'. The 'Define storage' section includes a 'Select storage type' dropdown with 'IBM Cloud Object Storage' selected, a 'Target Cloud Object Storage Instance' dropdown with 'cloud-object-storage-ck' selected, and a 'Define compute engine' section with 'Select Spark service' dropdown and 'DSX-Spark' selected. At the bottom right, there are 'Cancel' and 'Create' buttons.

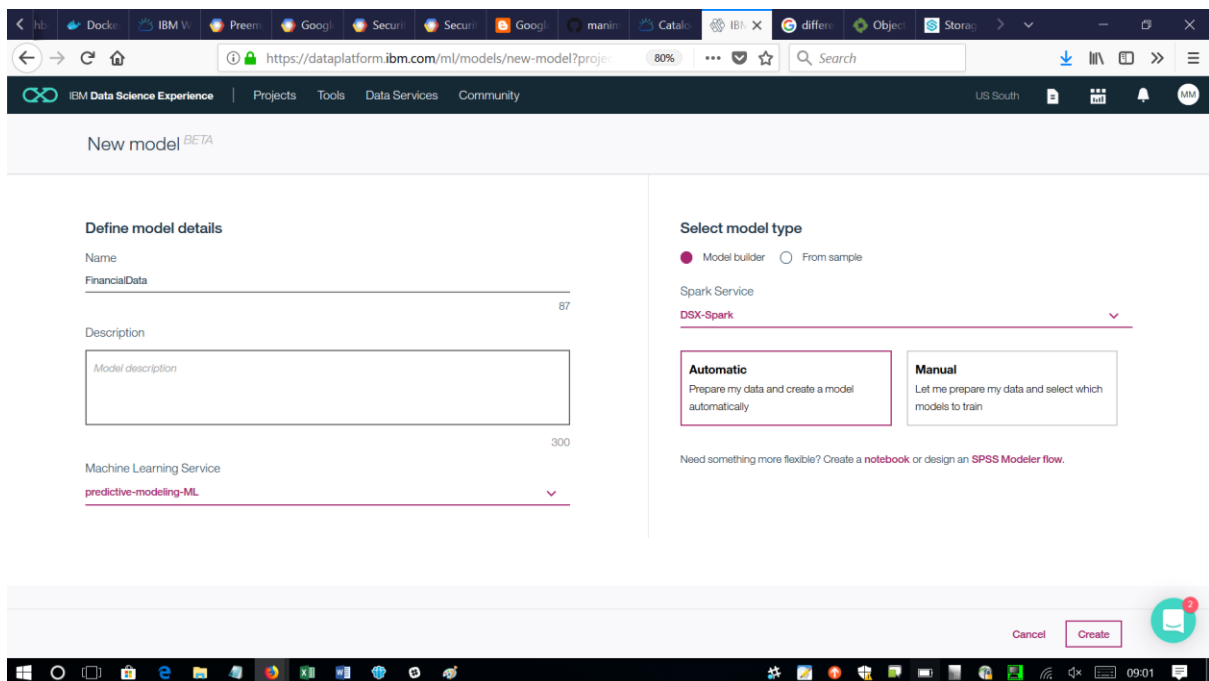
5. Once the project gets created, familiarize yourself with the interface and tabs on the platform.

The screenshot shows the project overview page for 'IBM Code Day' in the IBM Data Science Experience interface. The page has a dark blue header with the IBM Data Science Experience logo and navigation links. Below the header, there is a 'My Projects' section with a tab for 'IBM Code Day'. The main content area shows the project name 'IBM Code Day' and its last updated date 'Jan 31 2018'. There are three large numbers: '0 Assets', '0 Bookmarks', and '1 Collaborators'. Below these, there is a 'Date created' section with 'Jan 31 2018', a 'Description' section with 'No description available', a 'Storage' section, a 'Collaborators' section with 'Mani Madhukar Admin' and a 'View all (1)' link, and a 'Bookmarks' section with 'View all (0)' and a note 'You currently have 0 bookmarks'. On the right, there is a 'Recent activity' section with a placeholder for an alert.

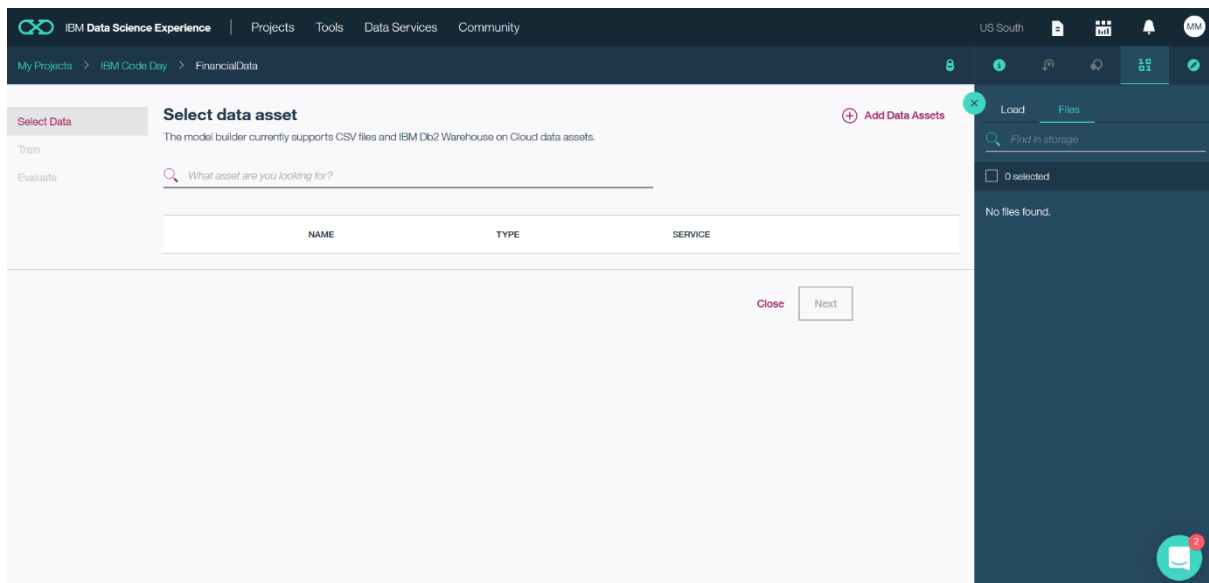
6. We are looking to leverage the model functionality provided by IBM DSX.



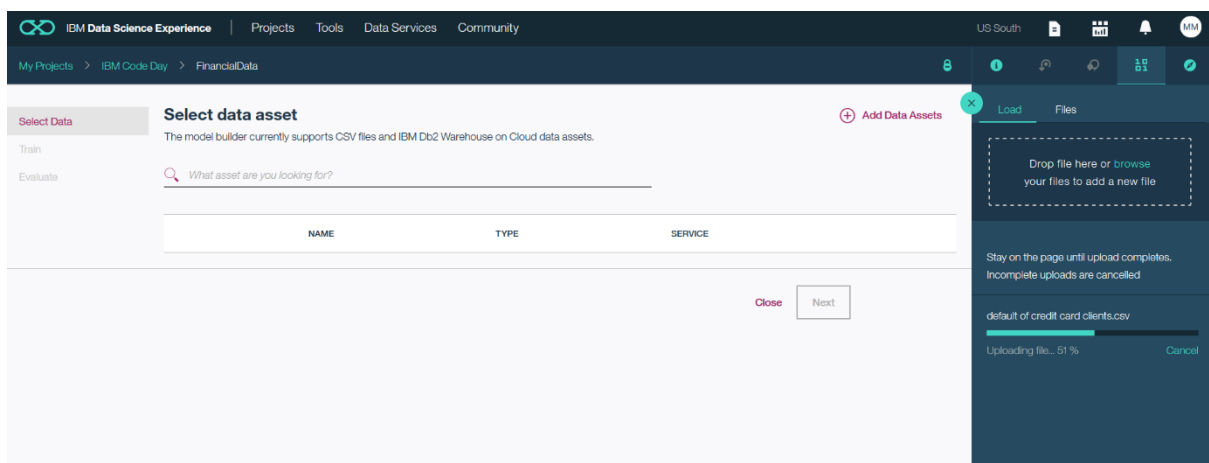
7. Define the various details like naming the model and selecting whether you would like IBM DSX to use Automatic mode or manual. For the sake of ease, we are using the Automatic model approach.



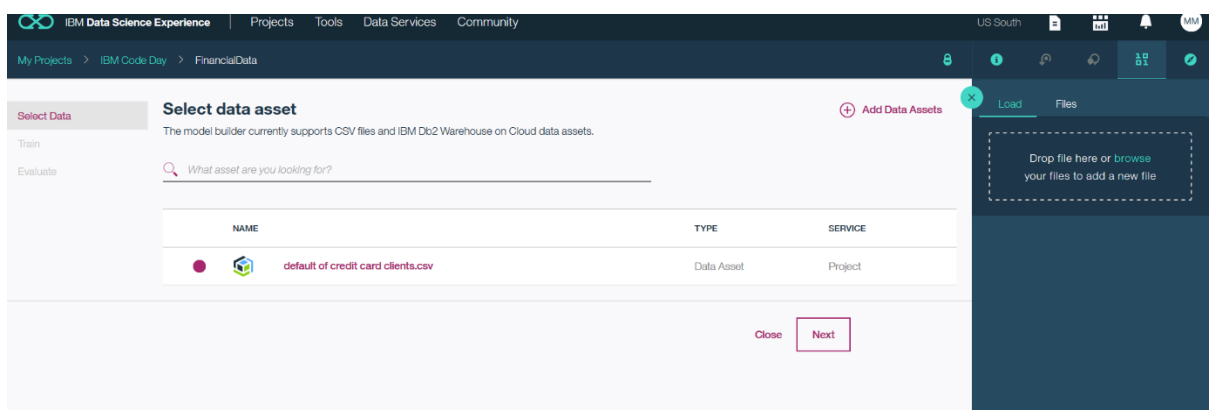
8. Add the Data source to project to build the model for. The dataset will be used to build and train the model.

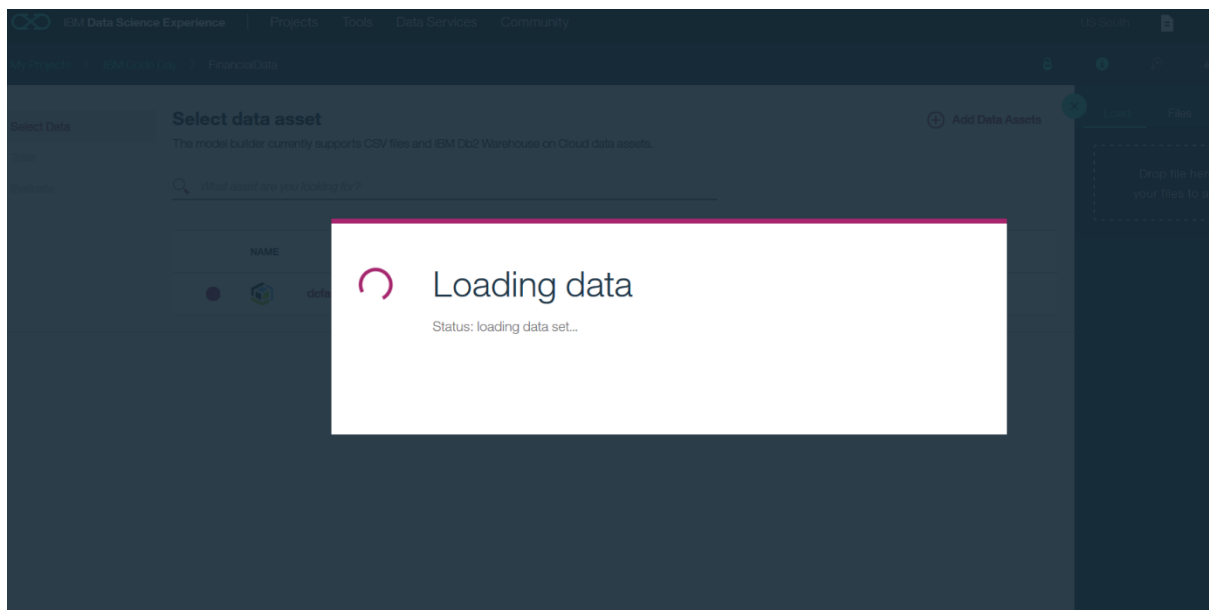


9. Notice that the data is being loaded.

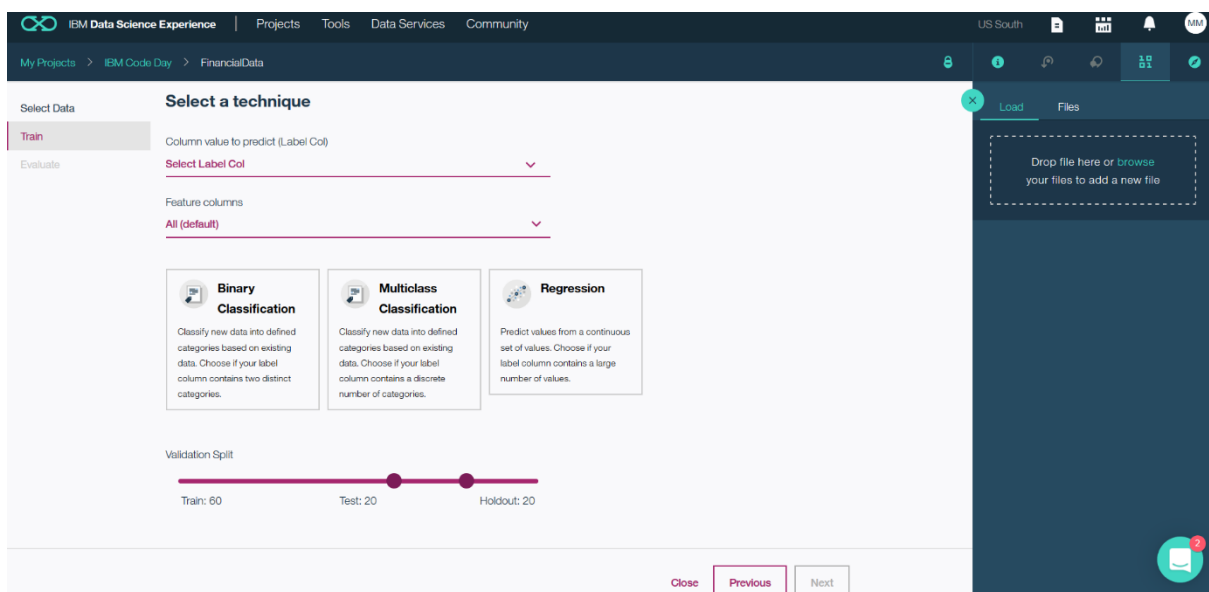


10. Once the dataset is loaded, it will be listed on the DSX platform.

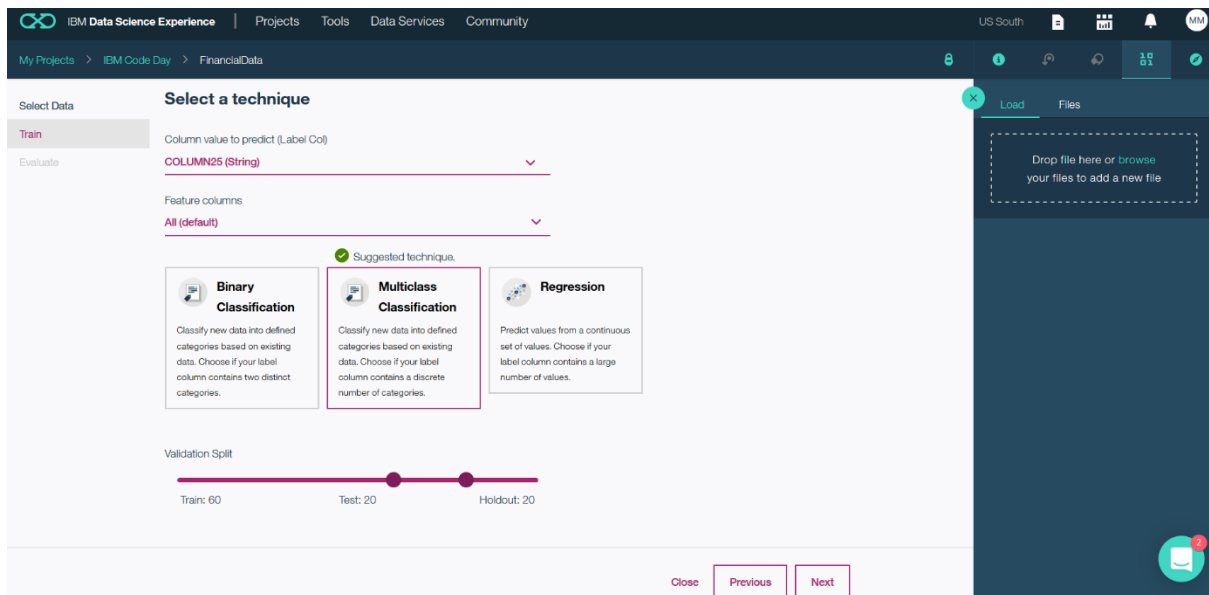




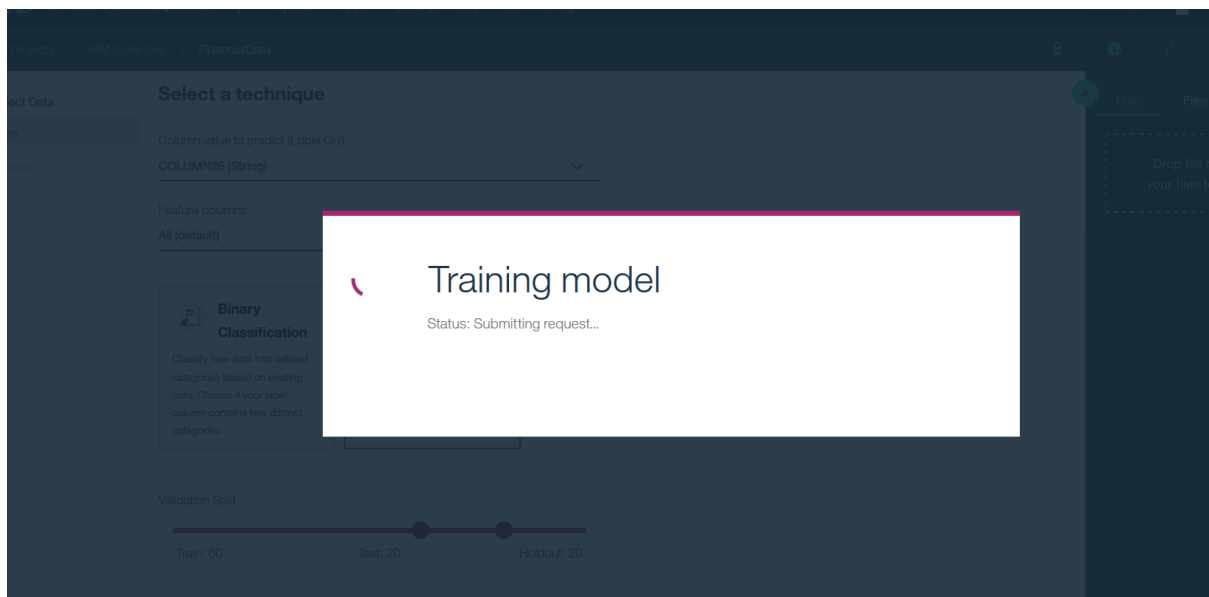
11. Select the field for which the prediction is to be made. Remember the second field refers to All default to take into account the remaining data fields.



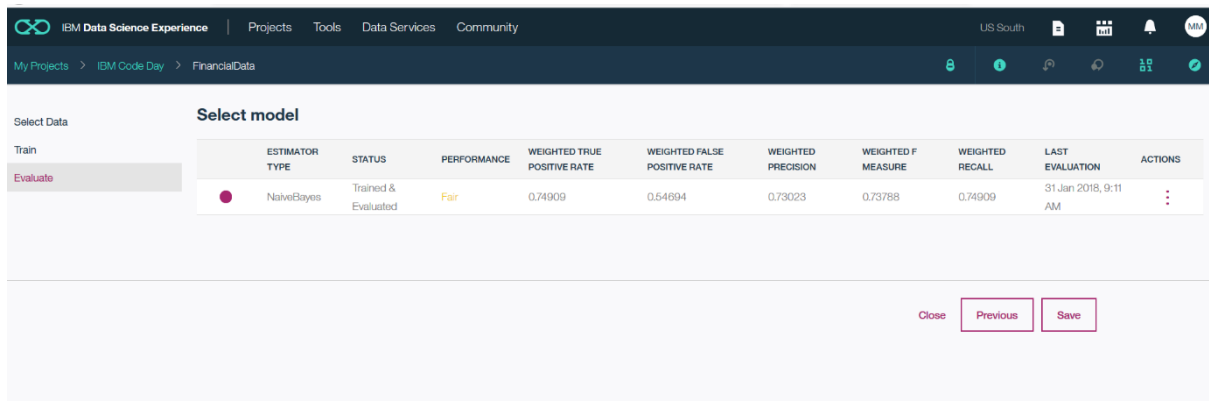
12. Also notice that the DSX, makes a suggestion for choice of algorithm to be used for analysing the dataset. Once done click on Next to proceed with training the model.



13. The IBM DSX trains the model selected with the data made available to the platform.



14. Once trained the model displays the statistics about the performance, weighed true positive rate and other parameters.

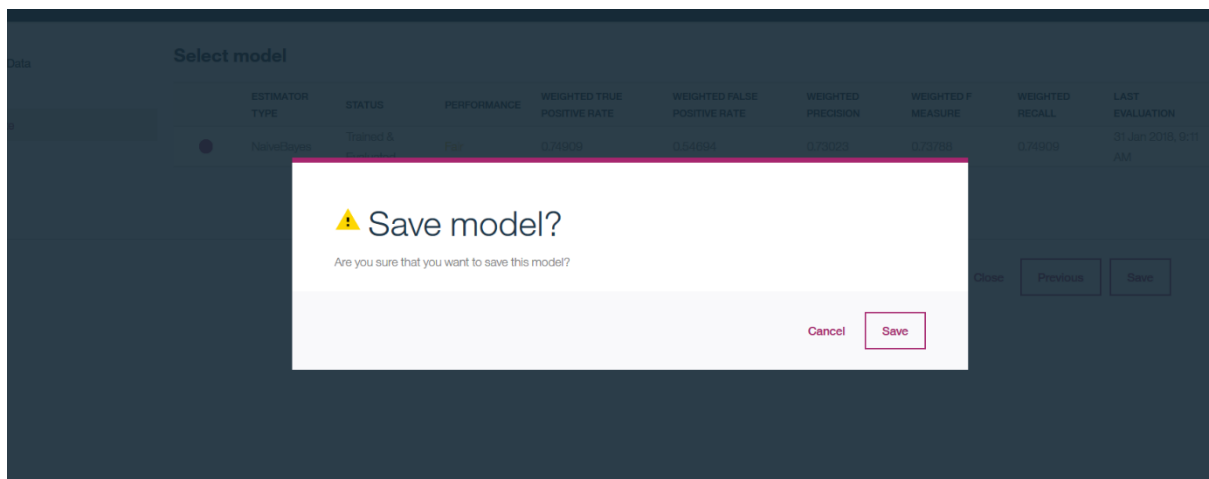


The screenshot shows the 'Select model' dialog in the IBM Data Science Experience interface. The 'Evaluate' tab is selected, displaying a table with the following data:

ESTIMATOR TYPE	STATUS	PERFORMANCE	WEIGHTED TRUE POSITIVE RATE	WEIGHTED FALSE POSITIVE RATE	WEIGHTED PRECISION	WEIGHTED F MEASURE	WEIGHTED RECALL	LAST EVALUATION	ACTIONS
NaiveBayes	Trained & Evaluated	Fair	0.74909	0.54694	0.73023	0.73788	0.74909	31 Jan 2018, 9:11 AM	

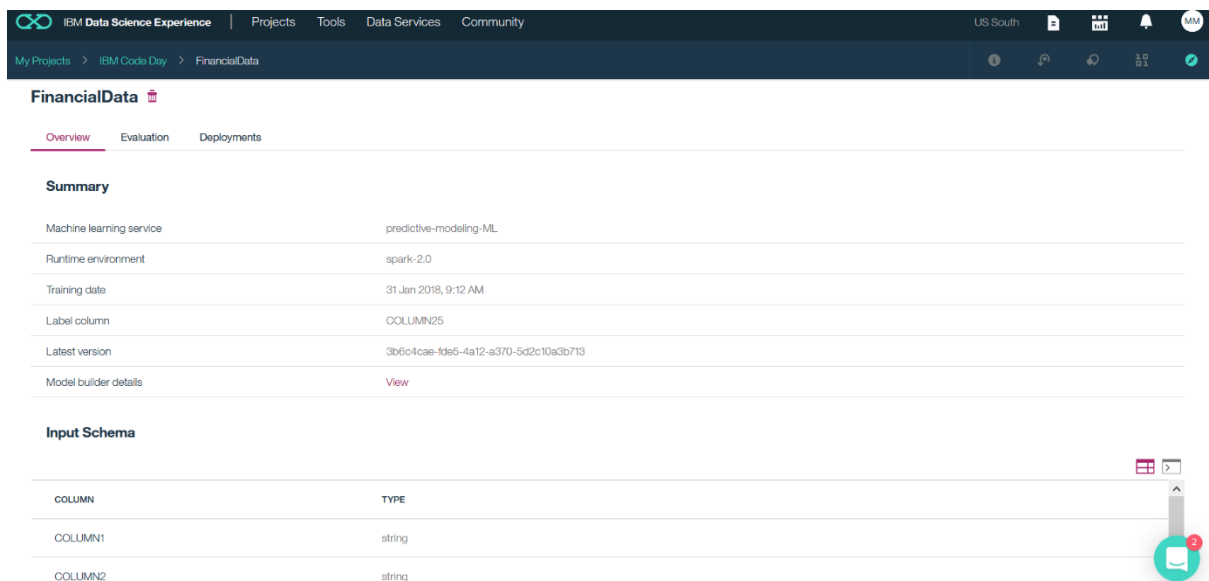
Buttons at the bottom: Close, Previous, Save.

15. Save the trained model for making predictions with test data.



The screenshot shows the 'Save model?' dialog in the IBM Data Science Experience interface. The dialog asks: 'Are you sure that you want to save this model?'. Buttons at the bottom: Cancel, Save.

16. Please notice the details about the model trained on the Overview and Evaluation tabs.



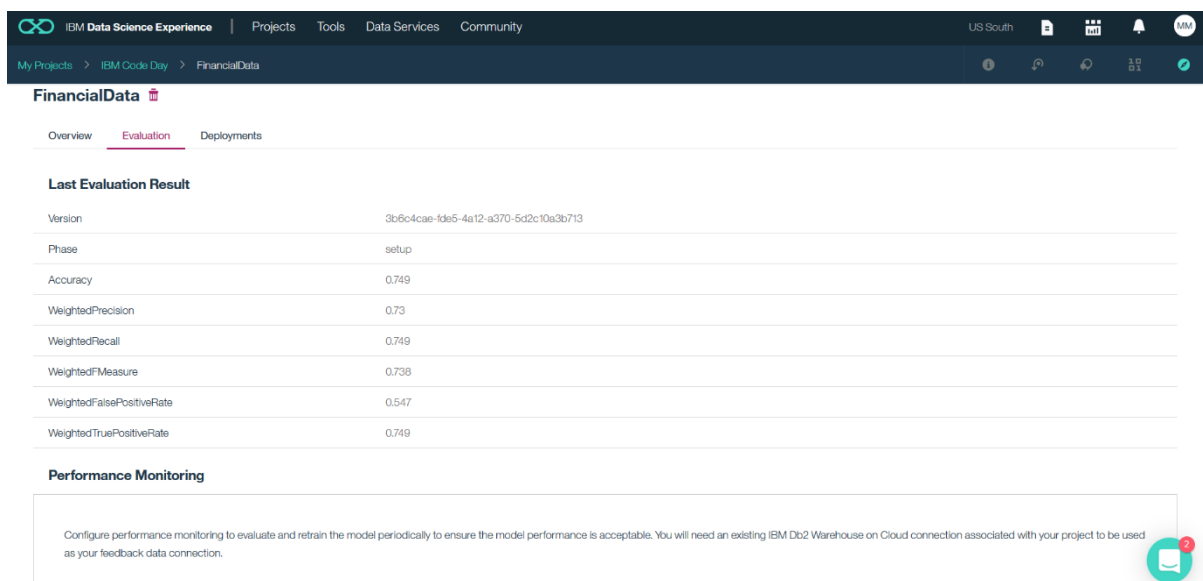
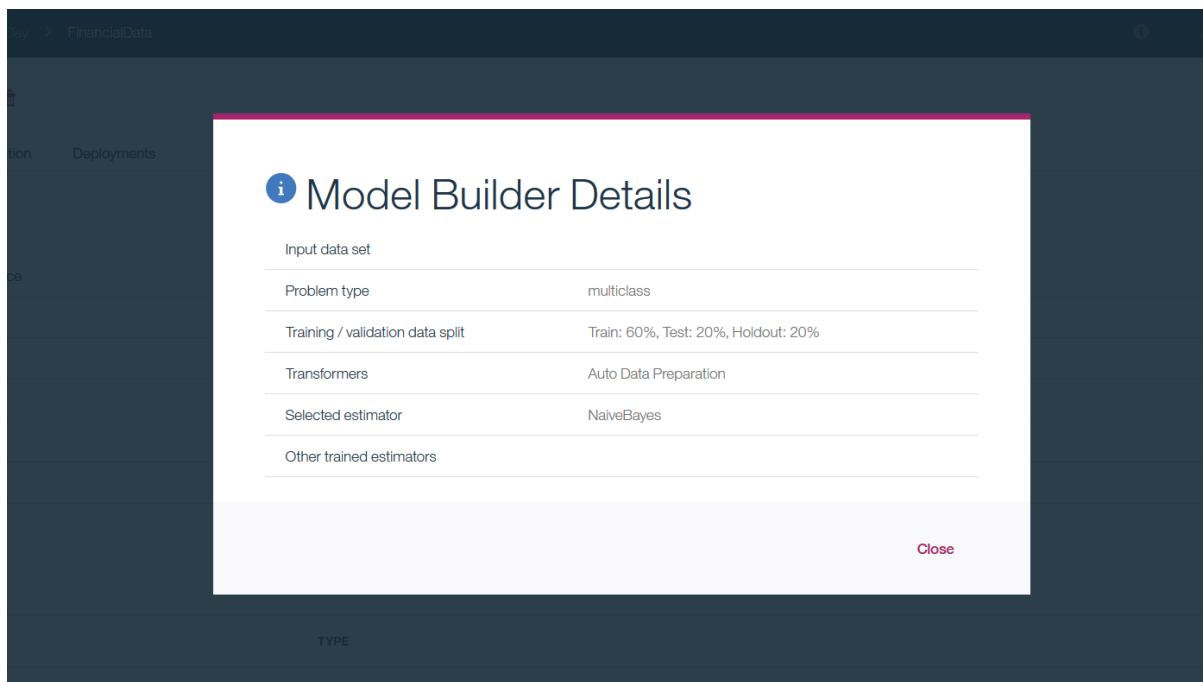
The screenshot shows the 'FinancialData' project details in the IBM Data Science Experience interface. The 'Overview' tab is selected, displaying a summary of the model and its input schema.

Summary

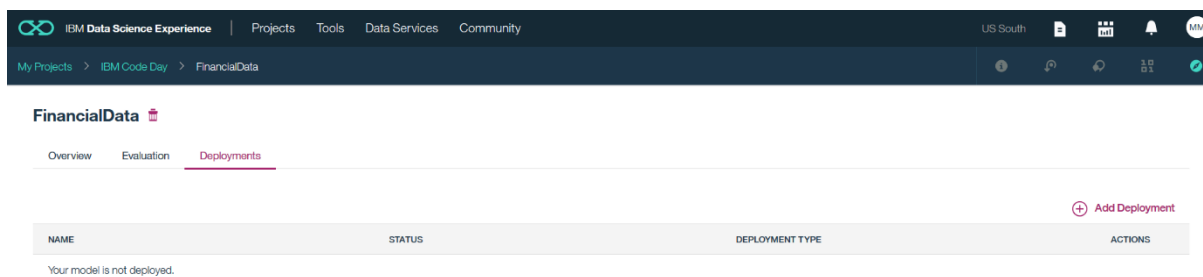
Property	Value
Machine learning service	predictive-modeling-ML
Runtime environment	spark-2.0
Training date	31 Jan 2018, 9:12 AM
Label column	COLUMN25
Latest version	3b6c4cae-1de5-4a12-a370-5d2c10a3b713
Model builder details	View

Input Schema

COLUMN	TYPE
COLUMN1	string
COLUMN2	string



17. We need to deploy the model, click on Add Deployment.



18. Select deployment as Web Service. Provide suitable name to the deployment.

IBM Data Science Experience | Projects Tools Data Services Community US South

Create Deployment

Web Service Batch Prediction Real-time Streaming Predictions

Name
Web Service Deployment Name

Description
Web Service Deployment Description

300

Cancel Save

19. Once the model is deployed as web service , wait for initialization post which it will be available as Active.

IBM Data Science Experience | Projects Tools Data Services Community US South

My Projects > IBM Code Day > FinancialData

FinancialData

Overview Evaluation Deployments

+ Add Deployment

NAME	STATUS	DEPLOYMENT TYPE	ACTIONS
FinancialDataPrediction	ACTIVE	Web Service	

20. Notice the details available on Overview and Implementation tab of the deployed model.

IBM Data Science Experience | Projects Tools Data Services Community US South

My Projects > IBM Code Day > FinancialData > FinancialDataPrediction

FinancialDataPrediction

Overview Implementation Test

Deployment

Name	FinancialDataPrediction
Type	Web Service
Deployment ID	b7547095-3509-45df-aae9-45fa060c0cde
Status	ACTIVE
Machine learning service	predictive-modeling-ML
Created	31 Jan 2018 09:15am
Last modified	31 Jan 2018 09:34am

Model

Name	FinancialData
Model ID	2bd7bc62-1e60-4c0b-9aee-266486666470
Version ID	3b6c4cae-1d46-4a12-a370-6d2c10a3b7f3

21. The model provides code snippets in various programming language to embed in application for execution.

IBM Data Science Experience | Projects Tools Data Services Community US South

My Projects > IBM Code Day > FinancialData > FinancialDataPrediction

FinancialDataPrediction

Overview **Implementation** Test

Implementation [View API Specification](#)

Scoring End-point	https://ibm-watson-ml.mybluemix.net/v3/wml_instances/4b5cddd-e0c8-49ba-995c-d06b3a76d358/published_models/2bd1bc62-1e60-4c6b-9a9c-268485656470/deployments/b7547095-3509-45df-ace9-451a050c9cde/online
Authorization: Bearer <token>	See code snippets below for information on how to retrieve the WML Authorization Token to be passed with scoring requests.
Content-type: application/json	Required if the request body is sent in JSON format.

Code Snippets

cURL Java JavaScript Python Scala

```
# retrieve your $WML_SERVICE_CREDENTIALS_USERNAME, $WML_SERVICE_CREDENTIALS_PASSWORD, and $WML_SERVICE_CREDENTIALS_URL from the
# Service credentials associated with your IBM Cloud Watson Machine Learning Service instance

curl --basic --user $WML_SERVICE_CREDENTIALS_USERNAME:$WML_SERVICE_CREDENTIALS_PASSWORD $WML_SERVICE_CREDENTIALS_URL/v3/identity/token

# the above CURL request will return an auth token that you will use as $WML_AUTH_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 $WML_AUTH_TOKEN' -d '{"fields": ["COLUMN1", "COLUMN2"]}'
```

22. Go to the Testing tab, and provide relevant data and click on Predict to see whether the test data for customer will default at the credit card payment for the forthcoming month.

IBM Data Science Experience | Projects Tools Data Services Community US South

My Projects > IBM Code Day > FinancialData > FinancialDataPrediction

FinancialDataPrediction

Overview Implementation **Test**

Enter input data

COLUMN1

COLUMN2

X1

COLUMN3

X2

COLUMN4

X3

[Predict](#)

IBM Data Science Experience | Projects Tools Data Services Community US South

My Projects > IBM Code Day > FinancialData > FinancialDataPrediction

FinancialDataPrediction

Overview Implementation **Test**

Enter input data

COLUMN21

20000

COLUMN22

20000

COLUMN23

10000

COLUMN24

X23

[Predict](#)

Predicted value for COLUMN25

1.00

20% 40% 60% 80% 100%

1 99.89%

default payment n. 0.09%

0 0.03%