

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

Aim : Using IBM Cloud, Create Auto A.I (Artificial Intelligence) using Machine Learning Instances to train a Model for various scenarios based on Cognitive Scale Cloud Industry. Demonstrate a service called automation in Artificial Intelligence (AI) to train custom models that will be utilized by the industry use case.

Steps and Screenshots :

1. Create a Machine Learning Instance on IBM Cloud

The screenshot shows the IBM Cloud catalog interface for the Watson Machine Learning service. The left sidebar lists categories like Type (Service), Provider (IBM), Last updated (04/19/2024), Category (AI / Machine Learning), and Compliance (HIPAA Enabled, IAM-enabled, Service Endpoint Supported). The main content area displays the 'Watson Machine Learning' service details. It includes a 'Create' button, a location dropdown set to 'London (eu-gb)', and a pricing plan table. The table shows a 'Lite' plan with a free price, listing features like 20 capacity unit-hours (CUH) per month and 50,000 tokens per month. The right side shows a summary panel with service name 'Watson Machine Learning', location 'London', plan 'Lite', and resource group 'Default'. At the bottom, there's a checkbox for accepting license agreements and a 'Create' button.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

2. Create Watson Studio instance in the same region as ML instance

The screenshot shows the IBM Cloud Catalog interface for creating a Watson Studio instance. The URL in the address bar is <https://cloud.ibm.com/catalog/services/watson-studio>. The left sidebar shows navigation options like Catalog, Services, and Resources. The main content area is titled "Watson Studio" and describes it as a service for developing machine learning models. It shows a "Create" button and an "About" link. A "Select a location" dropdown is set to "London (eu-gb)". On the right, there's a "Summary" panel with details: Watson Studio is free, located in London, plan is Lite, service name is "Watson Studio-ysl", and resource group is Default. Below this is a table comparing "Lite" and "Professional" plans. The "Lite" plan is free and includes 1 authorized user, 10 capacity unit-hours monthly limit, 1 vCPU + 4 GB RAM = 0.5, 2 vCPU + 8 GB RAM = 1, 4 vCPU + 16 GB RAM = 2, Decision Optimization + Watson NLP = Environment + 5, and Synthetic Data Generator, 2 vCPU + 8 GB RAM = 7 (requires Watson Machine Learning). The "Professional" plan costs ₹84.8334 INR/Capacity Unit-Hour and includes Unlimited collaborators, Unlimited elastic compute environments, Environment = # of capacity units required per hour, 1 vCPU + 4 GB RAM = 0.5, 2 vCPU + 8 GB RAM = 1, 4 vCPU + 16 GB RAM = 2, and Watson Machine Learning. At the bottom, there's a checkbox for accepting license agreements, a "Create" button, and an "Add to estimate" button.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

3. Provide information required

Preparing IBM Watson Studio... | IBM Watson Studio and 13 more pages - College - Microsoft Edge

https://eu-gb.dataplatform.cloud.ibm.com/registration/steptwo?context=cpdaas&apps=data_science_experience&sync_account_id=e147b91f...

IBM Watson Studio

Provide your information to continue

User account
Yash Lakhtariya's Account

Resource group
Default

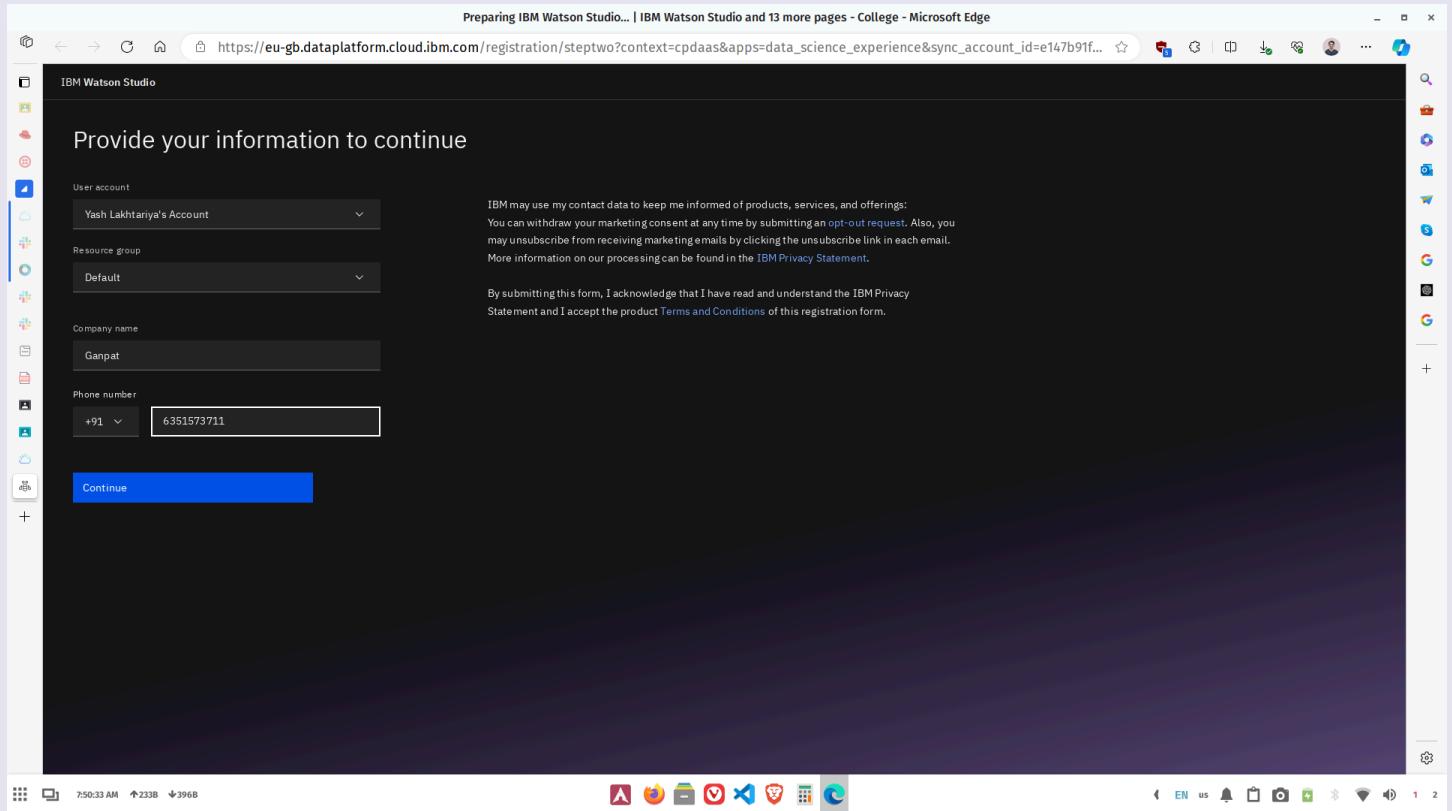
Company name
Ganpat

Phone number
+91 6351573711

IBM may use my contact data to keep me informed of products, services, and offerings.
You can withdraw your marketing consent at any time by submitting an [opt-out request](#). Also, you
may unsubscribe from receiving marketing emails by clicking the unsubscribe link in each email.
More information on our processing can be found in the [IBM Privacy Statement](#).

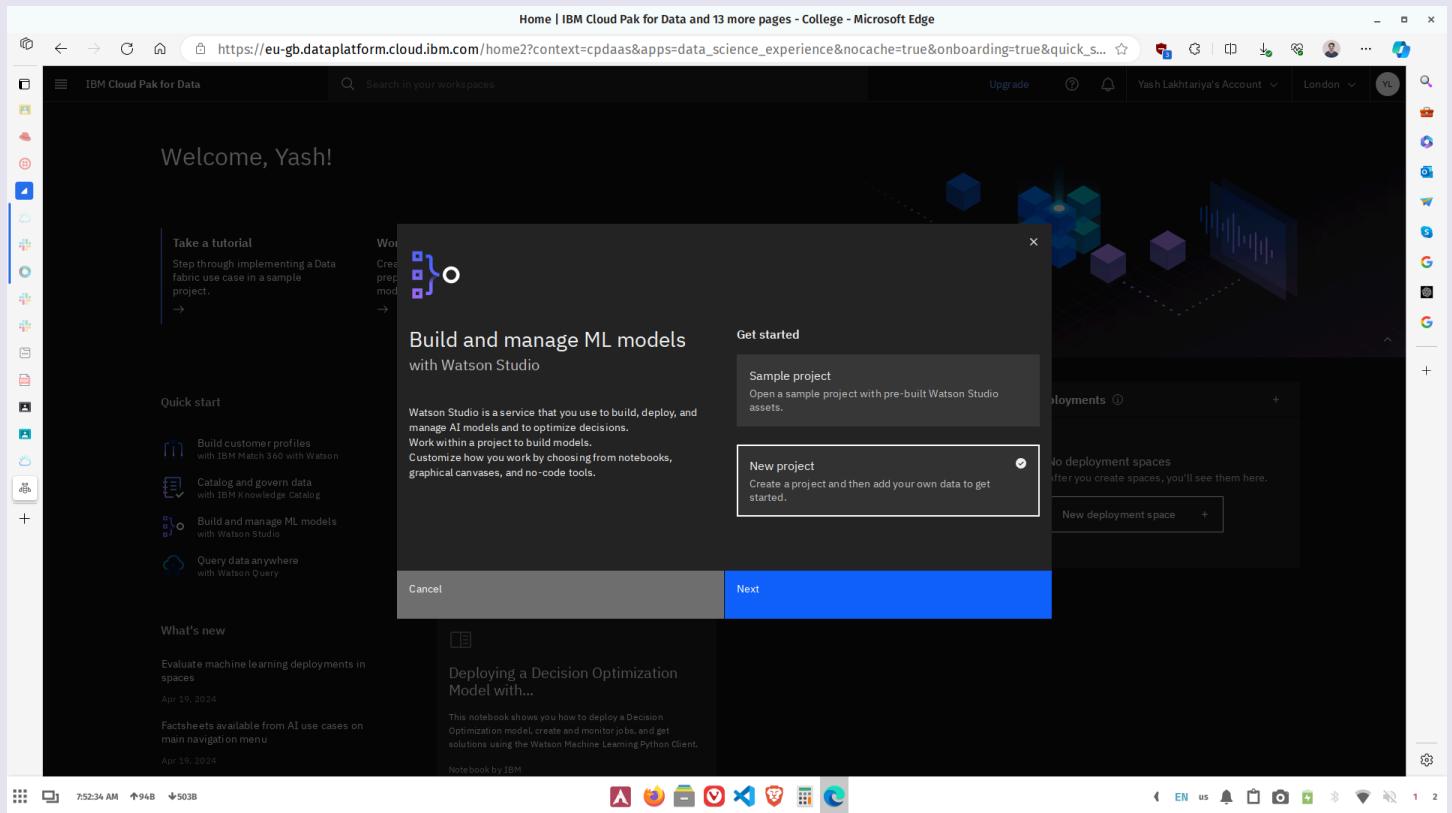
By submitting this form, I acknowledge that I have read and understand the IBM Privacy
Statement and I accept the product [Terms and Conditions](#) of this registration form.

Continue



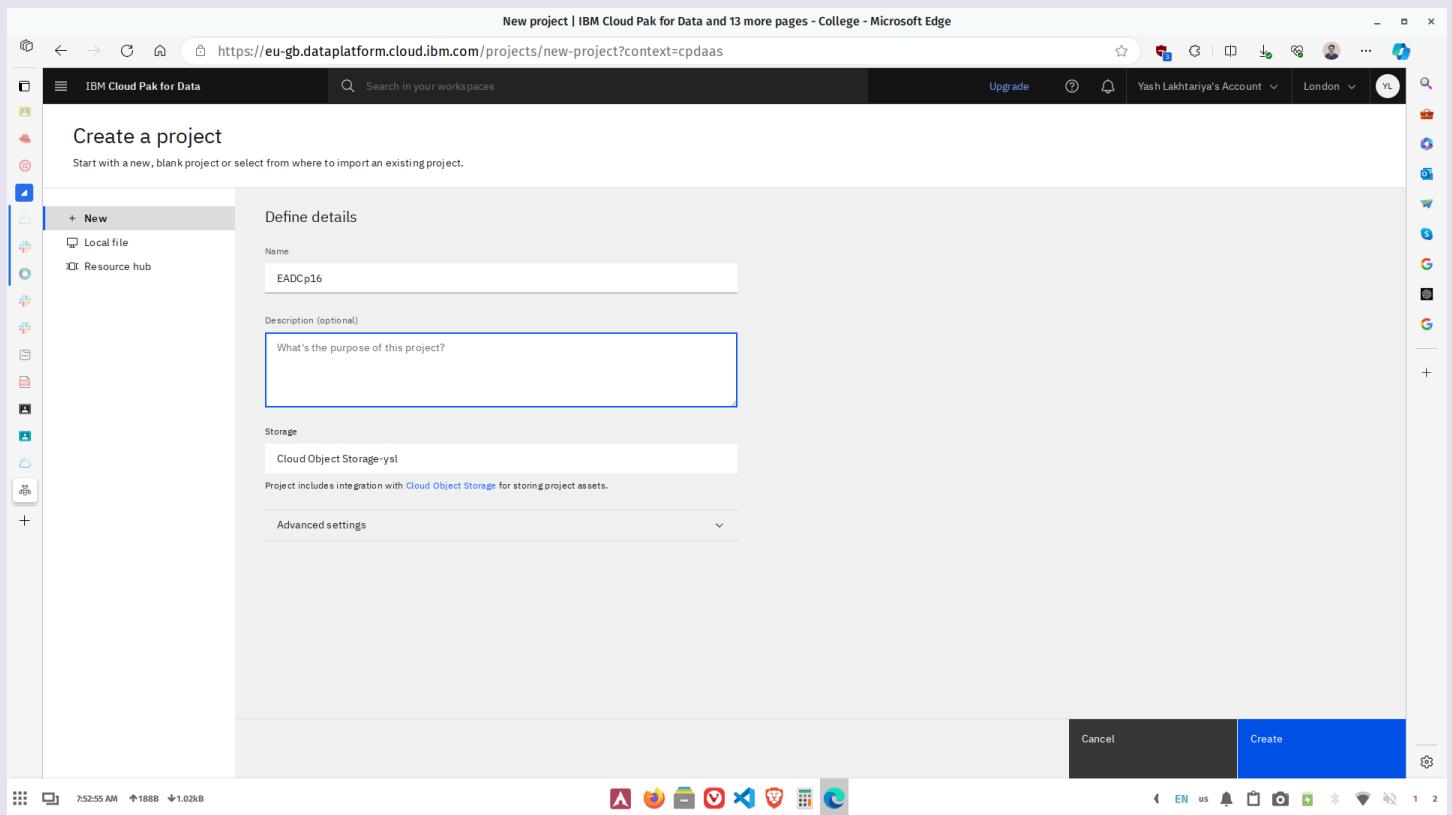
Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

4. Create new project in Watson Studio



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

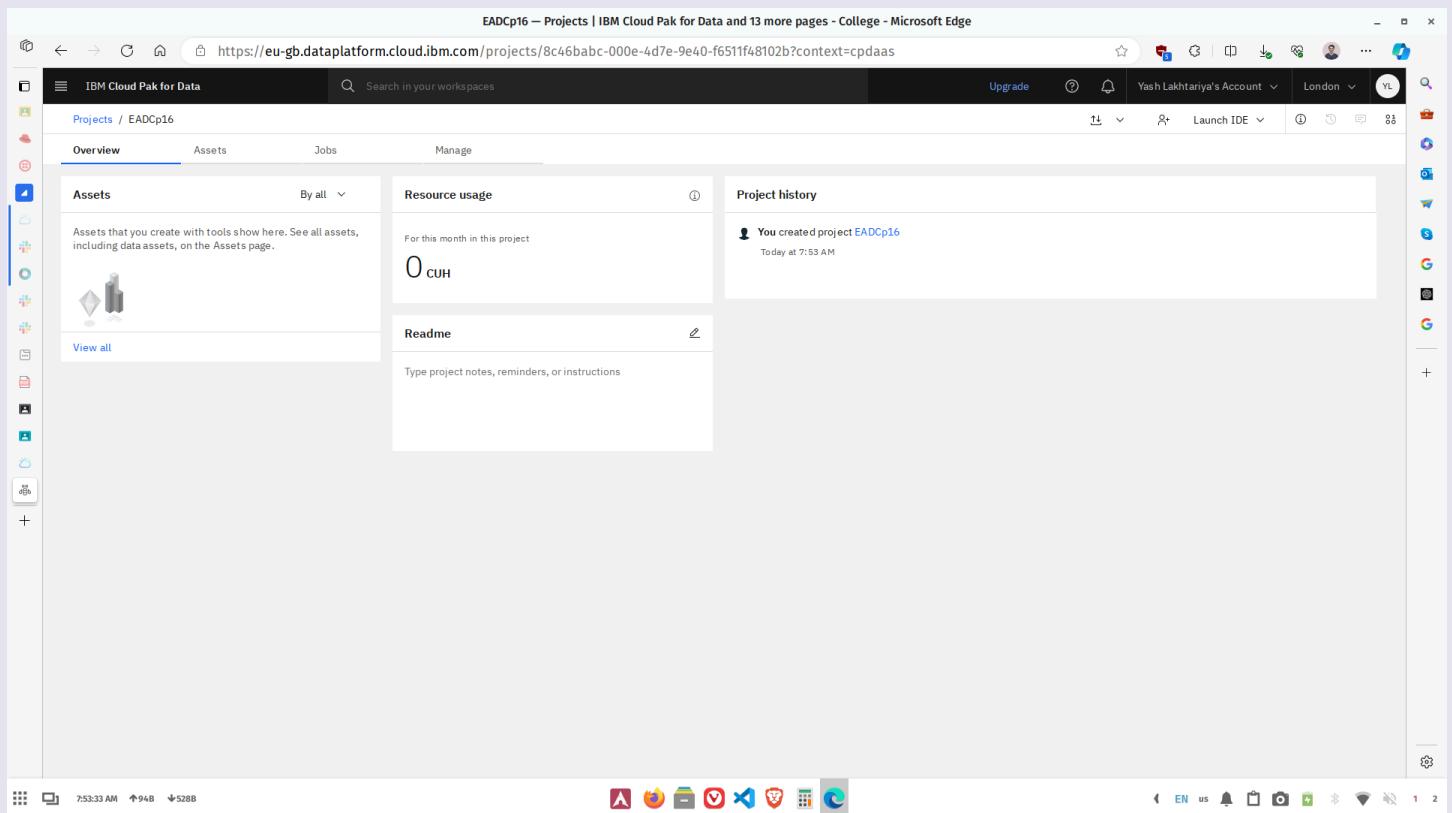
5. Assign the name and cloud object storage instance (required)



The screenshot shows the 'Create a project' page in the IBM Cloud Pak for Data interface. The URL in the address bar is <https://eu-gb.dataplatform.cloud.ibm.com/projects/new-project?context=cpdaas>. The page title is 'New project | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge'. On the left, there's a sidebar with various icons and a 'Create a project' section. The main area has a 'Define details' form. In the 'Name' field, 'EADCp16' is entered. Below it, there's a 'Description (optional)' field with the placeholder 'What's the purpose of this project?'. Under the 'Storage' section, 'Cloud Object Storage-ysl' is selected. At the bottom right, there are 'Cancel' and 'Create' buttons.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

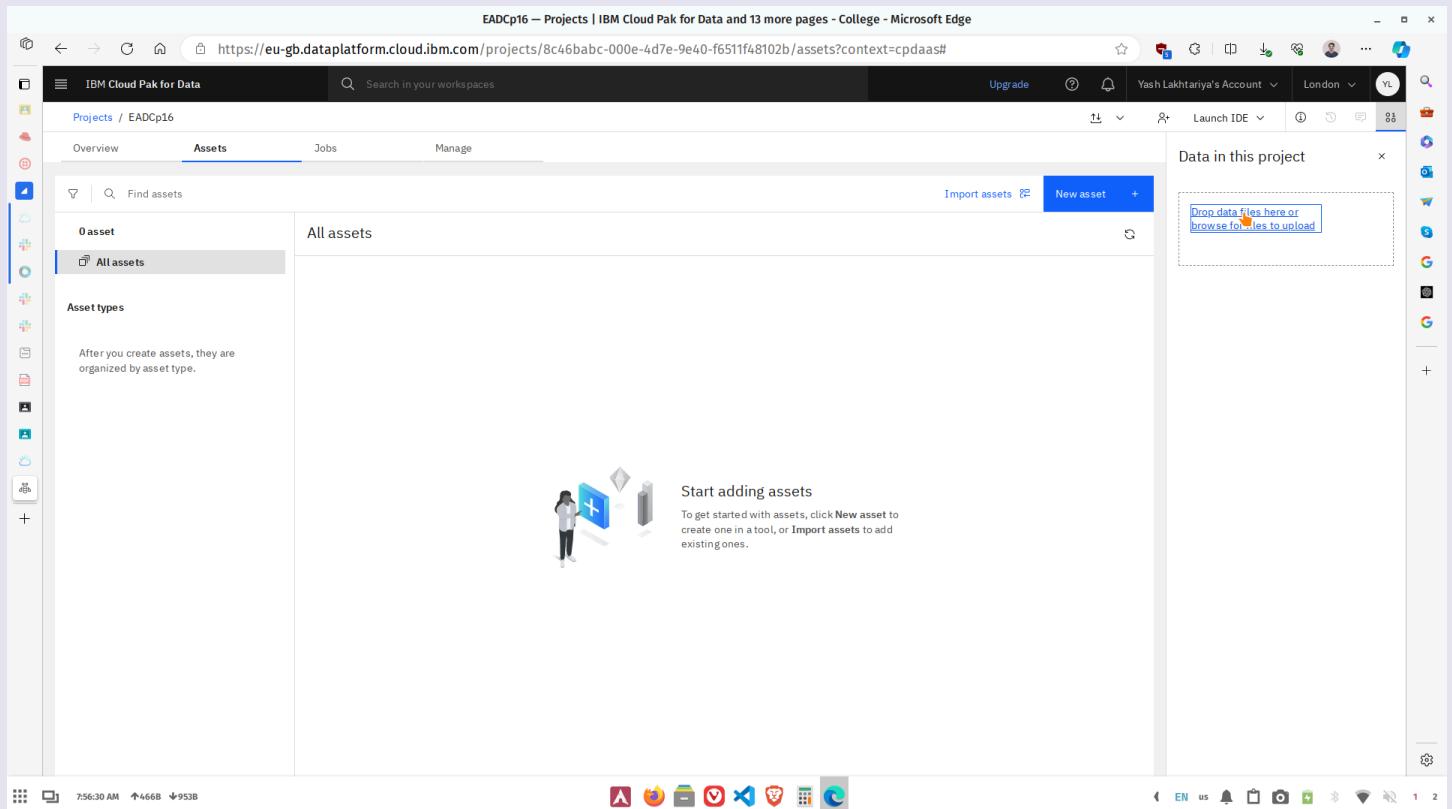
6. Go to Assets tab



The screenshot shows the 'EADCp16' project page in the IBM Cloud Pak for Data interface. The top navigation bar includes links for 'Upgrade', 'Yash Lakhtariya's Account', 'London', and 'Launch IDE'. The left sidebar has a 'Projects' section showing 'EADCp16' and various tool icons. The main content area has tabs for 'Overview', 'Assets', 'Jobs', and 'Manage', with 'Overview' selected. The 'Assets' section displays a message: 'Assets that you create with tools show here. See all assets, including data assets, on the Assets page.' It features a small icon of a diamond and a cylinder. Below this is a 'View all' button. The 'Resource usage' section shows '0 CUH' for the current month. The 'Project history' section shows a log entry: 'You created project EADCp16 Today at 7:53 AM'. The bottom of the screen shows the Windows taskbar with various pinned icons and system status.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

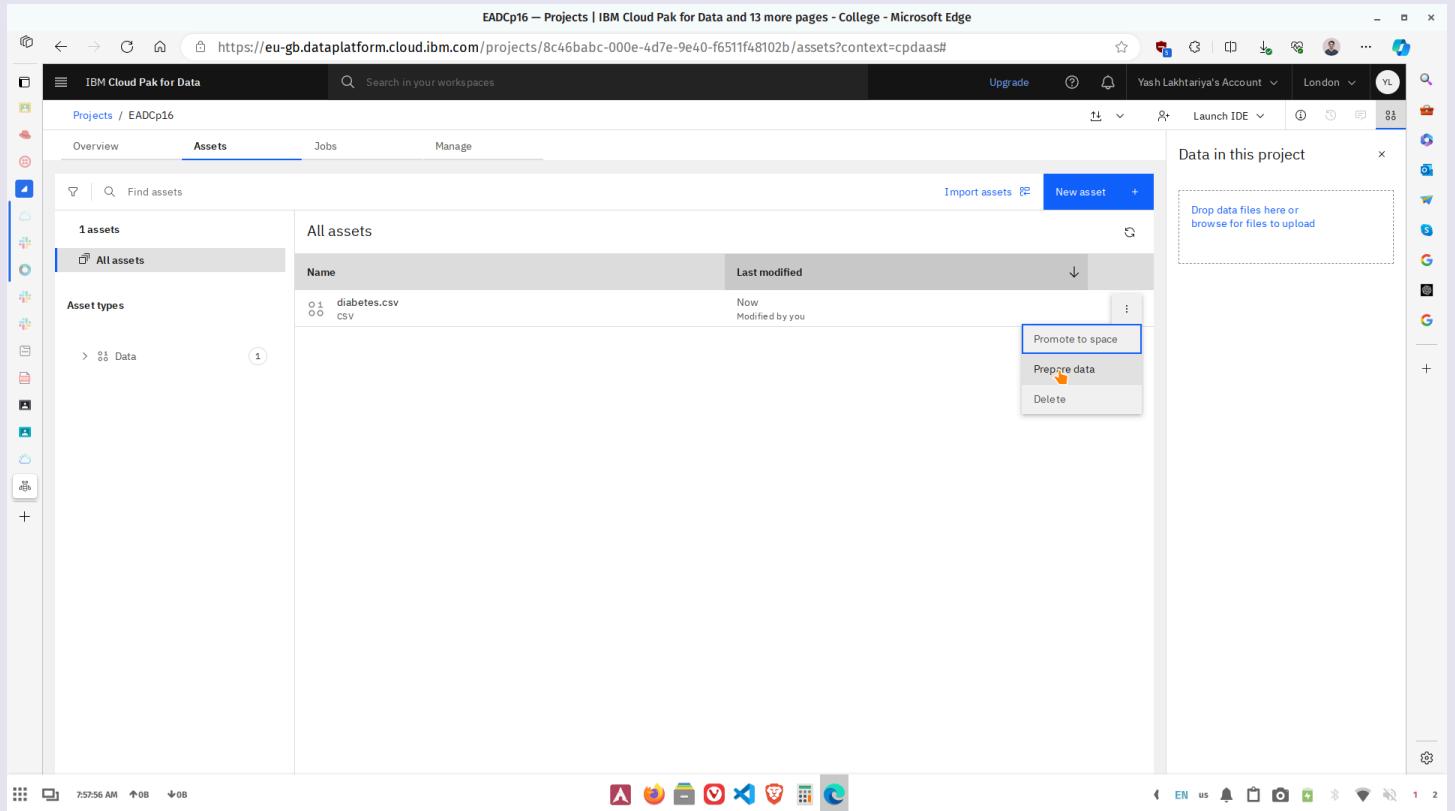
7. Drop file or upload the file (csv file here)



The screenshot shows the 'Assets' tab of the IBM Cloud Pak for Data interface. On the right side, there is a large dashed box labeled 'Data in this project' with the instruction 'Drop data files here or browse for files to upload'. The rest of the page is mostly empty, indicating no assets have been uploaded yet.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

8. Now, for data preprocessing, click Prepare Data option



The screenshot shows the IBM Cloud Pak for Data interface in Microsoft Edge. The URL is <https://eu-gb.dataplatform.cloud.ibm.com/projects/8c46babc-000e-4d7e-9e40-f6511f48102b/assets?context=cpdaas#>. The page title is "EADCp16 – Projects | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge". The left sidebar shows "IBM Cloud Pak for Data" with "Projects / EADCp16" selected. The main area is the "Assets" tab, showing "All assets" with one item: "diabetes.csv" (CSV). A context menu is open over this file, with "Prepare data" highlighted by a red arrow. Other options in the menu include "Promote to space" and "Delete". The right sidebar shows "Data in this project" with a placeholder "Drop data files here or browse for files to upload". The bottom status bar shows the time as 7:57:56 AM.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

9. Create New Step and select change column type

The screenshot shows the Data Refinery interface in IBM Cloud Pak for Data. On the left, a sidebar lists various operations under categories like CLEANSE, COMPUTE, and ORGANIZE. The main area displays a dataset named 'diabetes.csv' with 768 rows and 9 columns. The columns are: Pregnancies, Glucose, BloodPress..., SkinThickn..., Insulin, and BMI. The data preview shows numerical values for each row across these columns. To the right, there's a panel titled 'About this asset' with details such as Name (diabetes.csv_flow), Description (What is the purpose of this Data Refinery flow?), Asset details (Steps: 1), Associated assets (Source: diabetes.csv, Target: diabetes_csv_shaped), and Last modified (Not yet saved). The bottom status bar indicates the current time (8:01:32 AM) and system status (68.4kB/222kB).

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

10. Here, Outcome is changed from integer to boolean

The screenshot shows the IBM Cloud Pak for Data Data Refinery interface. The URL in the address bar is https://eu-gb.dataplatform.cloud.ibm.com/shaper?project_id=8c46babc-000e-4d7e-9e40-f6511f48102b&dataset_id=318f6237-6564-40ed-a5bb-b972.... The main window displays a 'Data Refinery' step titled 'CONVERSION 1'. The 'Outcome' column is selected for conversion, and its current type is 'Integer'. A dropdown menu is open over the 'Type' field, showing options like 'Boolean', 'Date', 'Double', etc., with 'Boolean' highlighted. The interface includes tabs for 'Data', 'Profile', and 'Visualizations'. On the right side, there's an 'About this asset' panel with details such as 'Name: diabetes.csv_flow', 'Description: What is the purpose of this Data Refinery flow?', 'Asset details: Steps: 1', and 'Associated assets: Source: diabetes.csv, Target: diabetes_csv_shaped'. The bottom status bar shows the system time as 8:03:19 AM and battery level at 94B.

**Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16**

IBM Cloud Pak for Data

Projects / EADCp16 / diabetes.csv / Data Refinery

All Operations / Convert column type

Convert the data type of the columns to a different data type.

Automatically convert the data to inferred data types

Select the columns and data types to convert.

CONVERSION 1

The 'Integer' data type most closely matches the column's data. Convert the column to a different data type.

Column Type

Outcome Boolean

Create a new column for results

Select column

Cancel Apply

Use a code template to add a step

Data Profile Visualizations

Outcome Integer

1
0
1
0
1
0
1
1
0
0
1
0
1
1
1
1
0
1
1
1
1
0
1

Configure Viewing: 768 rows, 9 columns Full dataset: 768 rows, 9 columns

About this asset

Name diabetes.csv_flow Data Refinery flow

Description What is the purpose of this Data Refinery flow?

Asset details Steps: 1

Associated assets Source: diabetes.csv Target: diabetes_csv_shaped

Last modified Not yet saved Created on Not yet saved

8:04:39 AM ↑1888 ↓08

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

11. Now click Save to save the new flow

Data Refinery — EADCp16 / diabetes.csv | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge

https://eu-gb.dataplatform.cloud.ibm.com/shaper?project_id=8c46babc-000e-4d7e-9e40-f6511f48102b&dataset_id=318f6237-6564-40ed-a5bb-b972...

IBM Cloud Pak for Data

Projects / EADCp16 / diabetes.csv / Data Refinery

Steps (1)

Use a code template to add a step

Data source: diabetes.csv

1. Convert column type

Automatically converted one or more columns to inferred data types. Strings that are converted to decimal use a dot(.) for the decimal symbol.

Auto-generated

Data Profile Visualizations

	Pregnancies	Glucose	BloodPress...	SkinThickn...	Insulin	BMI	DiabetesPe...
	Integer	Integer	Integer	Integer	Integer	Decimal	Decimal
1	6	148	72	35	0	33.6	0.627
2	1	85	66	29	0	26.6	0.351
3	8	183	64	0	0	23.3	0.672
4	1	89	66	23	94	28.1	0.167
5	0	137	40	35	168	43.1	2.288
6	5	116	74	0	0	25.6	0.201
7	3	78	50	32	88	31	0.248
8	10	115	0	0	0	35.3	0.134
9	2	197	70	45	543	30.5	0.158
10	8	125	96	0	0	0	0.232
11	4	110	92	0	0	37.6	0.191
12	10	168	74	0	0	38	0.537
13	10	139	80	0	0	27.1	1.441
14	1	189	60	23	846	30.1	0.398
15	5	166	72	19	175	25.8	0.587
16	7	100	0	0	0	30	0.484
17	0	118	84	47	230	45.8	0.551
18	7	107	74	0	0	29.6	0.254
19	1	103	30	38	83	43.3	0.183
20	1	115	70	30	96	34.6	0.529

New step +

Configure Viewing: 768 rows, 9 columns Full data set: 768 rows, 9 columns

About this asset

Name: diabetes.csv_flow Data Refinery flow

Description: What is the purpose of this Data Refinery flow?

Asset details: Steps: 1

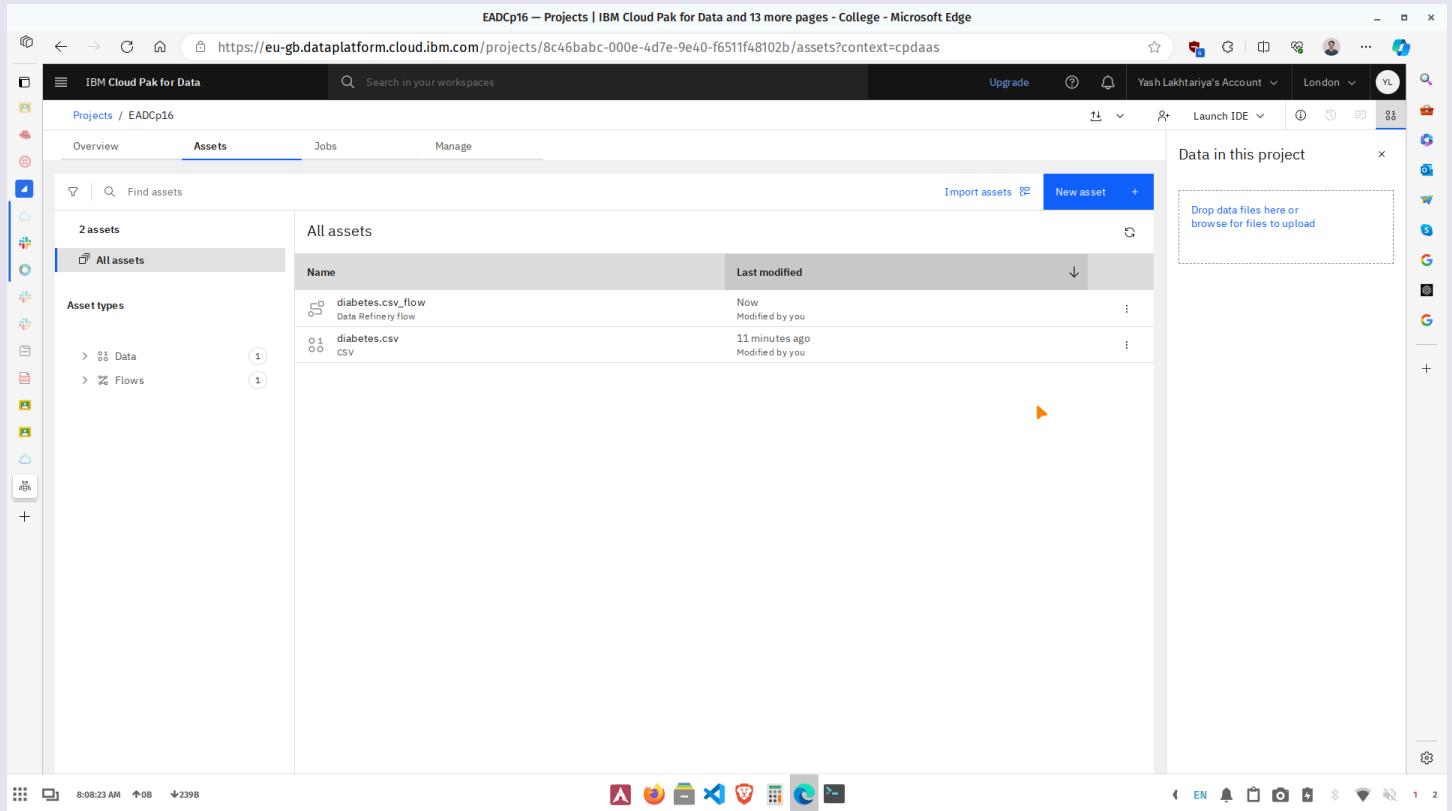
Associated assets: Source: diabetes.csv Target: diabetes_csv_shaped

Last modified: Not yet saved Created on: Not yet saved

8:07:14 AM 8:08 2.09KB EN

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

12. As seen, the new flow is created



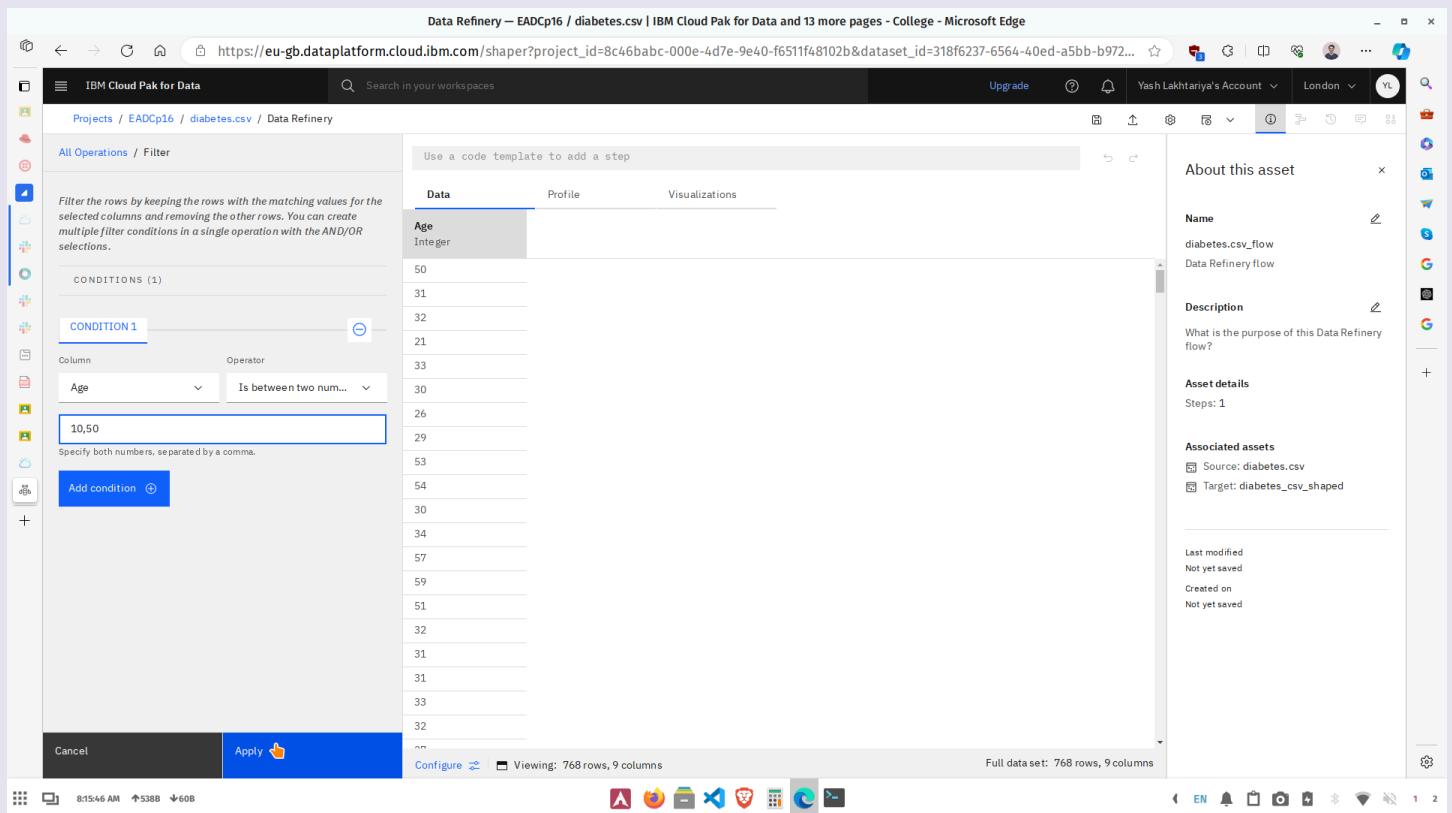
The screenshot shows the 'Assets' tab in the IBM Cloud Pak for Data interface. On the left sidebar, under 'Asset types', there are categories for 'Data' and 'Flows'. The main area displays a table titled 'All assets' with two entries:

Name	Last modified
diabetes.csv_flow Data Refinery flow	Now Modified by you
diabetes.csv CSV	11 minutes ago Modified by you

A large orange arrow points from the bottom right towards the 'diabetes.csv' asset entry.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

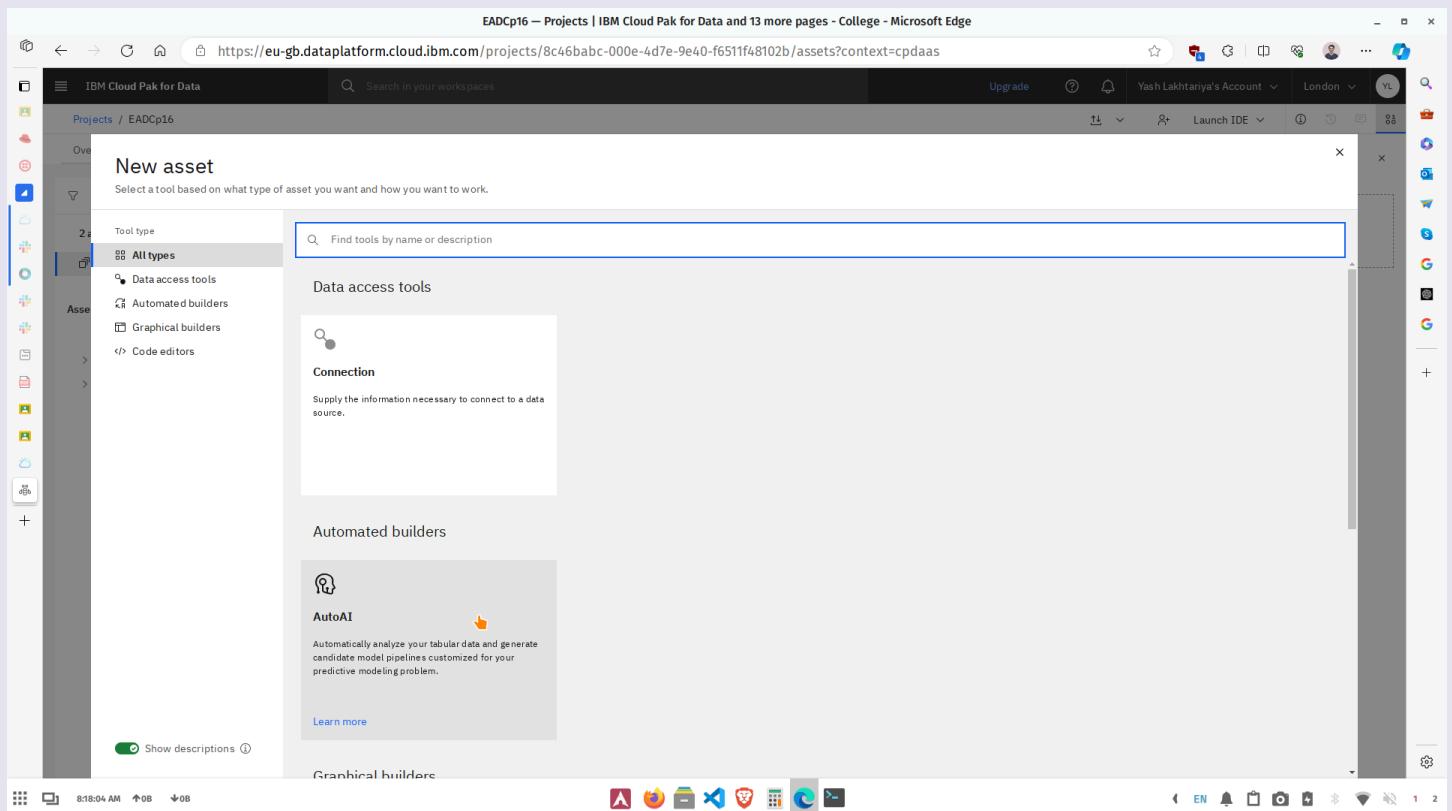
13. Other conditions can also be added like age filter



The screenshot shows the IBM Cloud Pak for Data Data Refinery interface. On the left, there's a sidebar with various icons and a search bar. The main area has tabs for 'Data', 'Profile', and 'Visualizations', with 'Data' selected. A large list of integers from 50 down to 31 is displayed. On the left, under 'CONDITIONS (1)', there's a section for 'CONDITION 1' where 'Age' is set to 'Is between two numbers...' with '10,50' entered. At the bottom, there are 'Cancel' and 'Apply' buttons. To the right, there's a panel titled 'About this asset' with details like Name (diabetes.csv_flow), Description (What is the purpose of this Data Refinery flow?), Asset details (Steps: 1), Associated assets (Source: diabetes.csv, Target: diabetes_csv_shaped), and Last modified (Not yet saved). The status bar at the bottom shows the time as 8:15:46 AM.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

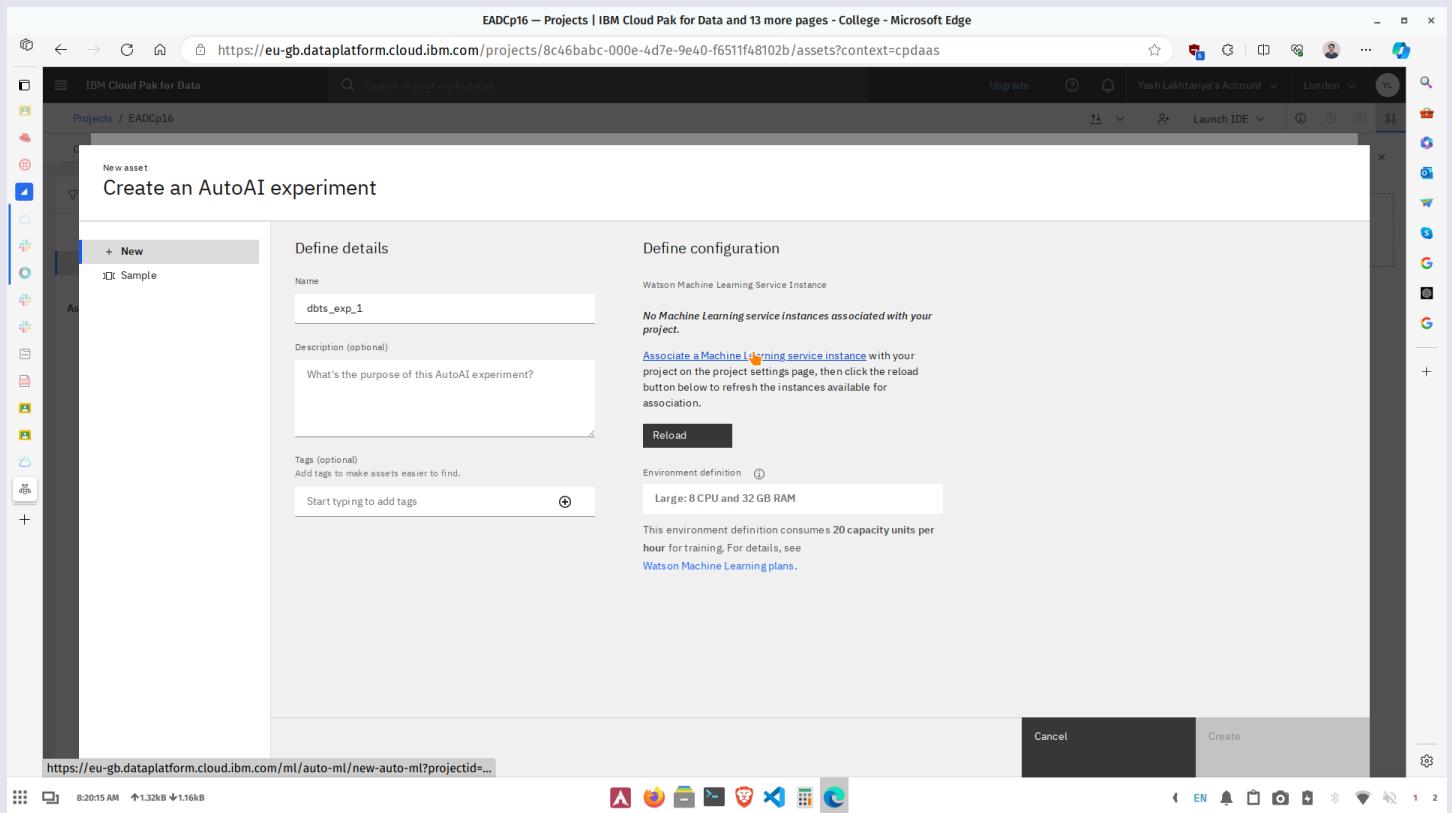
14. Now, click New Asset to add asset and choose AutoAI



The screenshot shows the 'New asset' creation interface in the IBM Cloud Pak for Data web application. The left sidebar lists categories like 'Tool type', 'Data access tools', 'Automated builders', 'Graphical builders', and 'Code editors'. The main area is titled 'New asset' with the sub-instruction 'Select a tool based on what type of asset you want and how you want to work.' A search bar at the top right says 'Find tools by name or description'. Below it, three sections are visible: 'Data access tools' (with a 'Connection' sub-section), 'Automated builders' (with a 'AutoAI' sub-section highlighted by a red arrow), and 'Graphical builders'. The status bar at the bottom shows the time as 8:18:04 AM.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

15. In AutoAI, it required ML instance, so associate the previously created ML instance



The screenshot shows the 'Create an AutoAI experiment' page in the IBM Cloud Pak for Data interface. The URL in the address bar is <https://eu-gb.dataplatform.cloud.ibm.com/projects/8c46babc-000e-4d7e-9e40-f6511f48102b/assets?context=cpdaas>. The left sidebar shows 'IBM Cloud Pak for Data' and 'Projects / EADCp16'. The main area has a title 'Create an AutoAI experiment' with a 'New asset' button. On the left, there's a 'Sample' section. The right side is divided into 'Define details' and 'Define configuration'. In 'Define details', the 'Name' field is filled with 'dbts_exp_1'. In 'Define configuration', it says 'Watson Machine Learning Service Instance' and 'No Machine Learning service instances associated with your project.' There's a 'Associate a Machine Learning service instance' link, a 'Reload' button, and an 'Environment definition' section set to 'Large: 8 CPU and 32 GB RAM'. At the bottom are 'Cancel' and 'Create' buttons.

Name - Yash Lakhtariya

Enrollment number - 21162101012

Branch - CBA Batch - 61

EADC Practical 16

EADCp16 — Projects | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge

https://eu-gb.dataplatform.cloud.ibm.com/projects/8c46bab-000e-4d7e-9e40-f651f48102b/assets?context=cpdaas

IBM Cloud Pak for Data

Projects / EADCp16

Create an AutoAI experiment

Associate service

Choose an existing or add a new service to associate with your project.

Default Locations

Find services

Name	Type	Plan	Location	Status	Group
Watson Machine Learning-ysl	Watson Machine Learning	Lite	London	Not associated	Default

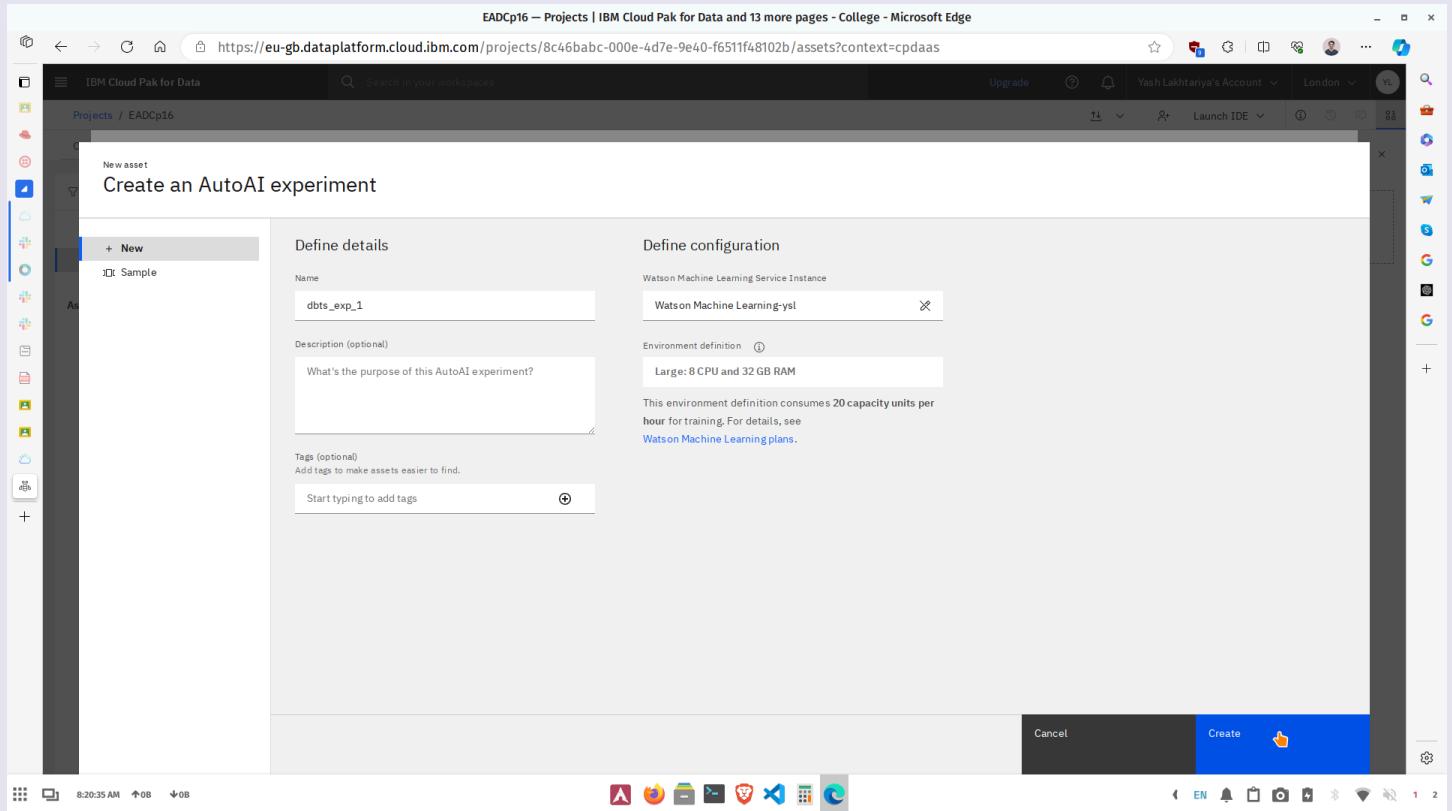
Cancel Associate

8:20:24 AM ↑466B ↓456B

The screenshot shows a Microsoft Edge browser window displaying the IBM Cloud Pak for Data interface. The user is in the 'EADCp16' project under the 'Assets' section, specifically creating a new AutoAI experiment. A modal dialog titled 'Associate service' is open, prompting the user to choose an existing or add a new service. A single service entry is listed: 'Watson Machine Learning-ysl' (Type: Watson Machine Learning, Plan: Lite, Location: London, Status: Not associated, Group: Default). The 'Associate' button at the bottom right of the dialog is highlighted with a yellow arrow. The browser's address bar shows the URL: https://eu-gb.dataplatform.cloud.ibm.com/projects/8c46bab-000e-4d7e-9e40-f651f48102b/assets?context=cpdaas. The system tray at the bottom indicates the time as 8:20:24 AM and network speeds of ↑466B and ↓456B.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

16. Create the AutoAI asset



The screenshot shows the 'Create an AutoAI experiment' dialog box within the IBM Cloud Pak for Data interface. The dialog is divided into two main sections: 'Define details' on the left and 'Define configuration' on the right.

Define details:

- Name:** dbts_exp_1
- Description (optional):** What's the purpose of this AutoAI experiment?
- Tags (optional):** Add tags to make assets easier to find. A placeholder text 'Start typing to add tags' is visible.

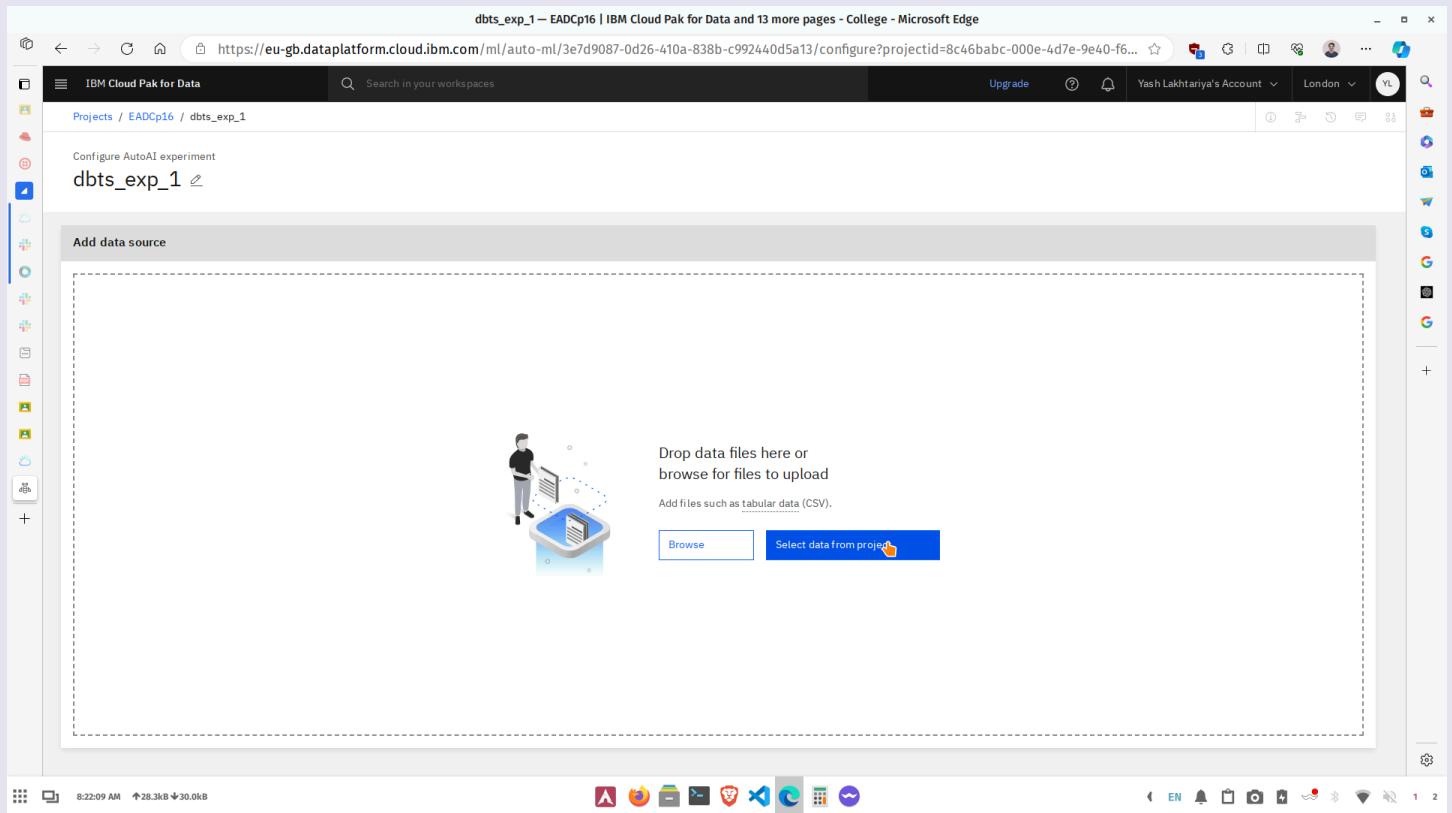
Define configuration:

- Watson Machine Learning Service Instance:** Watson Machine Learning-ysl
- Environment definition:** Large: 8 CPU and 32 GB RAM
- A note states: 'This environment definition consumes 20 capacity units per hour for training. For details, see Watson Machine Learning plans.'

At the bottom right of the dialog are 'Cancel' and 'Create' buttons. The 'Create' button is highlighted with a blue background and a yellow hand cursor icon.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

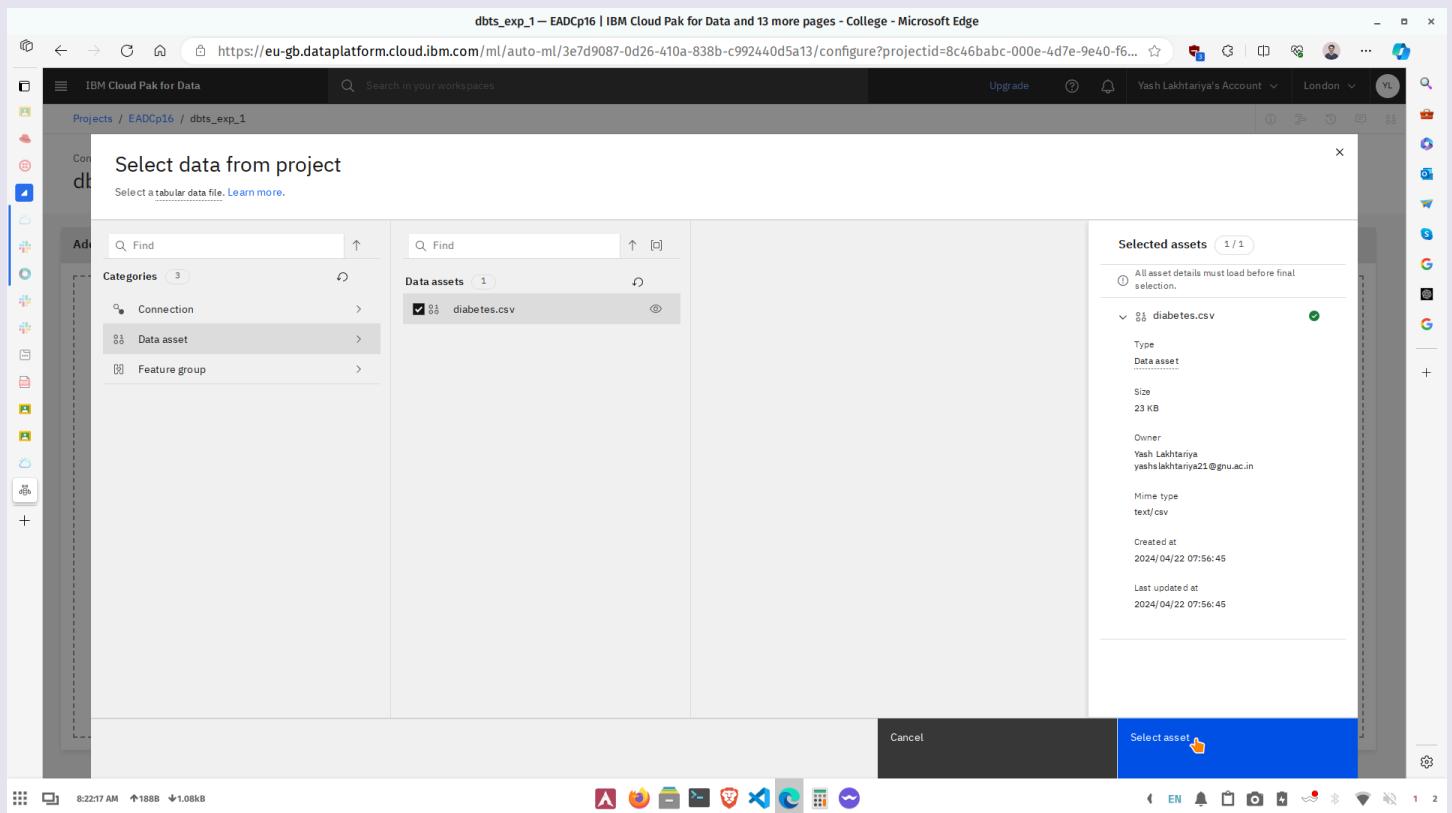
17. In it, add csv file uploaded to the project



The screenshot shows the IBM Cloud Pak for Data interface in Microsoft Edge. The URL in the address bar is <https://eu-gb.dataplatform.cloud.ibm.com/ml/auto-ml/3e7d9087-0d26-410a-838b-c992440d5a13/configure?projectId=8c46bab-000e-4d7e-9e40-f6...>. The page title is "ddts_exp_1 — EADCp16 | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge". The left sidebar shows "IBM Cloud Pak for Data" and "Projects / EADCp16 / dbts_exp_1". The main content area is titled "Configure AutoAI experiment dbts_exp_1" and contains a "Add data source" section. This section features a large dashed box for dropping files, a placeholder icon of a person interacting with a server, and instructions: "Drop data files here or browse for files to upload" and "Add files such as tabular data (CSV)". Below these are two buttons: "Browse" and "Select data from project" (the latter is highlighted with a mouse cursor). The bottom of the screen shows the Windows taskbar with various pinned icons and system status.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

18. Choose diabetes.csv file from data assets



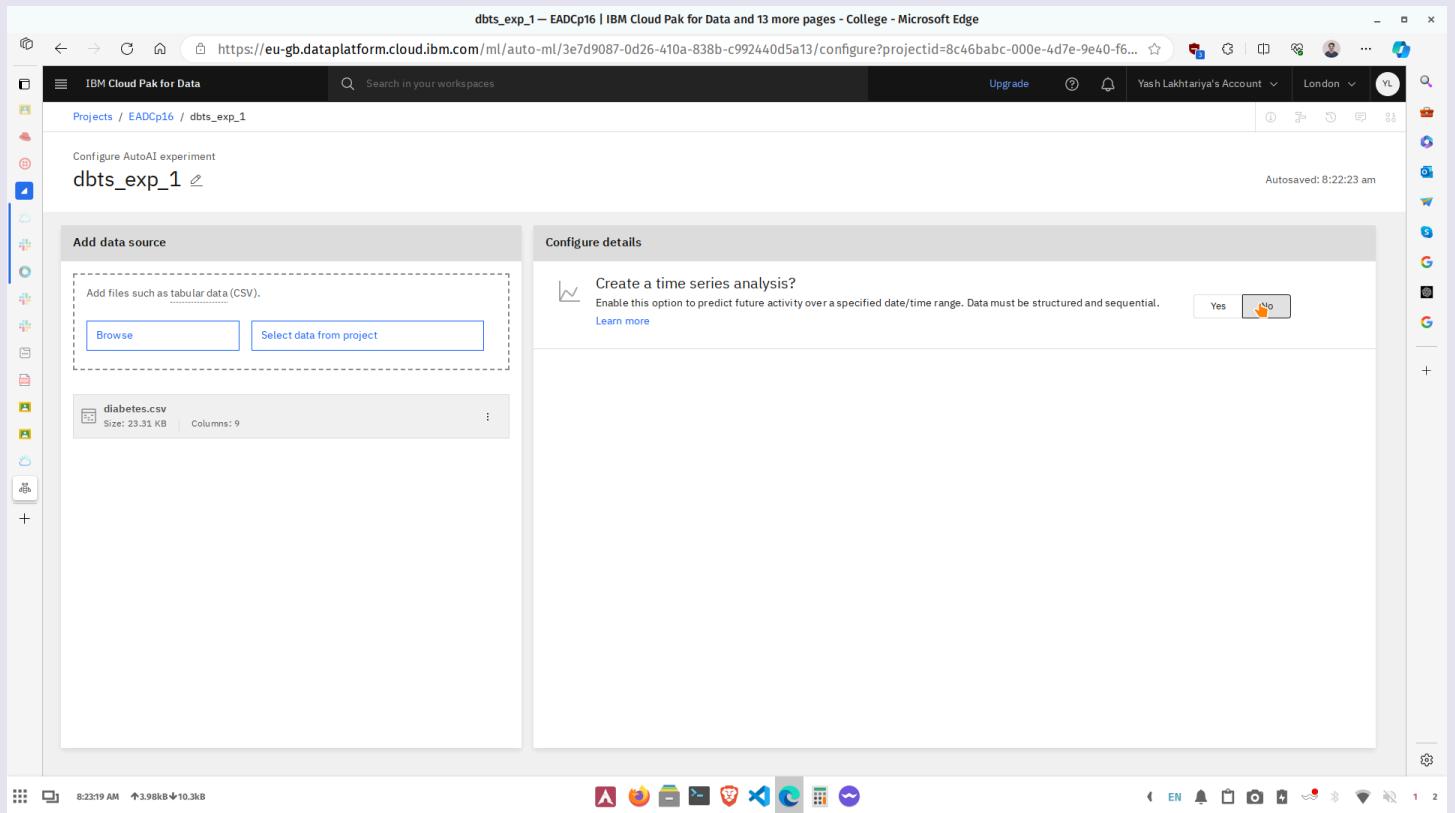
The screenshot shows the 'Select data from project' dialog box in the IBM Cloud Pak for Data interface. On the left, there are two search bars: 'Categories' (with 'Connection', 'Data asset', and 'Feature group' listed) and 'Data assets' (with 'diabetes.csv' selected). On the right, the details for the selected asset are displayed:

- Selected assets**: 1 / 1
- All asset details must load before final selection.
- diabetes.csv** (selected)
- Type**: Data asset
- Size**: 23 KB
- Owner**: Yash Lakhtariya (yash.lakhtariya21@gnu.ac.in)
- Mime type**: text/csv
- Created at**: 2024/04/22 07:56:45
- Last updated at**: 2024/04/22 07:56:45

At the bottom right of the dialog, there is a blue button labeled 'Select asset' with a small orange arrow pointing to it.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

19. Now, click No to time series analysis as it is not required in this case



The screenshot shows the 'Configure AutoAI experiment' page for 'dbts_exp_1'. On the left, under 'Add data source', there's a section for 'Add files such as tabular data (CSV)' with 'Browse' and 'Select data from project' buttons. A file named 'diabetes.csv' is listed with a size of 23.31 KB and 9 columns. On the right, under 'Configure details', there's a section titled 'Create a time series analysis?' with a description: 'Enable this option to predict future activity over a specified date/time range. Data must be structured and sequential.' Below this is a button group with 'Yes' and 'No'. The 'No' button is highlighted with a yellow arrow pointing to it. The top of the screen shows the URL 'https://eu-gb.dataplatform.cloud.ibm.com/ml/auto-ml/3e7d9087-0d26-410a-838b-c992440d5a13/configure?projectid=8c46bab-000e-4d7e-9e40-f6...', the title 'dbts_exp_1 — EADCp16 | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge', and the user 'Yash Lakhtariya's Account'. The bottom of the screen shows the Windows taskbar with various icons.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

20. Now, predict Outcome column and visit Experiment Settings to explore if required

The screenshot shows the IBM Cloud Pak for Data interface in Microsoft Edge. The URL is <https://eu-gb.dataplatform.cloud.ibm.com/ml/auto-ml/3e7d9087-0d26-410a-838b-c992440d5a13/configure?projectid=8c46bab-000e-4d7e-9e40-f6...>. The page title is "dbts_exp_1 — EADCp16 | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge".

The main content area displays the "Configure AutoAI experiment" section for "dbts_exp_1". On the left, there's a sidebar with icons for various workspace components like files, projects, and dashboards. The main panel has two main sections:

- Add data source:** A dashed box for adding files such as tabular data (CSV). It includes "Browse" and "Select data from project" buttons. Below this is a preview of a file named "diabetes.csv" (Size: 23.31 KB, Columns: 9).
- Configure details:** This section includes:
 - A checkbox for "Create a time series analysis?" with a "Learn more" link. The "Yes" button is selected.
 - A section titled "What do you want to predict?" where "Outcome" is listed as the prediction column.
 - A "Prediction column: Outcome" section showing "PREDICTION TYPE: Binary Classification", "POSITIVE CLASS: 1", and "CUH remaining: 20 CUH".
 - An "Experiment settings" button at the bottom left and a "Run experiment" button at the bottom right.

The status bar at the bottom shows the date and time as 8:26:17 AM and network activity as 5.26kB ↔ 617kB.

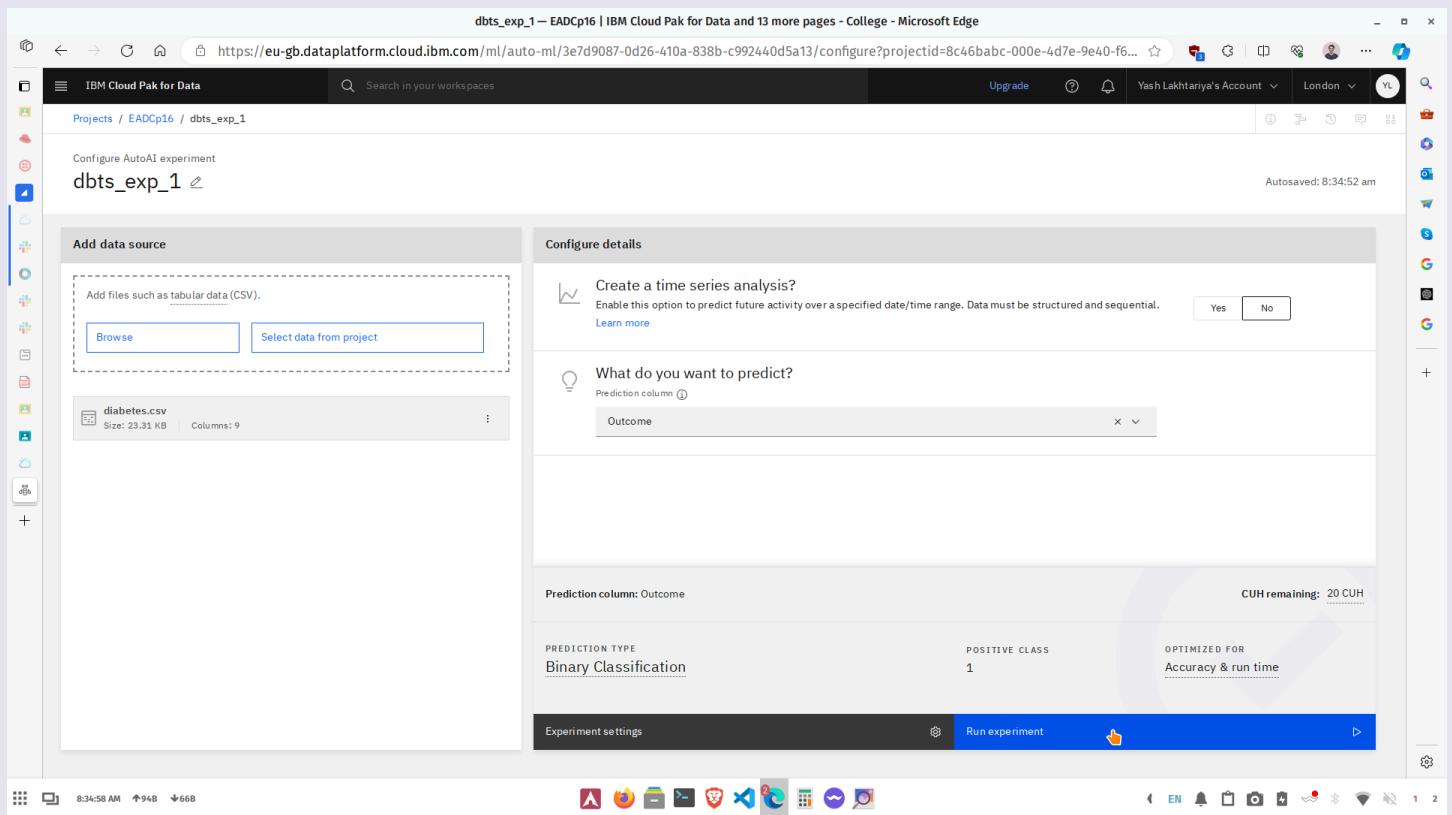
Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

21. Here, the setting useful in this case is the number of algorithms to use

The screenshot shows the 'Experiment settings' dialog box in Microsoft Edge. The URL is <https://eu-gb.dataplatform.cloud.ibm.com/ml/auto-ml/3e7d9087-0d26-410a-838b-c992440d5a13/configure?projectid=8c46bab-000e-4d7e-9e40-f6...>. The 'Prediction' tab is selected. Under 'General', several algorithms are listed with checkboxes: LGBM Classifier, Logistic Regression, Random Forest Classifier, Snap Boosting Machine Classifier, Snap Decision Tree Classifier, Snap Logistic Regression, Snap Random Forest Classifier, Snap SVM Classifier, and XGB Classifier. All checkboxes are checked. Below this, a progress bar indicates 'Algorithms to use 2 / 4'. A note states: 'AutoAI will test the specified algorithms and use the top performers to create model pipelines. Choose how many top algorithms to apply. Each algorithm generates 4-5 pipelines and more algorithms increase the runtime.' At the bottom right are 'Cancel' and 'Save settings' buttons.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

22. Now, Run experiment



The screenshot shows the IBM Cloud Pak for Data interface for configuring an AutoAI experiment. The experiment name is "dbts_exp_1".

Add data source: A file named "diabetes.csv" (Size: 23.31 KB, Columns: 9) is selected.

Configure details:

- Create a time series analysis?**: A checkbox labeled "Yes" is checked.
- What do you want to predict?**: The prediction column is set to "Outcome".
- PREDICTION TYPE**: Binary Classification
- POSITIVE CLASS**: 1
- CUH remaining:** 20 CUH
- OPTIMIZED FOR**: Accuracy & run time

Experiment settings and **Run experiment** buttons are at the bottom. The "Run experiment" button has a blue background and a white hand cursor icon.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

23. Wait for pipeline to complete and click any Pipeline to save (here Pipeline 4 with highest accuracy)

dbts_exp_1 — EADCp16 | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge

IBM Cloud Pak for Data

Projects / EADCp16 / dbts_exp_1

Experiment summary Pipeline comparison

Relationship map ⚑ Prediction column: Outcome

FEATURE TRANSFORMERS

PIPELINES

1 | Pipeline 4 | THMS

Extra Trees Classifier

diabetes.cs...

Accuracy: 0.771

90% Training data 3 Folds 10% Holdout data

Progress map Swap view

Experiment completed 8 PIPELINES GENERATED

8 pipelines generated from algorithms. See pipeline leaderboard below for more detail.

Time elapsed: 2 minutes

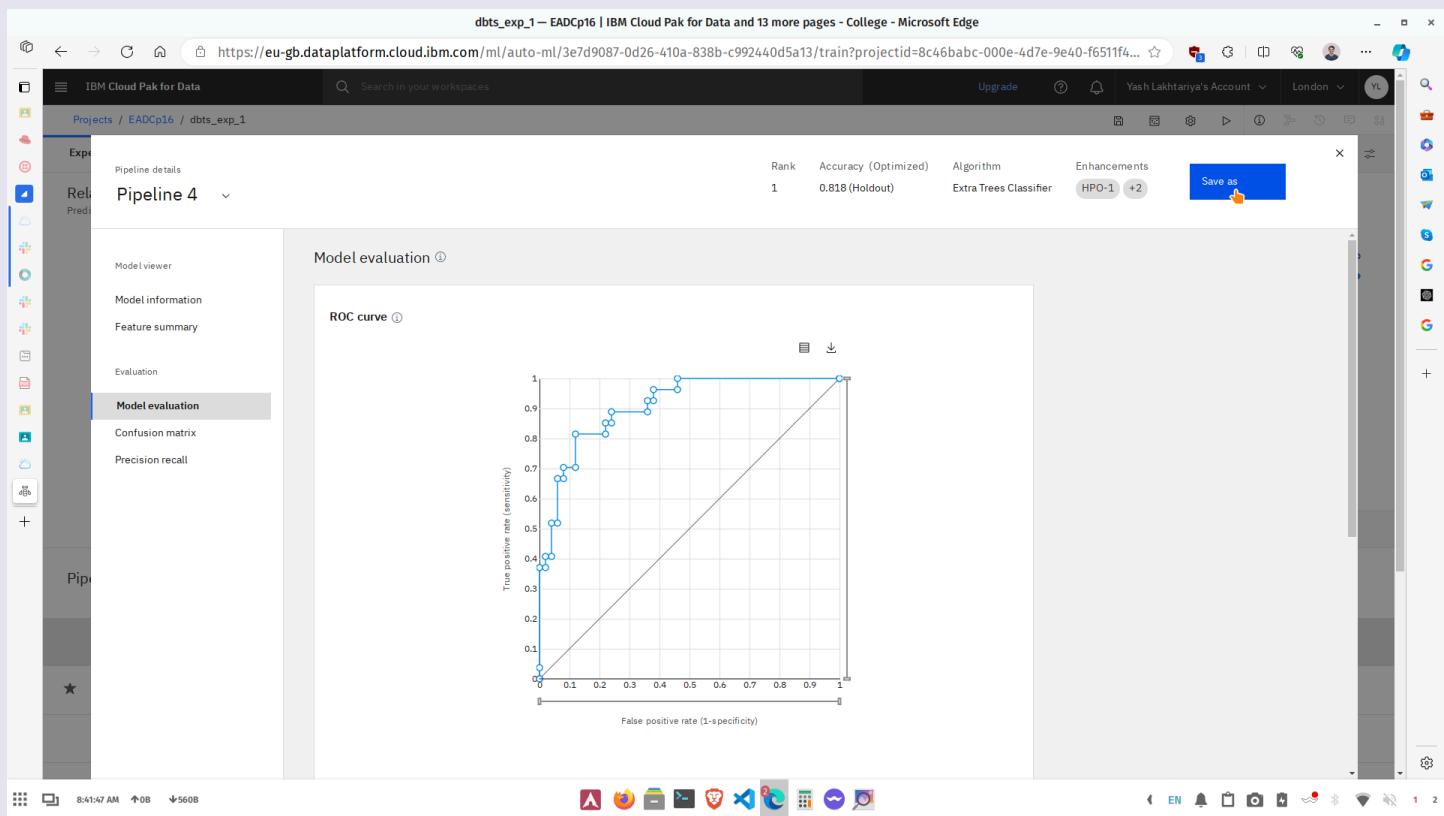
View log Save code

Pipeline leaderboard

Rank	↑	Name	Algorithm	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 4	Extra Trees Classifier	0.771	HPO-1 FE HPO-2	00:00:54
	2	Pipeline 3	Extra Trees Classifier	0.767	HPO-1 FE	00:00:40

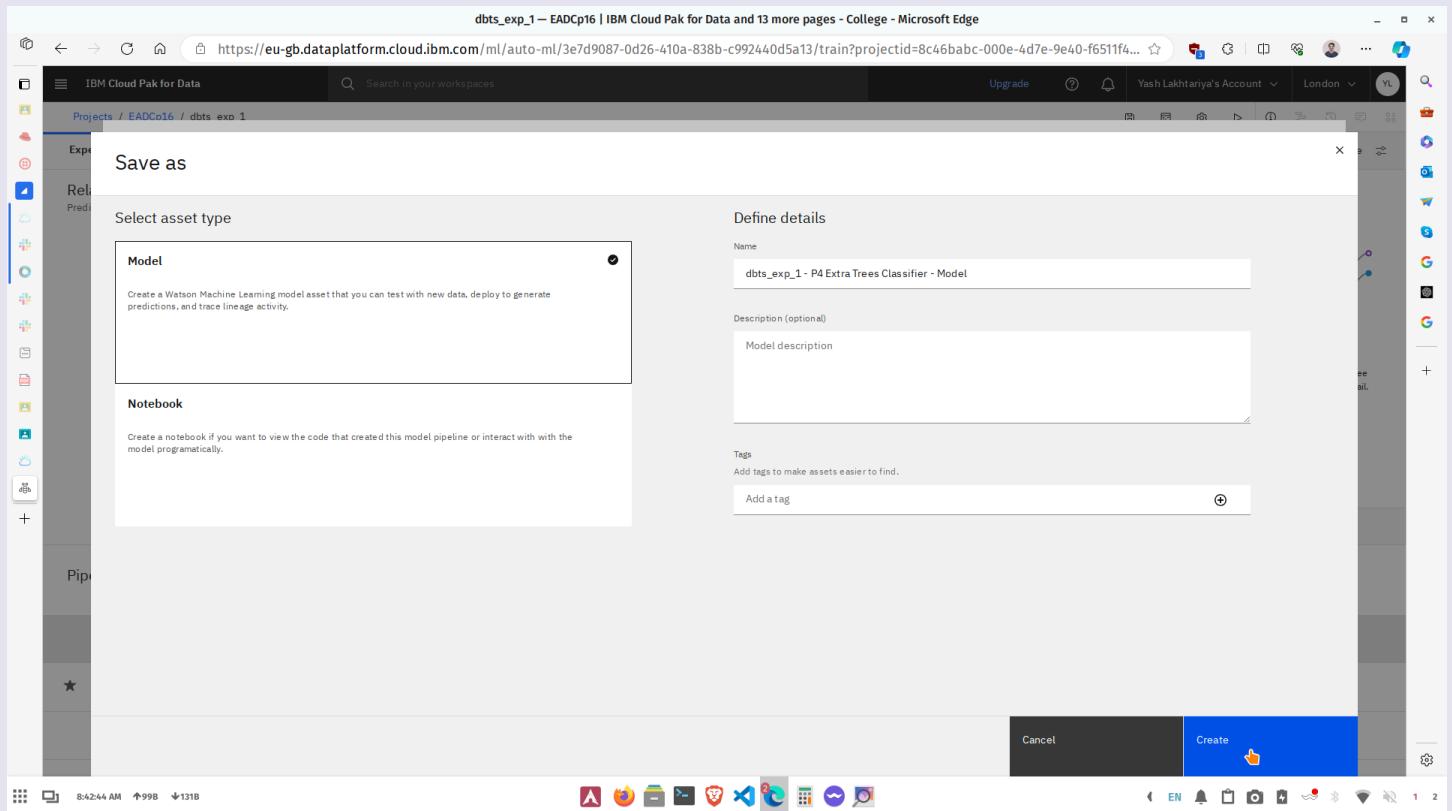
8:41:36 AM 446B 3.12kB

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

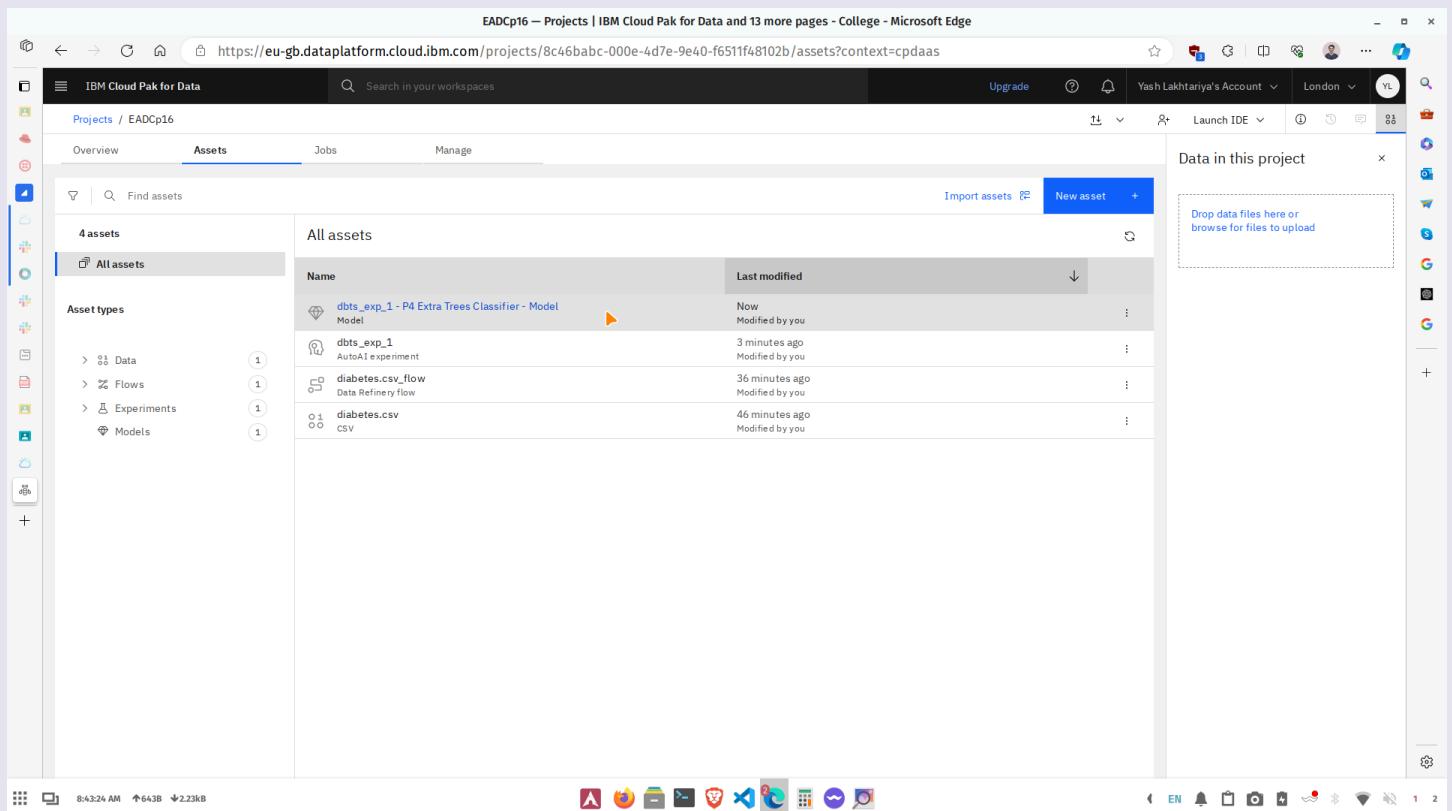
24. Save as Model



The screenshot shows the IBM Cloud Pak for Data interface in Microsoft Edge. A modal dialog titled "Save as" is open, prompting the user to select an asset type. The "Model" option is selected, with a description: "Create a Watson Machine Learning model asset that you can test with new data, deploy to generate predictions, and trace lineage activity." Below this is the "Notebook" option, which is not selected. The right side of the dialog is titled "Define details" and contains fields for "Name" (set to "dbts_exp_1 - P4 Extra Trees Classifier - Model"), "Description (optional)" (set to "Model description"), and "Tags" (a field to add tags). At the bottom right of the dialog are "Cancel" and "Create" buttons, with the "Create" button being highlighted with a blue background and a hand cursor icon.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

25. It is visible in Assets tab



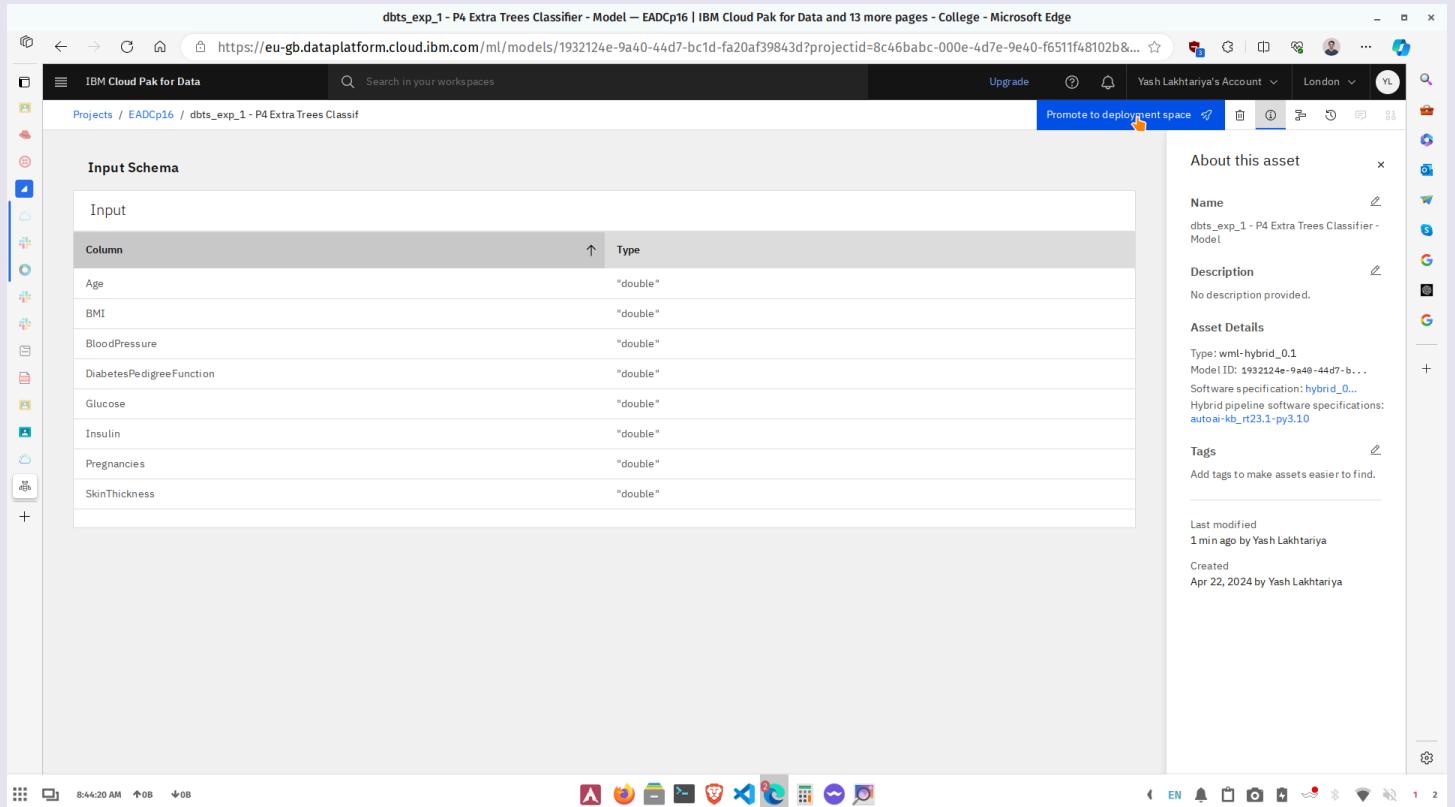
The screenshot shows the 'Assets' tab in the IBM Cloud Pak for Data interface. On the left, there's a sidebar with 'Asset types' listed: Data (1), Flows (1), Experiments (1), and Models (1). The main area displays a table titled 'All assets' with columns for 'Name' and 'Last modified'. The assets listed are:

Name	Last modified
dbts_exp_1 - P4 Extra Trees Classifier - Model	Now Modified by you
dbts_exp_1	3 minutes ago Modified by you
diabetes.csv.flow	36 minutes ago Modified by you
diabetes.csv	46 minutes ago Modified by you

To the right of the table is a 'Data in this project' section with a placeholder 'Drop data files here or browse for files to upload'. The browser address bar shows the URL: <https://eu-gb.dataplatform.cloud.ibm.com/projects/8c46babc-000e-4d7e-9e40-f6511f48102b/assets?context=cpdaas>.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

26. Click it and promote to deployment space



The screenshot shows the IBM Cloud Pak for Data interface. On the left, there's a sidebar with various icons. The main area has a header bar with the URL <https://eu-gb.dataplatform.cloud.ibm.com/ml/models/1932124e-9a40-44d7-bc1d-fa20af39843d?projectId=8c46babbc-000e-4d7e-9e40-f651f48102b&...>. Below the header, there's a search bar and a 'Promote to deployment space' button. The central part of the screen displays the 'Input Schema' for a model named 'dbts_exp_1 - P4 Extra Trees Classifier - Model'. The schema table lists columns: Age, BMI, BloodPressure, DiabetesPedigreeFunction, Glucose, Insulin, Pregnancies, and SkinThickness, each with a type of "double". To the right, there's a panel titled 'About this asset' with sections for Name, Description, Asset Details, Tags, Last modified, and Created. The 'Name' section shows 'dbts_exp_1 - P4 Extra Trees Classifier - Model'. The 'Description' section says 'No description provided.'. The 'Asset Details' section includes 'Type: wml-hybrid_0.1', 'Model ID: 1932124e-9a40-44d7-bc1d-fa20af39843d', 'Software specification: hybrid_0..', and 'Hybrid pipeline software specifications: autoai-kb_r123.1-py3.10'. The 'Tags' section is empty. The 'Last modified' section shows '1 min ago by Yash Lakhtariya'. The 'Created' section shows 'Apr 22, 2024 by Yash Lakhtariya'. At the bottom of the interface, there are system status icons and a taskbar.

Column	Type
Age	"double"
BMI	"double"
BloodPressure	"double"
DiabetesPedigreeFunction	"double"
Glucose	"double"
Insulin	"double"
Pregnancies	"double"
SkinThickness	"double"

About this asset

Name
dbts_exp_1 - P4 Extra Trees Classifier - Model

Description
No description provided.

Asset Details
Type: wml-hybrid_0.1
Model ID: 1932124e-9a40-44d7-bc1d-fa20af39843d
Software specification: hybrid_0..
Hybrid pipeline software specifications: [autoai-kb_r123.1-py3.10](#)

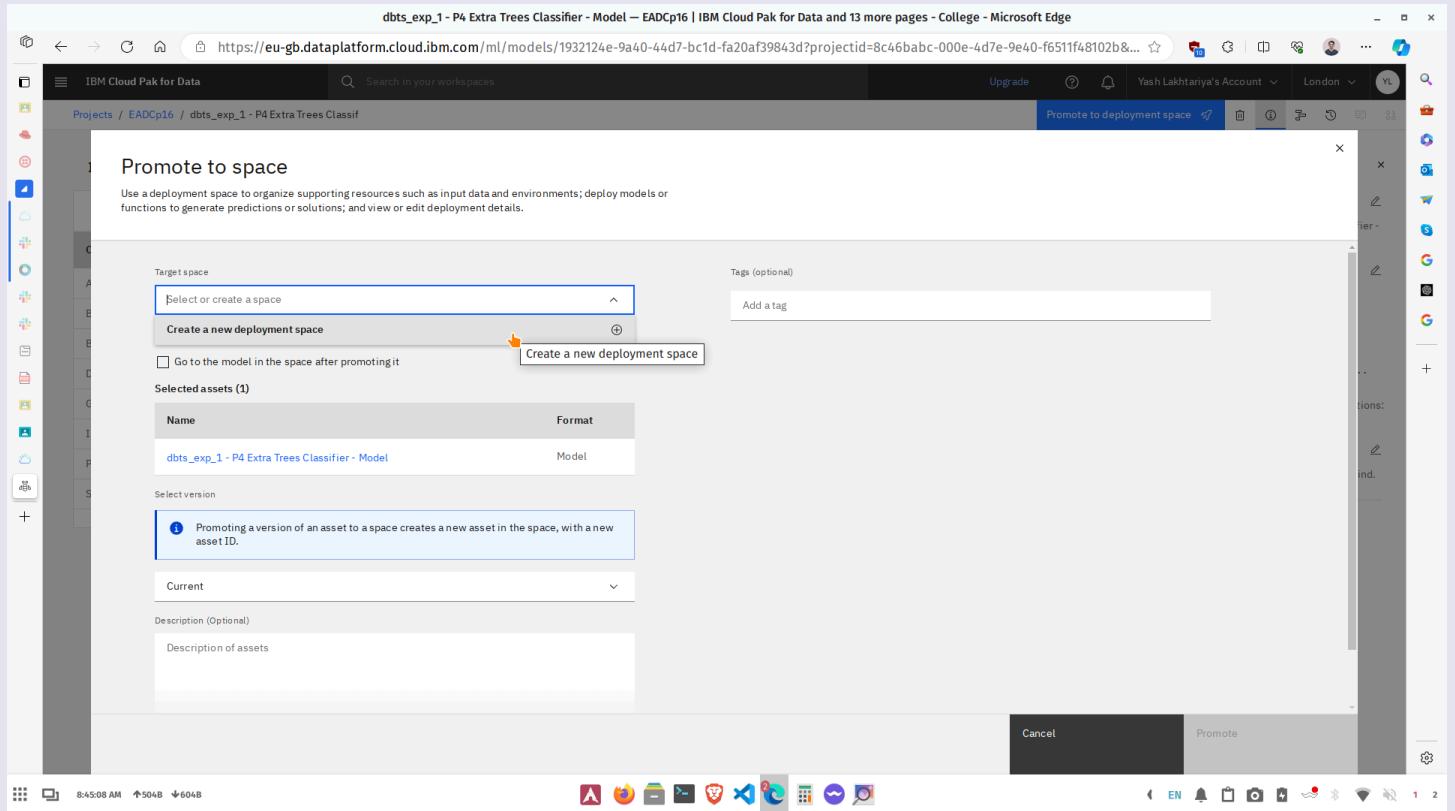
Tags
Add tags to make assets easier to find.

Last modified
1 min ago by Yash Lakhtariya

Created
Apr 22, 2024 by Yash Lakhtariya

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

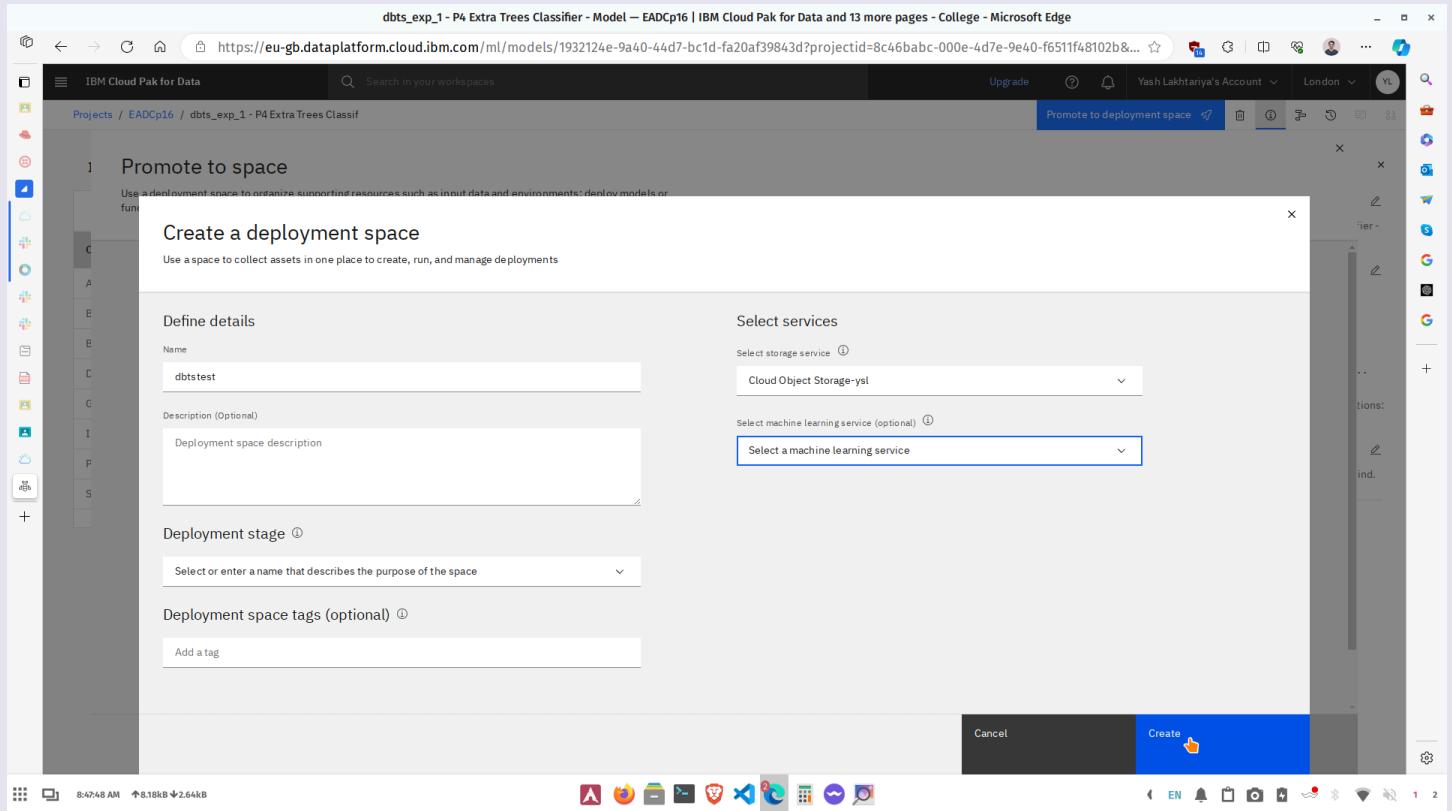
27. Create a new deployment space for that



The screenshot shows the 'Promote to space' dialog box in Microsoft Edge. The URL in the address bar is <https://eu-gb.dataplatform.cloud.ibm.com/ml/models/1932124e-9a40-44d7-bc1d-fa20af39843d?projectId=8c46babbc-000e-4d7e-9e40-f651f48102b&...>. The dialog has a title 'Promote to space' and a sub-instruction: 'Use a deployment space to organize supporting resources such as input data and environments; deploy models or functions to generate predictions or solutions, and view or edit deployment details.' It contains fields for 'Target space' (with a dropdown menu showing 'Select or create a space' and a button 'Create a new deployment space'), 'Tags (optional)' (with a 'Add a tag' input), and 'Selected assets (1)'. A table lists one asset: 'dbts_exp_1 - P4 Extra Trees Classifier - Model' (Format: Model). Below this is a note: 'Promoting a version of an asset to a space creates a new asset in the space, with a new asset ID.' Under 'Select version', 'Current' is selected. There is also a 'Description (Optional)' field containing 'Description of assets'. At the bottom are 'Cancel' and 'Promote' buttons.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

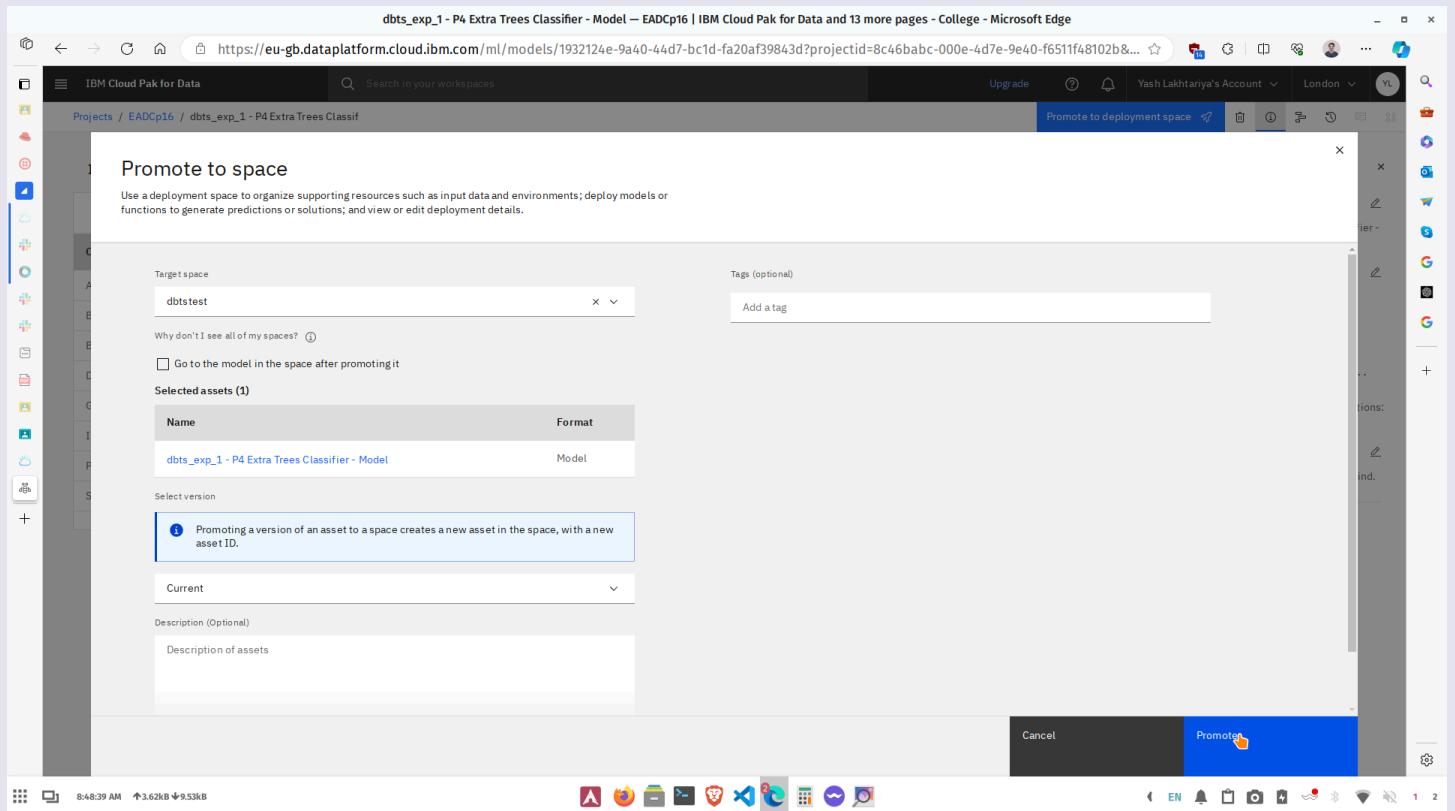
28. Create the deployment space finally



The screenshot shows the 'Promote to space' dialog box in the IBM Cloud Pak for Data interface. The dialog is titled 'Create a deployment space'. It contains two main sections: 'Define details' and 'Select services'. In the 'Define details' section, the 'Name' field is set to 'dbtstest'. The 'Description (Optional)' field is empty. In the 'Select services' section, the 'Select storage service' dropdown is set to 'Cloud Object Storage-ysl'. The 'Select machine learning service (optional)' dropdown is empty. Below these sections are fields for 'Deployment stage' and 'Deployment space tags (optional)'. The 'Create' button at the bottom right is highlighted with a yellow arrow. The background shows a sidebar with project names A through S and a top navigation bar.

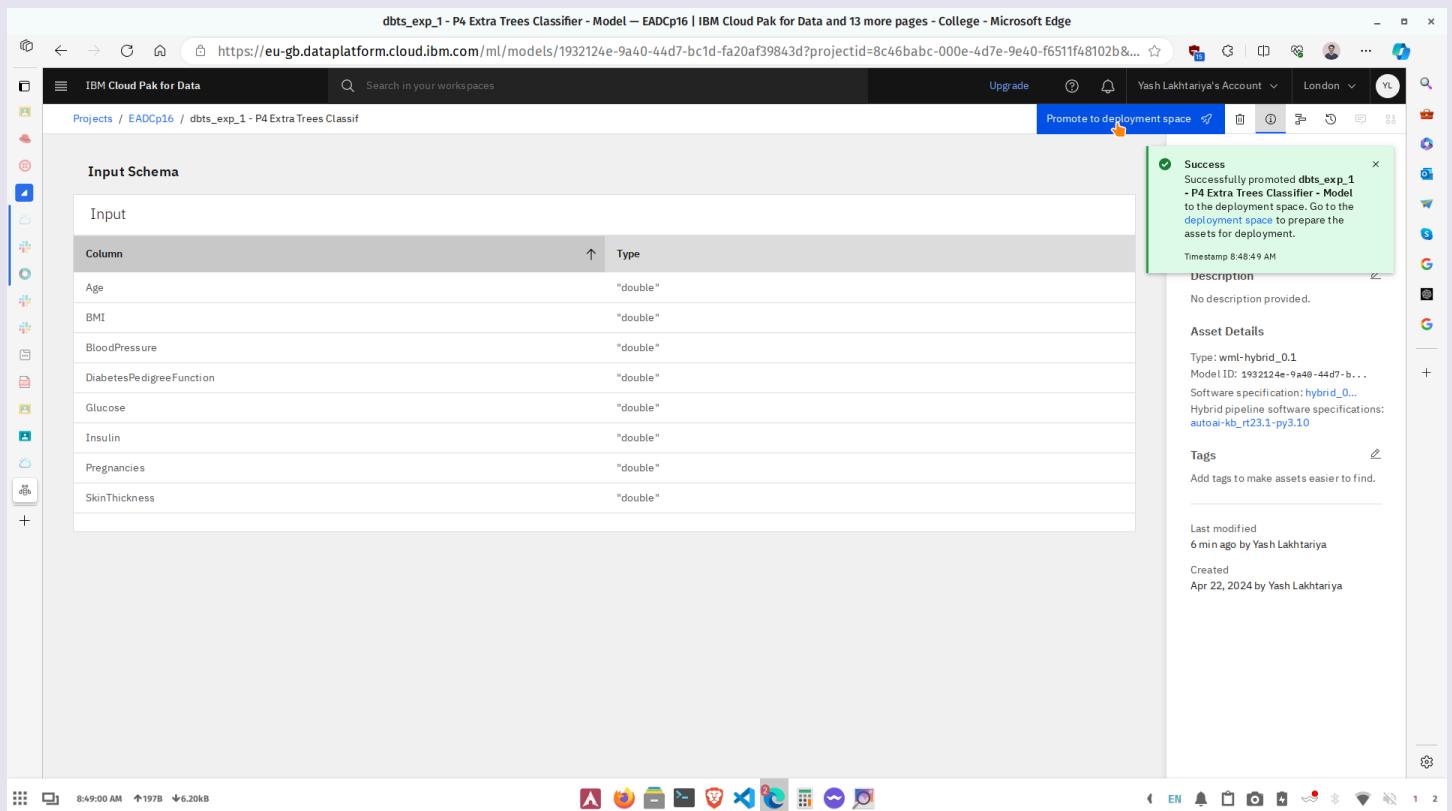
Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

29. After reviewing the changes, Promote to Space



Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

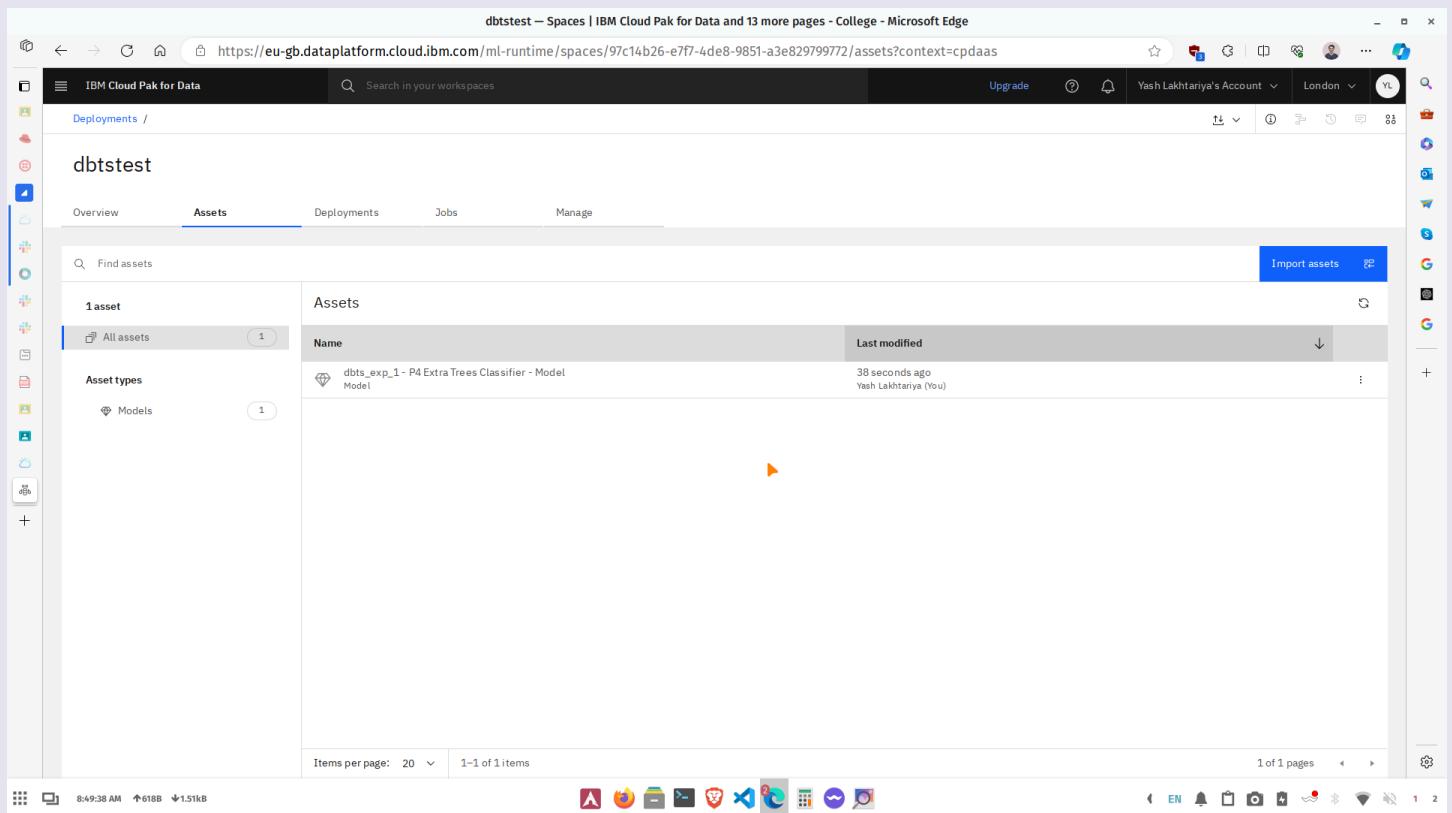
30. Now go to Assets after successful promotion



The screenshot shows the IBM Cloud Pak for Data interface. The URL in the address bar is <https://eu-gb.dataplatform.cloud.ibm.com/ml/models/1932124e-9a40-44d7-bc1d-fa20af39843d?projectId=8c46babbc-000e-4d7e-9e40-f651f48102b&...>. The main content area shows the 'Input Schema' for a 'dbts_exp_1 - P4 Extra Trees Classifier - Model'. The schema includes columns: Age, BMI, BloodPressure, DiabetesPedigreeFunction, Glucose, Insulin, Pregnancies, and SkinThickness, all of type "double". To the right, a sidebar displays 'Asset Details' for the model, including its ID, type (wm1-hybrid_0.1), and creation date (Apr 22, 2024). A green success message box states: 'Success: Successfully promoted dbts_exp_1 - P4 Extra Trees Classifier - Model to the deployment space. Go to the deployment space to prepare the assets for deployment.' The timestamp for this message is 8:48:49 AM.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

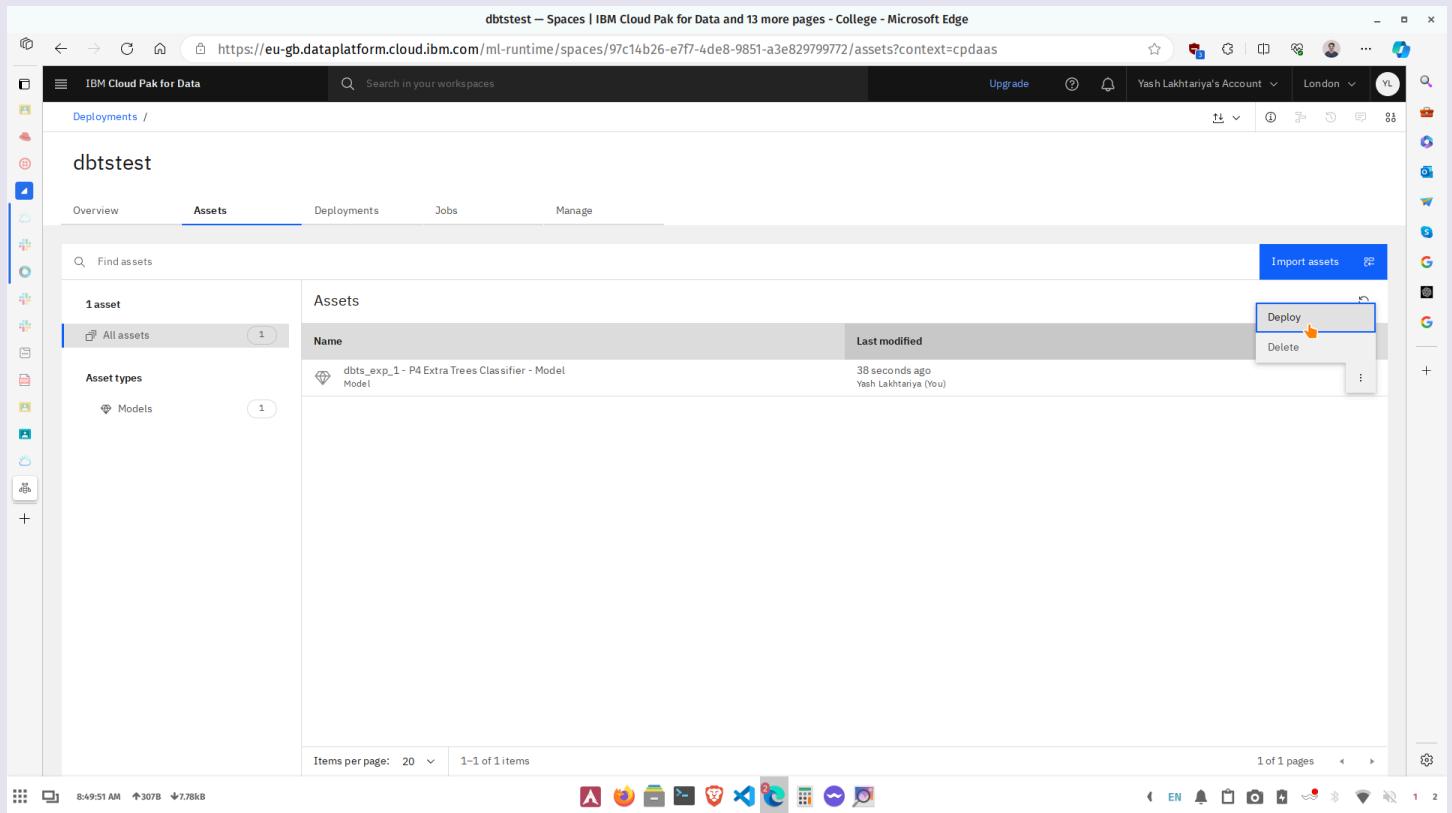
31. Here, the promotion is successful



The screenshot shows the 'Assets' tab in the IBM Cloud Pak for Data interface. A single asset named 'dbts_exp_1 - P4 Extra Trees Classifier - Model' is listed under the 'Assets' section. The asset was last modified 38 seconds ago by 'Yash Lakhtariya (You)'. The interface includes a sidebar with various icons for managing workspaces, deployments, and jobs. The top navigation bar shows the URL 'https://eu-gb.dataplatform.cloud.ibm.com/ml-runtime/spaces/97c14b26-e7f7-4de8-9851-a3e829799772/assets?context=cpdaas' and the account information 'Yash Lakhtariya's Account' and 'London'.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

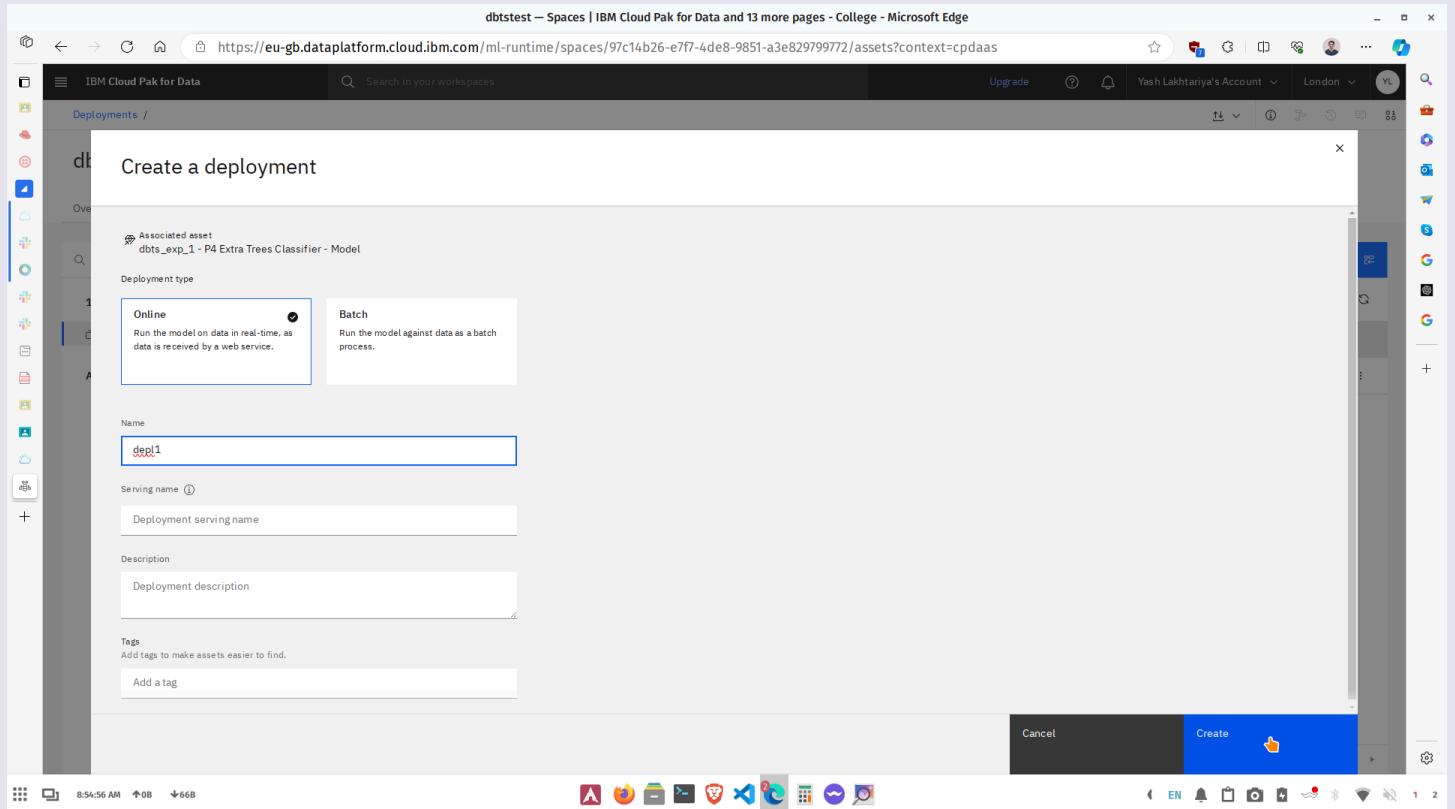
32. Click Deploy in the context menu



The screenshot shows the IBM Cloud Pak for Data interface in Microsoft Edge. The user is on the 'Assets' tab of the 'dbtstest' workspace. A context menu is open over a single asset named 'dbts_exp_1 - P4 Extra Trees Classifier - Model'. The menu options are 'Deploy' (highlighted with a red box and cursor) and 'Delete'. The left sidebar shows '1 asset' under 'All assets' and 'Asset types' with 'Models' selected. The bottom status bar shows network activity: 8:49:51 AM, ↑307B, ↓7.78kB.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

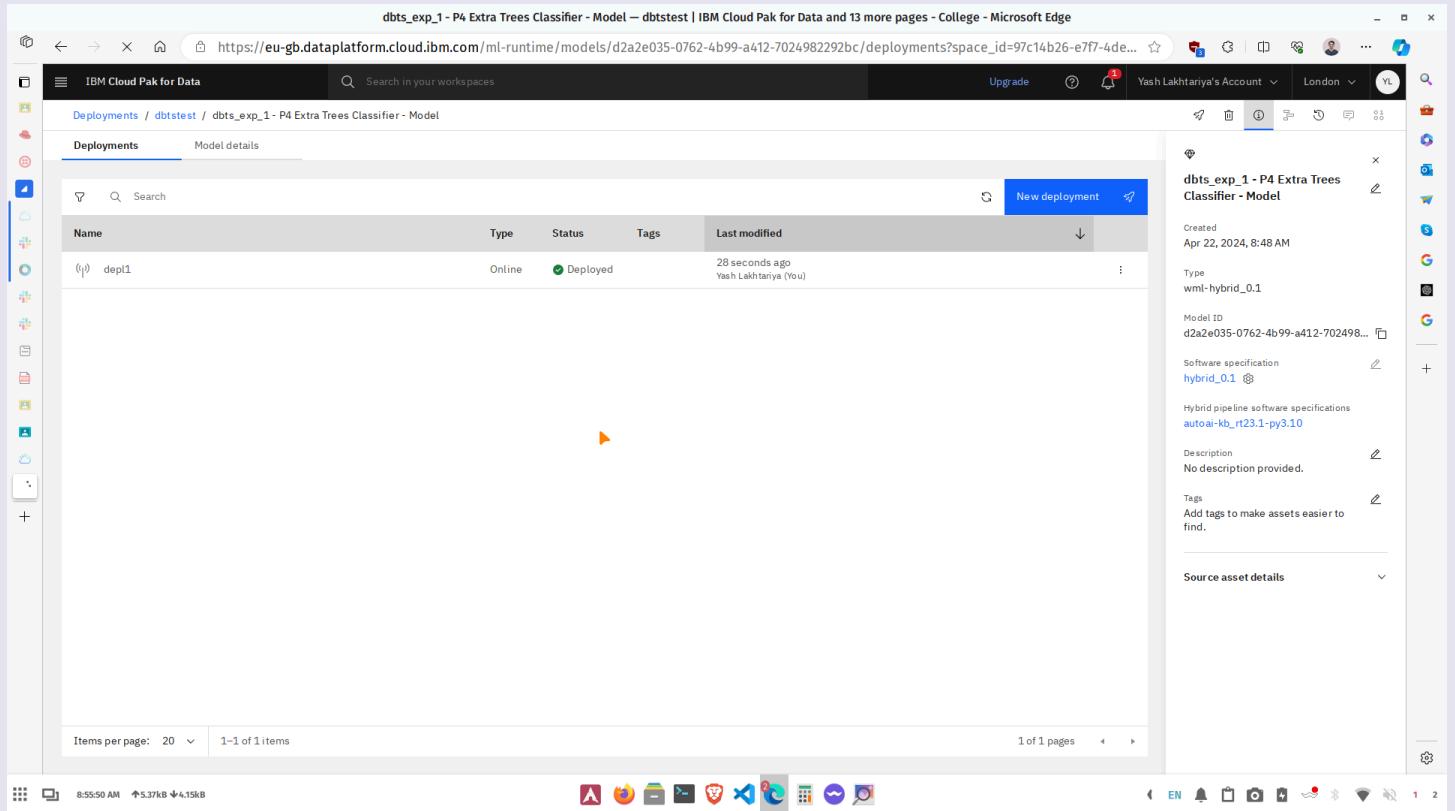
33. Choose Online type and deploy



The screenshot shows a Microsoft Edge browser window displaying the 'Create a deployment' dialog box from the IBM Cloud Pak for Data interface. The URL in the address bar is <https://eu-gb.dataplatform.cloud.ibm.com/ml-runtime/spaces/97c14b26-e7f7-4de8-9851-a3e829799772/assets?context=cpdaas>. The dialog box has a title 'Create a deployment'. It shows an associated asset 'dots_exp_1 - P4 Extra Trees Classifier - Model'. Under 'Deployment type', the 'Online' option is selected, described as 'Run the model on data in real-time, as data is received by a web service.' The 'Batch' option is also listed as 'Run the model against data as a batch process.'. The 'Name' field contains 'depl1'. The 'Serving name' field is empty. The 'Description' field contains 'Deployment description'. The 'Tags' section is empty. At the bottom right, there are 'Cancel' and 'Create' buttons, with the 'Create' button being highlighted.

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

34. Wait for deployment to be finished



The screenshot shows the IBM Cloud Pak for Data interface in Microsoft Edge. The URL is https://eu-gb.dataplatform.cloud.ibm.com/ml-runtime/models/d2a2e035-0762-4b99-a412-7024982292bc/deployments?space_id=97c14b26-e7f7-4de.... The page displays a deployment list and detailed information for a specific deployment.

Deployments / dbttest / dbts_exp_1 - P4 Extra Trees Classifier - Model

Name	Type	Status	Tags	Last modified
dep1	Online	Deployed		28 seconds ago Yash Lakhtariya (You)

Model details for dbts_exp_1 - P4 Extra Trees Classifier - Model:

- Created: Apr 22, 2024, 8:48 AM
- Type: wml-hybrid_0.1
- Model ID: d2a2e035-0762-4b99-a412-7024982292bc
- Software specification: hybrid_0.1
- Hybrid pipeline software specifications: autoai-kb_rt23.1-py3.10
- Description: No description provided.
- Tags: Add tags to make assets easier to find.

Source asset details

Items per page: 20 | 1-1 of 1 items | 1 of 1 pages

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

35. Use any set of values (record) from dataset to test the outcome is accurate or not

depl1 — dbttest | IBM Cloud Pak for Data and 13 more pages - College - Microsoft Edge

IBM Cloud Pak for Data

Deployments / dbttest / dbts_exp_1 - P4 Extra Trees Clas... /

depl1 Deployed Online

API reference Test

Enter input data

Text JSON

Enter data manually or use a CSV file to populate the spreadsheet. Max file size is 50 MB.

Download CSV template Browse local files Search in space Clear all X

	Pregnancies (double)	Glucose (double)	BloodPressure (double)	SkinThickness (double)	Insulin (double)	BMI (double)	DiabetesPedigreeFunction (double)	Age (double)
1	1	103	30	38	83	43.3	0.183	33
2								
3								
4								
5								
6								
7								
8								
9								
10								

1 row, 8 columns

Predict

8:58:27 AM ↑8.88kB ↓3.26kB

Name - Yash Lakhtariya

Enrollment number - 21162101012

Branch - CBA Batch - 61

EADC Practical 16

diabetes.csv — LibreOffice Calc

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	I														
2	6	148	72	35	0	33.6	0.627	50	1														
3	1	85	66	29	0	26.6	0.351	31	0														
4	8	183	64	0	0	23.3	0.672	32	1														
5	1	89	66	23	94	28.1	0.167	21	0														
6	0	137	40	35	168	43.1	2.288	33	1														
7	5	116	74	0	0	25.6	0.201	30	0														
8	3	78	50	32	86	31	0.248	26	1														
9	10	15	0	0	0	35.5	0.134	29	0														
10	2	167	70	46	543	30.5	0.158	35	1														
11	8	125	96	0	0	0	0.232	54	1														
12	4	110	92	0	0	37.6	0.191	30	0														
13	10	168	74	0	0	38	0.537	34	1														
14	10	139	80	0	0	27.1	1.441	57	0														
15	1	189	60	23	846	30.1	0.398	59	1														
16	5	166	72	19	175	25.8	0.587	51	1														
17	7	100	0	0	0	30	0.484	32	1														
18	0	118	84	47	230	45.8	0.551	31	1														
19	7	107	74	0	0	29.6	0.254	31	1														
20	1	145	82	38	93	43.8	0.145	34	1														
21	15	70	0	36	106	30.6	0.529	25	1														
22	3	126	88	41	225	39.3	0.704	27	0														
23	8	99	84	0	0	35.4	0.388	50	0														
24	7	196	90	0	0	39.8	0.451	41	1														
25	9	119	80	35	0	29	0.263	29	1														
26	11	143	94	33	146	36.6	0.254	51	1														
27	10	125	70	26	115	31.1	0.205	41	1														
28	7	147	76	0	0	39.4	0.257	43	1														
29	1	97	66	15	140	23.2	0.487	22	0														
30	13	145	82	19	118	22.2	0.245	57	0														
31	5	117	92	0	0	34.6	0.331	36	0														
32	5	109	75	26	0	36	0.546	60	0														
33	3	158	76	36	245	31.6	0.851	28	1														
34	3	88	58	11	54	24.8	0.267	22	0														
35	6	92	92	0	0	19.9	0.188	28	0														
36	10	122	78	31	0	27.6	0.512	45	0														
37	4	103	60	33	192	24	0.966	33	0														
38	11	138	76	0	0	33.2	0.42	35	0														
39	9	102	76	37	0	32.9	0.665	46	1														
40	2	68	42	0	38.2	0.033	27	1															
41	4	111	72	47	207	37.1	1.39	56	1														
42	3	180	64	25	70	34	0.271	26	0														
43	7	133	84	0	0	40.2	0.696	37	0														

diabetes

Sheet 1 of 1

Average: 0; Sum: 0

EN 8:58:41 AM 8:58:41 AM 3.98kB 9.01kB 100%

Name - Yash Lakhtariya
Enrollment number - 21162101012
Branch - CBA Batch - 61
EADC Practical 16

36. As seen, the prediction is 100% accurate in this case

The screenshot shows the 'Prediction results' page in IBM Cloud Pak for Data. The prediction type is 'Binary classification'. A large purple circle indicates '1 Record'. Below it, a bar chart shows the 'Confidence level distribution' with one bar at 60-70% confidence. The main table lists 16 predictions, all with a confidence of 62%.

Prediction	Confidence
1	62%
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	