NewgenONE

Intelligent Document Processing Studio

User Guide

Version: 1.0

Disclaimer

This document contains information proprietary to Newgen Software Technologies Ltd. User may not disclose or use any proprietary information or use any part of this document without written permission from Newgen Software Technologies Ltd.

Newgen Software Technologies Ltd. makes no representations or warranties regarding any software or to the contents or use of this guide. It also specifically disclaims any express or implied warranties of merchantability, title, or fitness for any particular purpose. Even though Newgen Software Technologies Ltd. has tested the hardware and software and reviewed the documentation, it does not guarantee or imply that this document is error free or accurate regarding any particular specification. As a result, this product is sold as it is and user, the purchaser, is assuming the entire risk as to its quality and performance. Further, Newgen Software Technologies Ltd. reserves the right to revise this publication and make changes in its content without any obligation to notify any person, of such revisions or changes. Newgen Software Technologies Ltd. authorizes no Newgen agent, dealer or employee to make any modification, extension, or addition to the above statements.

Newgen Software Technologies Ltd. has attempted to supply trademark information about company names, products, and services mentioned in this document. Trademarks indicated below were derived from various sources.

Newgen Software, Registered Office, New Delhi
E-44/13
Okhla Phase - II
New Delhi 110020
India
Phone: +91 1146 533 200
info@newgensoft.com

Contents

Preface

This section provides information on the revision history, about this guide, and intended audience for the Intelligent Document Processing (IDP) Studio user guide.

Revision history

Revision date	Description
May 2024	Initial publication

About this guide

This guide explains how to create, publish, deploy and maintain data models using the IDP Studio platform.

Intended audience

This guide is intended for data scientists and data analysts. The reader must have an understanding of data manipulation techniques.

Documentation feedback

To provide feedback or any improvement suggestions on technical documentation, write an email to docs.feedback@newgensoft.com.
To help capture your feedback effectively, share the following information in your email:

- Document name
- Version
- Chapter, topic, or section
- Feedback or suggestions

Introduction to IDP Studio

IDP Studio is a visual data science platform designed to automate the extraction, understanding, and processing of information from various types of documents. It can read and extract typed content from the documents. Based on the content in documents it can categorize and save them to a specified location or directory. Additionally, it can extract meaningful information from documents, facilitating the training of Neural Network (NN) and Machine Learning (ML) models.

Business use case

Refer to the following use case to gain a better understanding of the NewgenONE IDP Studio platform.

Problem Statement: To classify documents under different classes or types and extract the required feature values from them.

The following are the classification of documents:

- BillOfExchange
- CoveringSchedule
- CertificateOfOrigin
- CommercialInvoice
- BillOfLading
- InsuranceCertificate
- PackingList
- PerformaInvoice
- AirWayBill
- PurchaseOrder

Each of the documents has some different field (feature) values that are to be extracted. For example, A document for a certificate of origin has the following fields that need to be extracted:

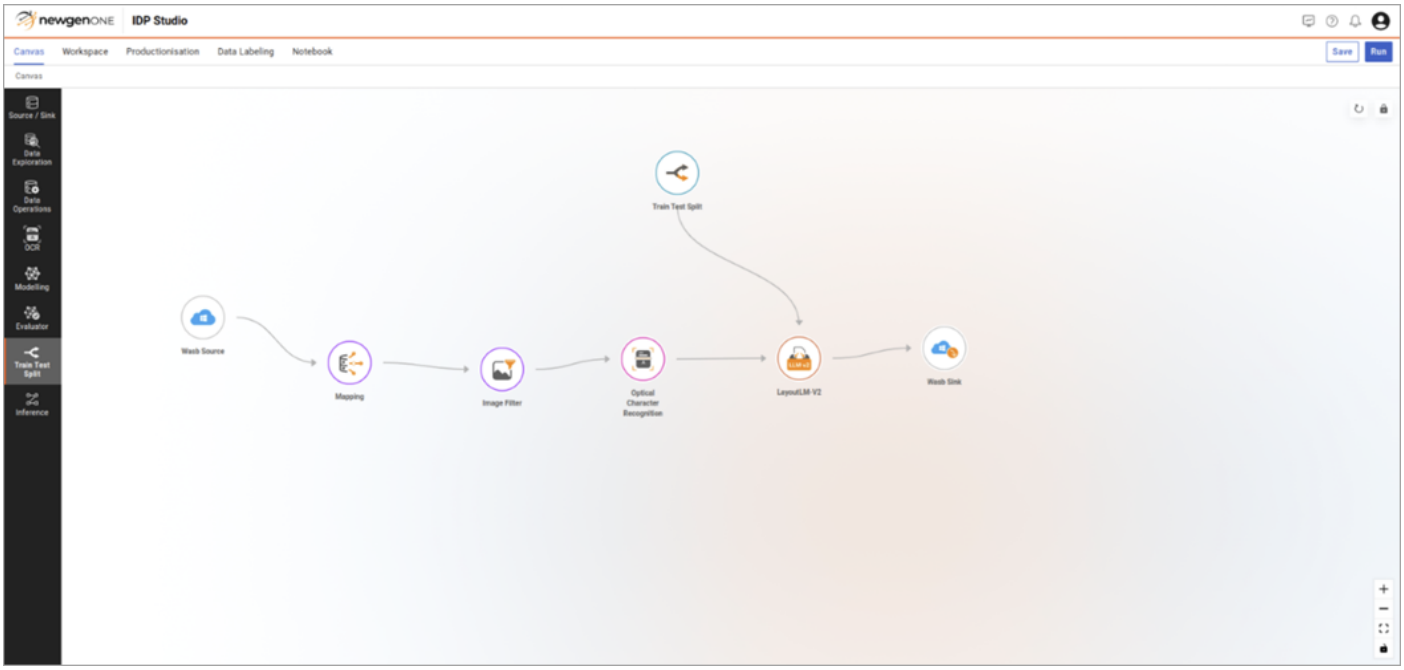
- declaration_by_exporter
- page_no
- invoice_no
- ref_no
- consignee_name
- consignor_address
- means_of_transport
- net_weight
- original_number
- certificate_no
- coo_issuer_address"

Pipelines: Development and modelling

A document classification pipeline is being created.

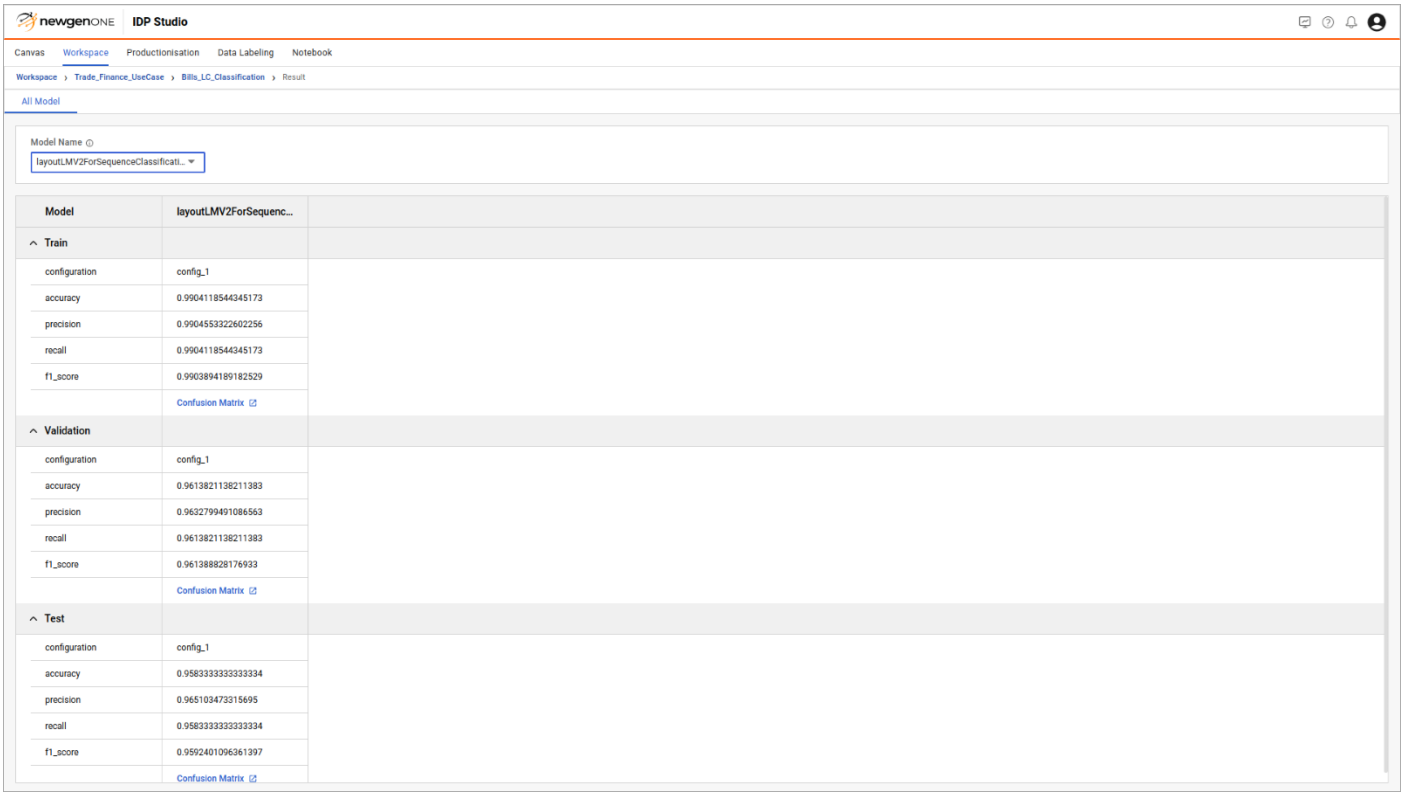
1. **Document classification model training (dev) pipeline**:
    a. Utilizing WASB (Windows Azure Storage Blob) as the data source.
    b. The second node in the pipeline maps the folders containing the images of the documents to a class. As this is the model training (Dev) pipeline, the folders containing the scanned copies of documents (images) and their respective classes are predetermined. The preceding node to this is the WASB source node.
    c. The image filter node is next in the pipeline, where noise is removed from the images using the Remove Noise parameter. This step eliminates pixel irregularities. The preceding node to this is the Mapping node.
    d. The subsequent node used in the pipeline is OCR, as the LLM-V2 (LayoutLM-V2) model for document classification, which includes OCR as a primary component is used. Various OCR engines like Tessaract, GV, Amazon, or ABBYY can be selected by the user. In this case, GV (Google Vision) is selected in the OCR node settings. The preceding node to this is the Image Filter.
    e. The LayoutLM-V2 model (found under the Document Classification category in the Modelling section) is used here. It employs image and OCR data to create embeddings and features for document classification. Therefore, the subsequent node after OCR is for LayoutLM-V2.
    f. An Apply node is used for Train Test Split on the model, where the data is split into training, validation, and test data sets.
    g. Finally, there is a sink where the output of the pipeline, which can be images or models, is stored. WASB is used as the data sink.

This completes the document classification pipeline. Refer to the below screenshot.



Save and run the pipeline.

Following is the result screen:



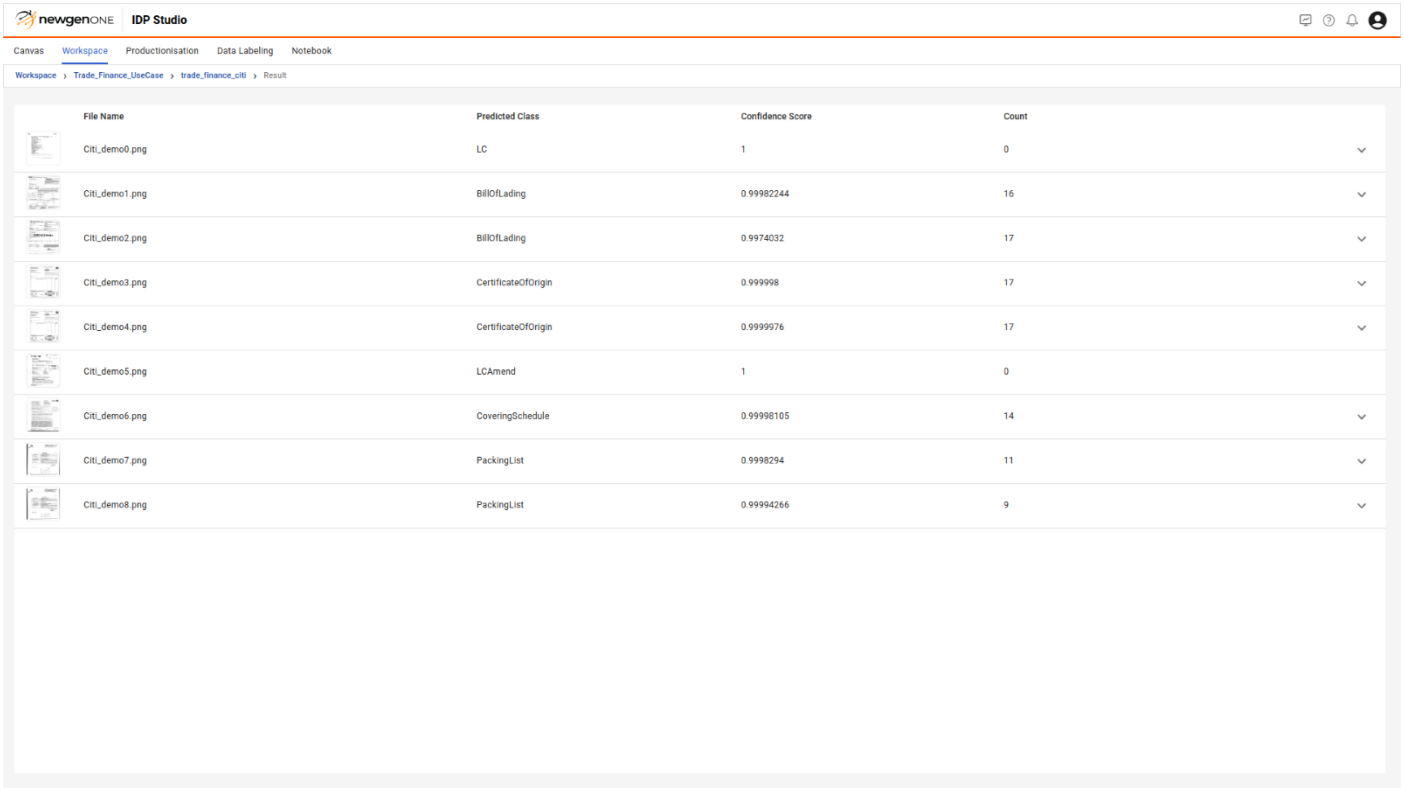
2. **Feature Extraction Model training (Dev) pipeline**:
   Now, feature extraction model development and training pipelines are being created.

a. Here, WASB as the data source is used. Select the folder for the class for which the extraction pipeline is being created.
b. The next node in the pipeline is the image filter where we are removing noise from the images using the “Remove Noise” parameter in the node. It removes the contours of pixel irregularity. The preceding node to this node is the “WASB” source node.
c. The next node used in the pipeline is OCR to use the LLM-V2 (LayoutLM-V2) model for document extraction, which has OCR as one of the main components. The user can select various OCR engines like Tessaract, GV, Amazon, or ABBYY. In this case, GV (Google Vision) is selected in the OCR node settings. The preceding node to this node is Image Filter.
d. Here, the LayoutLM-V2 (the one found under the Document Extraction category in the Modelling section) model is used which uses the image and OCR data to create embedding and features to classify documents. So, the next node to the OCR node is for LayoutLM-V2.
e. Apply node for Train Test Split on the model where it splits the data into training, validation, and test data.
f. Finally, there is a sink where the output of the pipeline, which can be images or models, is stored. WASB is used as the data sink.

This completes the document classification and feature extraction pipeline.

Save and run the pipeline.

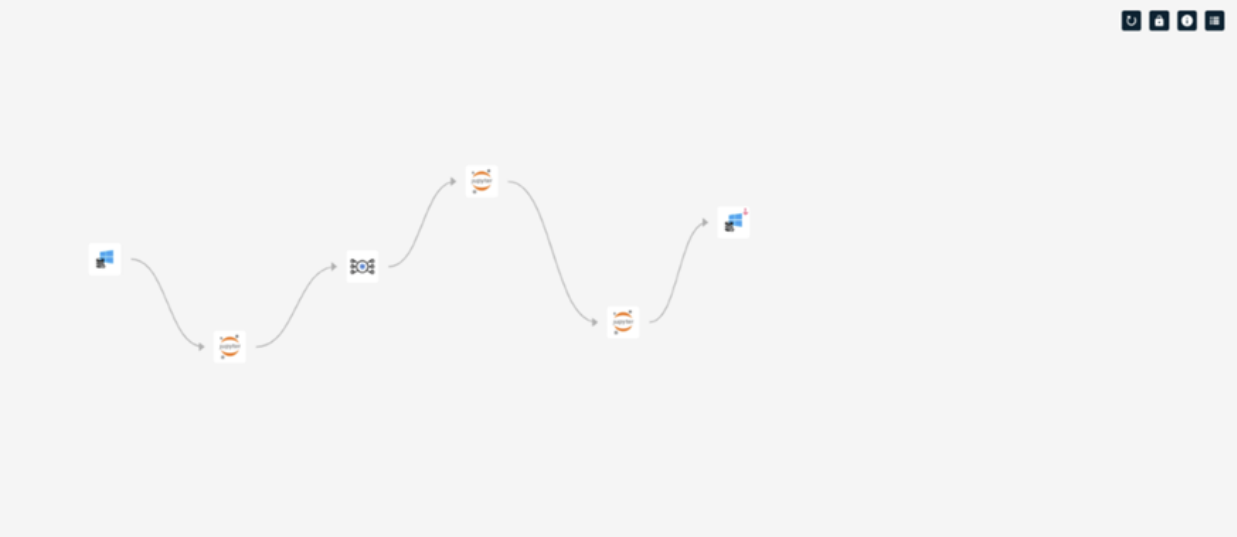
Following is the result screen for such a pipeline:



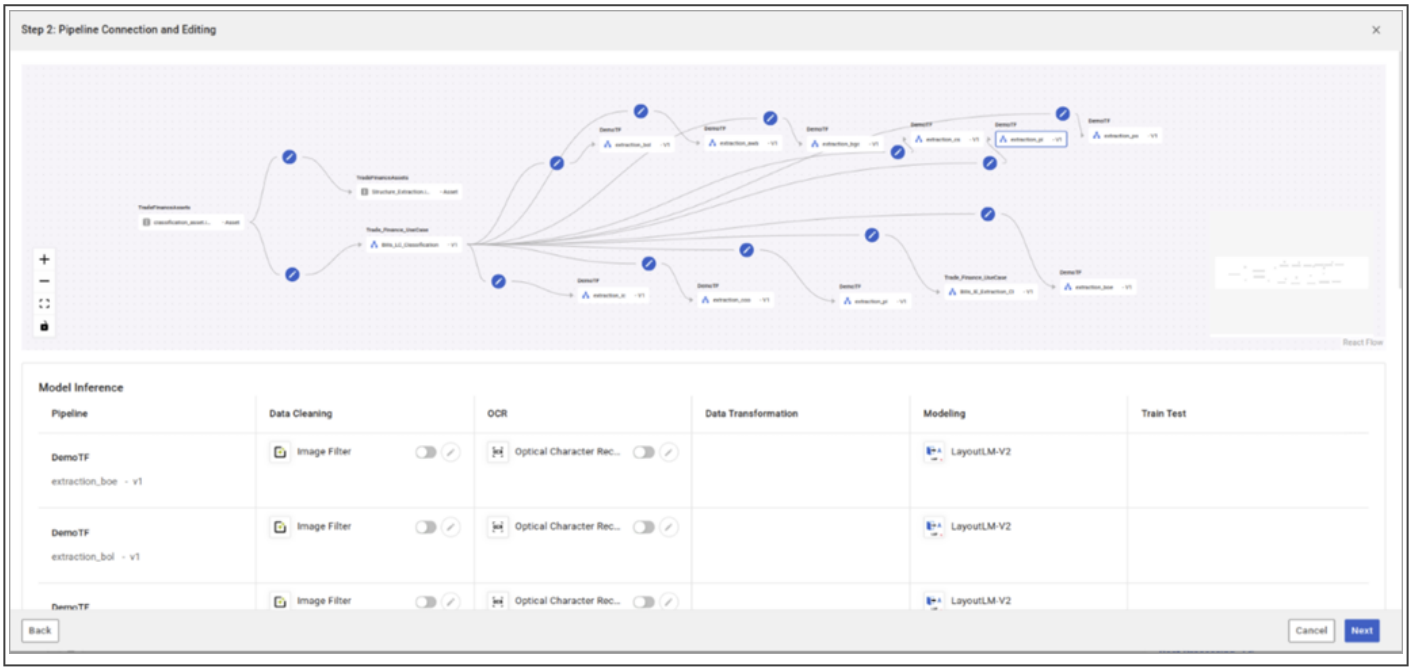
Create one extraction pipeline for each of the classes given in the problem statement. Since there are ten classes, there must be ten pipelines for data extraction.

3. **Jupyter Notebook APIs**: This APIs is required during inference. To create such APIs refer to the Notebook section. Different Jupyter Notebook APIs are created for classification as well as for result transformations.
    a. **Structured Classification**: This Asset is written in Jupyter Notebook which uses OCR data and regex is applied to classify documents.
    b. **Structured Extraction**: This Module is purely based on Computer Vision Techniques using OpenCV. No model is used in this. This Extraction consists of the extraction of LC documents. LC Documents are extracted using regex based filtering using the OCR all text data. LC Application, LC Ammendment, and Export Lodgement Letter documents are extracted using template matching. In this, we detect blocks perform line segmentation, and extract information from each block.
    c. **Indexing and Result Transformation**: Here indexing and set generation are created and the result is transformed so that it can be consumed by Trade Finance Desktop.
4. **Inference Pipeline**:

a. This is the inference pipeline which consists of a source, sink, inference engine, and multiple Jupyter notebooks
b. Sources: Here we select the images or PDF on which we want to do inference
c. Inference Engine: In this Engine, we can create a bundle inference where we can combine multiple models and notebooks API
d. Jupyter Notebook: Each notebook is defined for different tasks like classification, extraction, and result transformation.



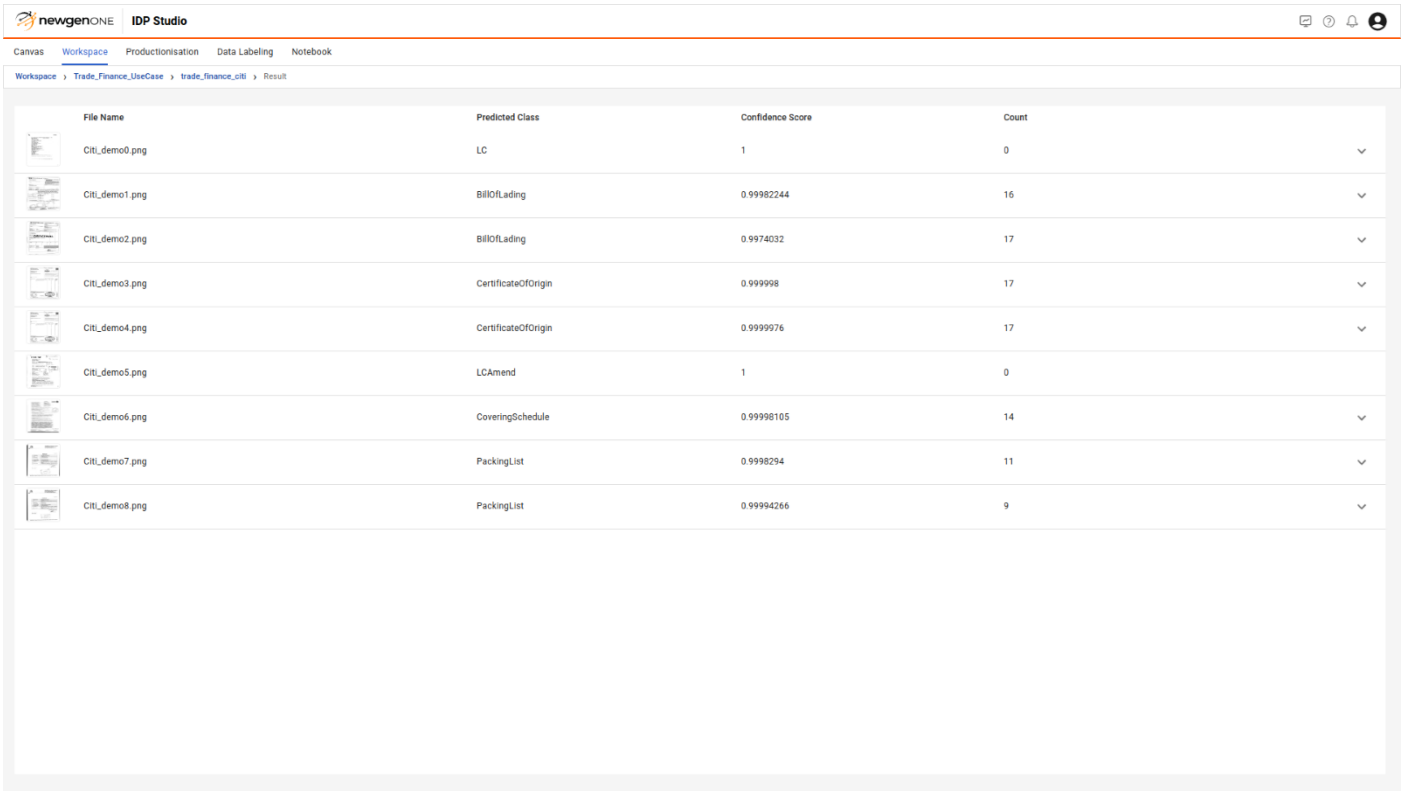
Data in the wasb source in this use case is a PDF containing images of different documents.



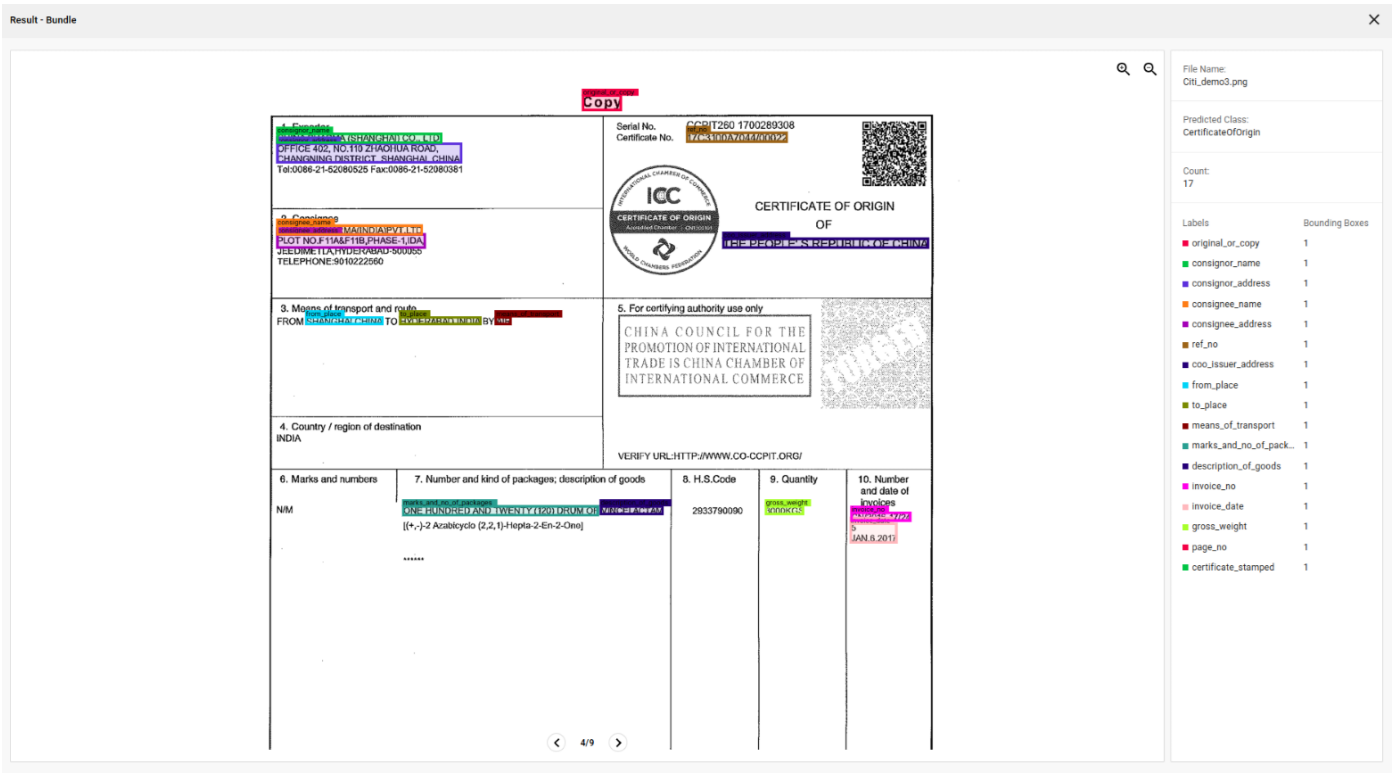
It is not model retraining in the pipeline nodes inside the inference engine node here, but only the trained models are used in the pipeline.

The above image shows the connection inside the inference engine. This engine first classifies documents and extracts information from each document based on the document classified.

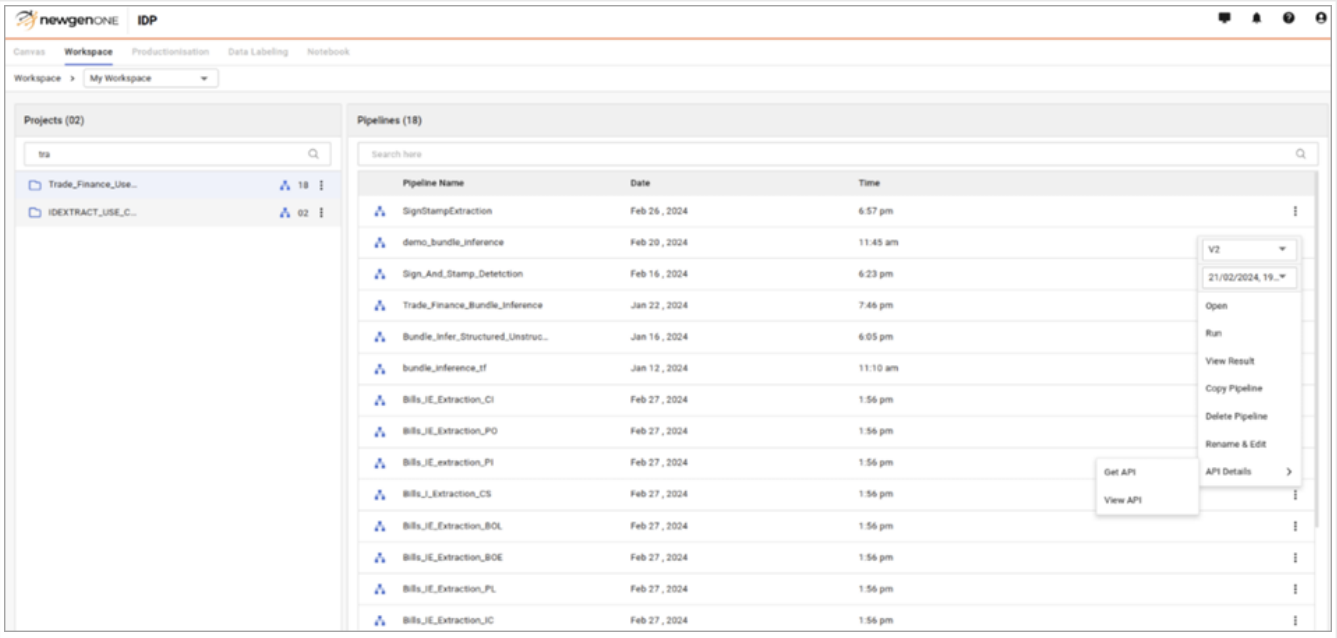
Result Screens:

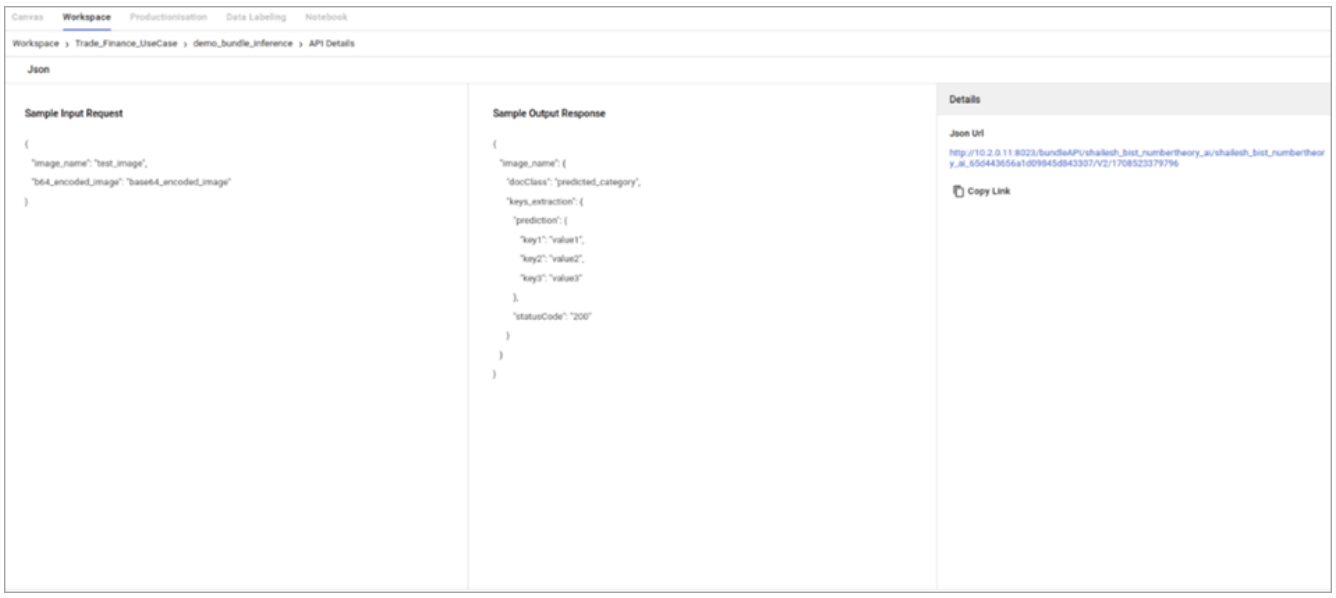


File Name	Predicted Class	Confidence Score	Count
Citi_demo0.png	LC	1	0
Citi_demo1.png	BillOfLading	0.99982244	16
Citi_demo2.png	BillOfLading	0.9974032	17
Citi_demo3.png	CertificateOfOrigin	0.999998	17
Citi_demo4.png	CertificateOfOrigin	0.9999976	17
Citi_demo5.png	LCAmend	1	0
Citi_demo6.png	CoveringSchedule	0.99998105	14
Citi_demo7.png	PackingList	0.9998294	11
Citi_demo8.png	PackingList	0.99994266	9



Each Pipeline can be exposed as an API using GetAPI and View API:





The API address can be used in any API client like Postman.

Accessing IDP Studio

The IDP Studio is accessible online through any of the recommended web browsers. This chapter describes the prerequisites and the process of signing in to IDP Studio.

Prerequisites

To access IDP Studio, you must meet the following prerequisites:

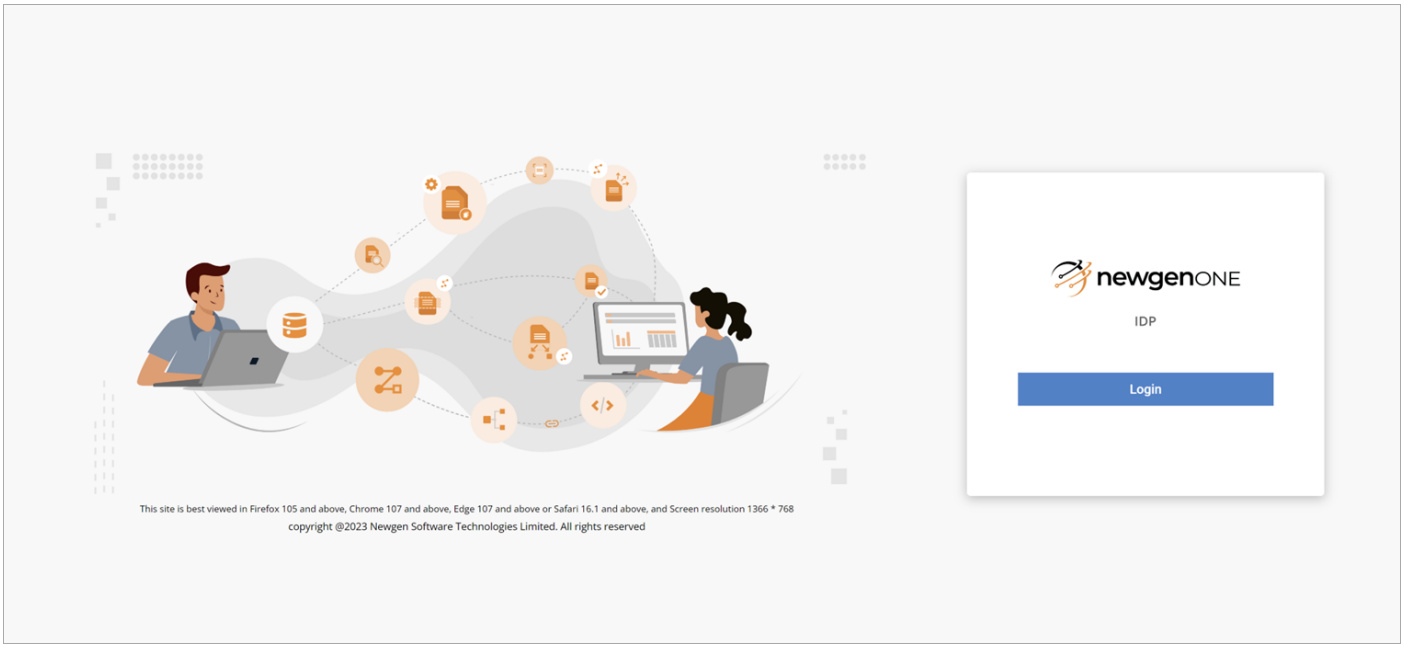
- Valid IDP Studio URL
- Registered username and password
- Supported browsers are:
  - Microsoft Edge 113 and later
  - Google Chrome 113 and later
  - Mozilla Firefox 113 and later

Signing in to IDP Studio

To sign in to IDP Studio, perform the following steps:

1. Launch a web browser.
2. Enter the IDP Studio URL in the following format:
   HTTP or https://<Domain or IP address>
   Here, the *<Domain or IP address>* represents the location of the server hosting the IDP Studio application and its services.
   For example,
   https://sample.abcsoftware.net/

The NewgenONE IDP Studio Login page appears.



3. Enter your Username and Password and click the **Login** button. On successful sign in, the landing page appears.
4. In case of On-Premise deployment enter the username and password
5. Click the **Login** button. Your login username must be username@yourdomain.com. On successful login, the landing page appears.

In case the IDP Studio is deployed on Microsoft Azure Cloud, make sure that you are logged out from any other Microsoft account on the same browser. You can also use the private or incognito mode of the same browser, or switch to a different browser. Ensure that the browser is not blocking the third-party cookies.

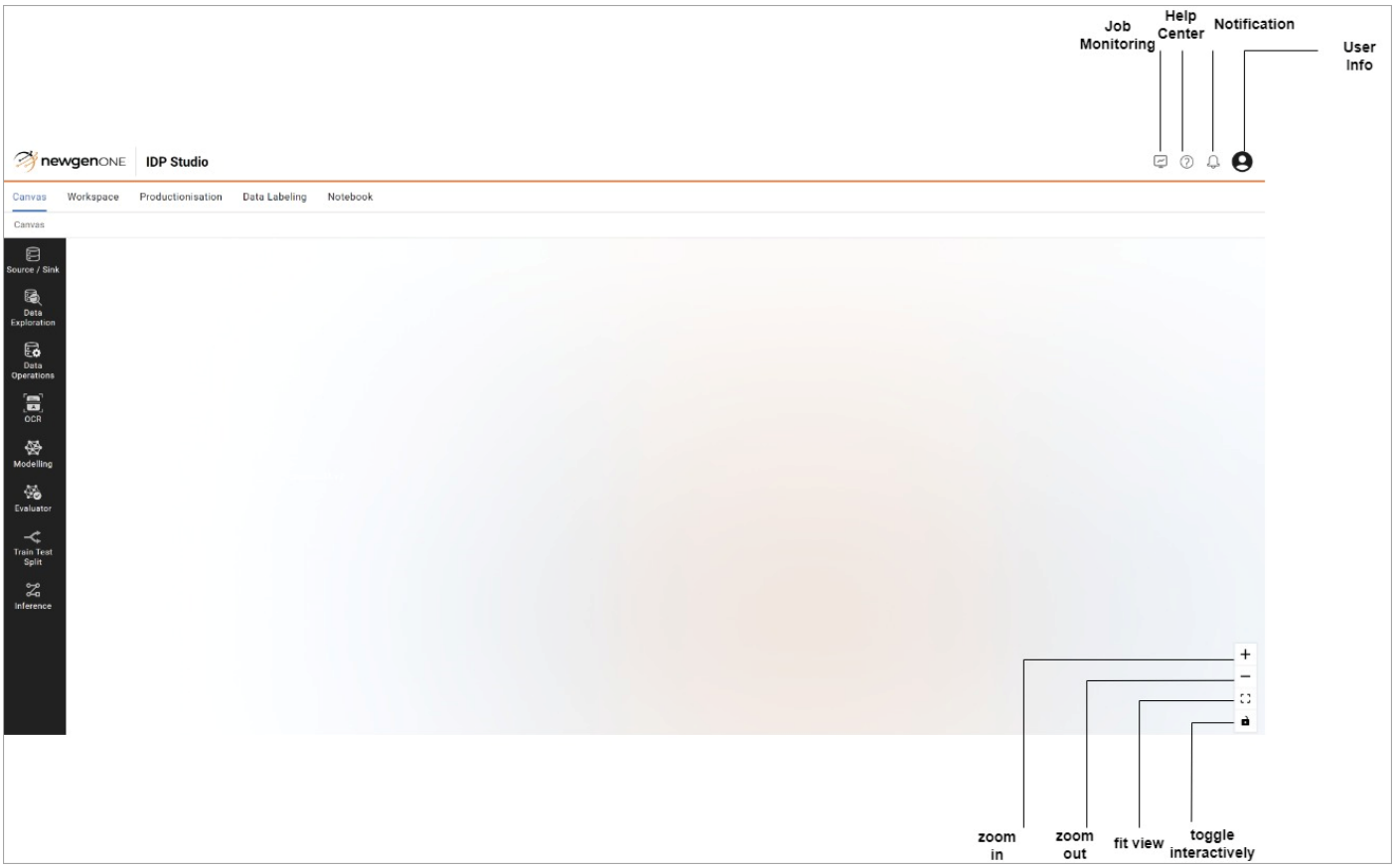
Exploring IDP Studio

When you sign into IDP, you see the following navigation options on the landing page:

- Navigation options are located on the upper row.

  The Canvas option selected by default.

- The options under the Canvas window provide different types of node clusters which consist of various nodes used to construct pipelines.
- The middle area of the canvas allows you to drag and drop the required nodes and then connect them to design a pipeline.
- The options on the upper right corner of the screen provide additional navigation, along with the option to sign out from the application..



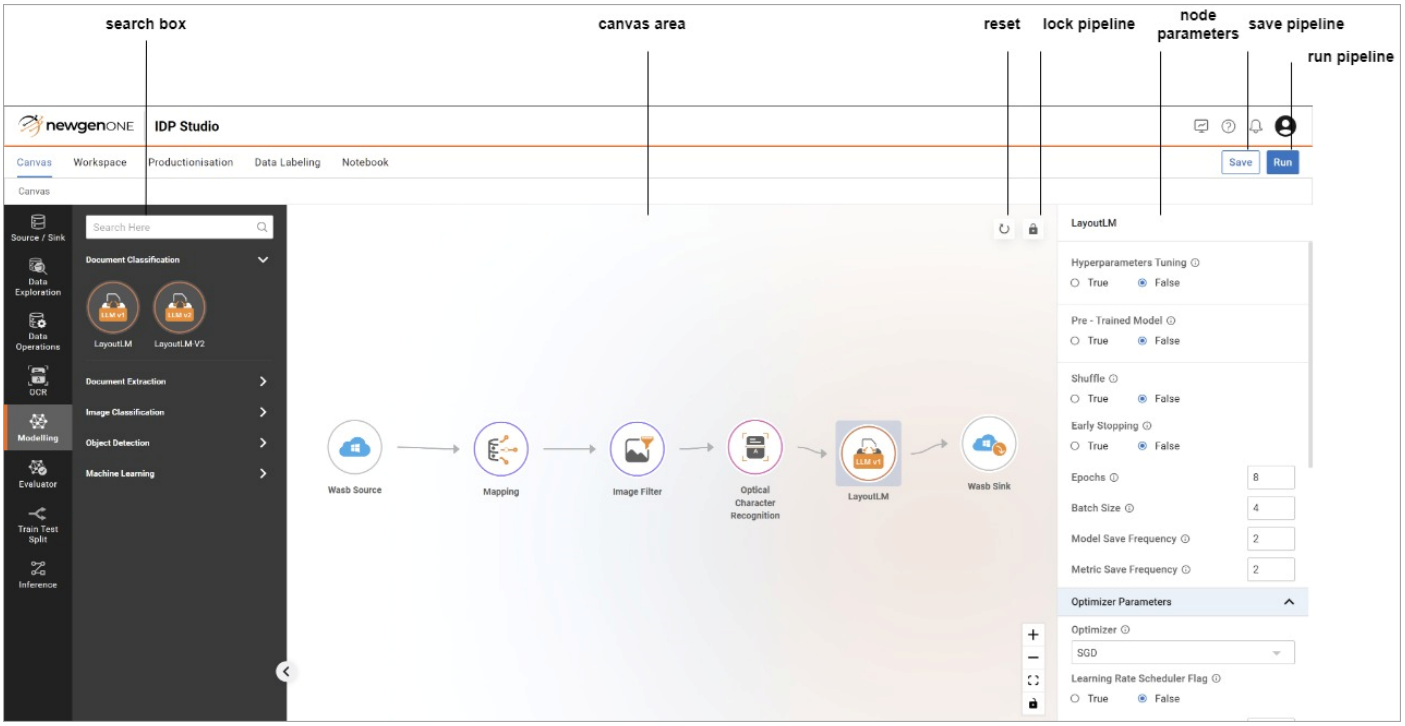
Following are the additional options available on the landing page:

Option	Description
Job Monitoring	Allows you to monitor the pipeline executions. For procedural details, see Designing model pipelines.
Notification	Allows you to browse the notification of pending approval to publish a pipeline. For procedural details, see Productionisation. This option is visible to all users, but only selected administrators by the user receive the notifications when publishing a pipeline.
Help Center	Allows you to open the document and training concepts to understand the IDP Studio.
User Profile	Allows you to sign out from the IDP Studio platform.
Toggle interactively	Allows you to toggle between grouping or clearing grouping of pipeline nodes and connectors.
Fit View	Allows you to fit a pipeline into the canvas area.
Zoom Out	Allows you to zoom out of the canvas area.
Zoom In	Allows you to zoom in on the canvas area.

Designing model pipelines

Canvas allows you to design pipelines. This page shows the following options:

Option	Description
Search box	Allows you to search the nodes by their name.
Canvas area	Allows you to drag and drop nodes on it where nodes can be connected.
Reset	Allows you to reload the canvas and reset it to blank.
Lock pipeline	Allows you to lock the pipeline so any updates made to a node's parameters do not affect the succeeding nodes.
Node parameters	Allows you to fill parameters for the selected nodes on the canvas area.
Save	Allows you to the save pipeline. For procedural details, see saving a pipeline.
Run	Allows you to execute the pipeline. A pipeline must have at least one saved version to run.

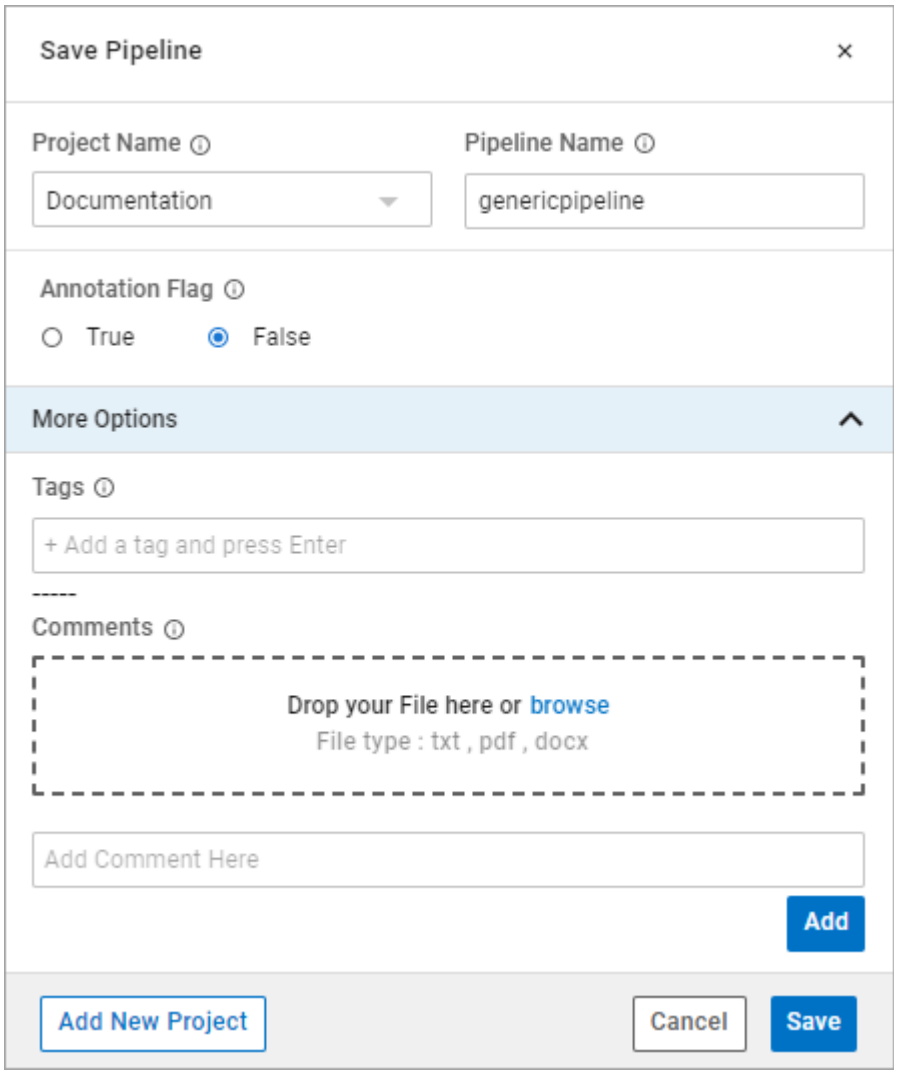


Saving a pipeline

The section describes the procedure for saving a pipeline.

To save a pipeline, perform the following steps:

1. From the Canvas, click the **Save** button on the upper-right corner. The Save Pipeline dialog appears.



2. Specify the following details:

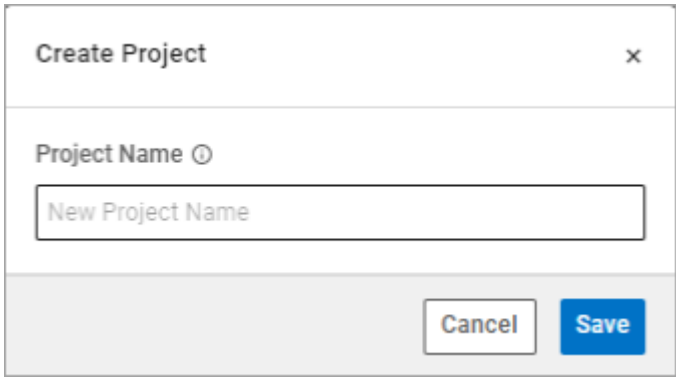
Option	Description
Project Name	Allows you to select the Project Name in which you want to save the pipeline. To save in a new Project Name, click **Add New Project.** For procedural details, see Add New Project.
Pipeline Name	Allows you to enter the Pipeline Name.
Annotation Flag	Allows you to create a new project under Data Labeling. For procedural details see Data Labeling. • The default value is False • Annotation Flag appears only for the non-model pipelines
More Options	Allows you to add more details to the pipeline. • **Tags** — Allows you to add tags in the pipeline. Categorize and organize pipelines with keywords for easy management and filtering. • **Comments** — Allows you to add comments to the pipeline by either dropping the file or browsing for it.

3. Click **Save** to save pipeline configurations.
   If a pipeline with the same name already exists, an alert message appears. In that case, either rename the pipeline or click **Create New Version** to create a new version, or click **Update** to save the modified details.

Adding a new project

To add a new project, perform the following steps:

1. Click the **Add New Project** button, the Create Project dialog appears.



2. Enter the **Project Name** and click **Save**.

Saving an existing pipeline

To save an existing pipeline, perform the following steps:

1. From the Canvas, click the **Save** button on the upper-right corner. The Save Pipeline dialog appears
2. Toggle the Annotation Flag to **True** or **False** as required.

   Annotation Flag option is available when you try to save a pipeline without a model node.

3. Click **Update** to save the updated changes in the pipeline.

Opening an existing pipeline shows you the following options in the upper right corner of the canvas area:

Option	Description
Reload	Allows you to reload the canvas and reset it to blank.
Lock Pipeline	Allows you to lock the pipeline so any updates made to a node's parameters do not affect the succeeding nodes.
Details	Allows you to see pipeline details such as project name, pipeline name, and version.
Go to Workspace	Allows you to navigate to the My Workspace.
Select Device	Allows you to select whether to execute the pipeline on CPU or GPU. The default value is CPU. This option appears only when the pipeline has a model node(s).

Monitoring pipelines

Monitoring pipelines allow you to check the status of the running pipeline. To access the pipeline monitoring, open the canvas and click the Job Monitoring icon .

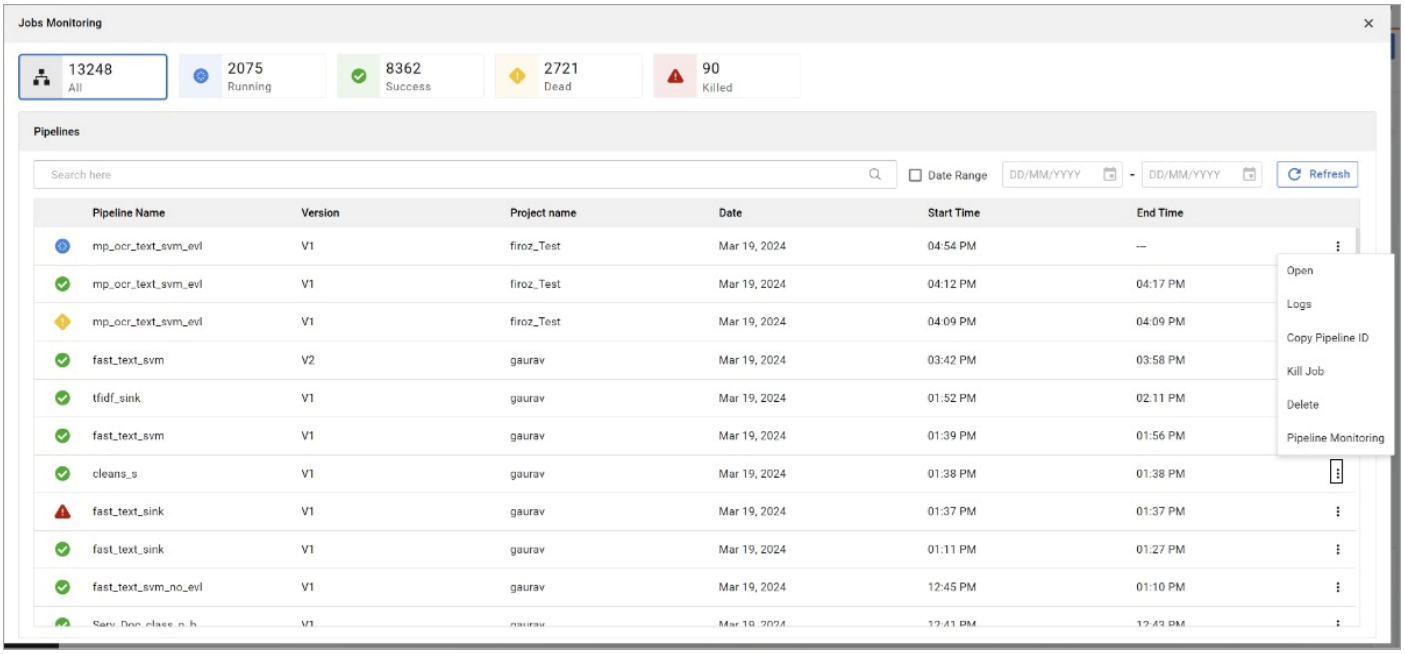
Specify the following details:

Options	Description
Pipeline status	Allows you to monitor all the pipeline status. Indicates the pipeline status as follows: • **All** — Shows the count of all available pipelines. • **Running** — Shows the count of running pipelines. • **Success** — Shows the count of successfully executed pipelines. • **Dead** — Shows the count of pipelines that encountered errors during execution. • **Killed** — Shows the count of pipelines stopped by the system or the signed-in user. For procedural details, see Operations on the existing pipeline.
Search box	Allows you to search any pipeline by its name.
Date Range	Allows you to view the pipelines by applying the date filter using the date range.
Refresh	Allows you to reload the pipeline data.
More options	Allows you to perform multiple operations such as viewing logs and opening pipelines in Canvas. For procedural details, see Viewing an executed or dead pipeline.

Operations on the existing pipeline

This section describes the various operations that can be performed on the pipelines on the monitoring screen.

On the Pipeline Monitoring dialog, click the More Options icon ⋮ against the required pipeline to open a dropdown list.



The options are as follows:

Option	Description
Open	Allows you to open the pipeline in the canvas and modify the node information according to your requirements.
Logs	Allows you to display the logs that consist type of error along with its description, including the date and time of the error.
Copy Pipeline ID	Allows to copy the pipeline ID.
Kill Job	Allows you to stop the running pipelines. This option appears only for the currently running pipelines.
Delete	Allows you to delete the pipeline from the monitoring.
Pipeline Monitoring	Allows you to monitor a pipeline in view-only mode. This view lets you know which node was executed successfully and which one got failed. • The color codes on the node show failed, currently executing, and successful nodes. • When a node fails while executing a pipeline, the pipeline becomes dead. In this case, succeeding nodes in the pipeline do not execute.

Viewing an executed or dead pipeline:

Click the More Options icon ⋮ against the pipeline and then click the pipeline monitoring option from the dropdown list to monitor the pipeline. The Job Monitoring screen appears.

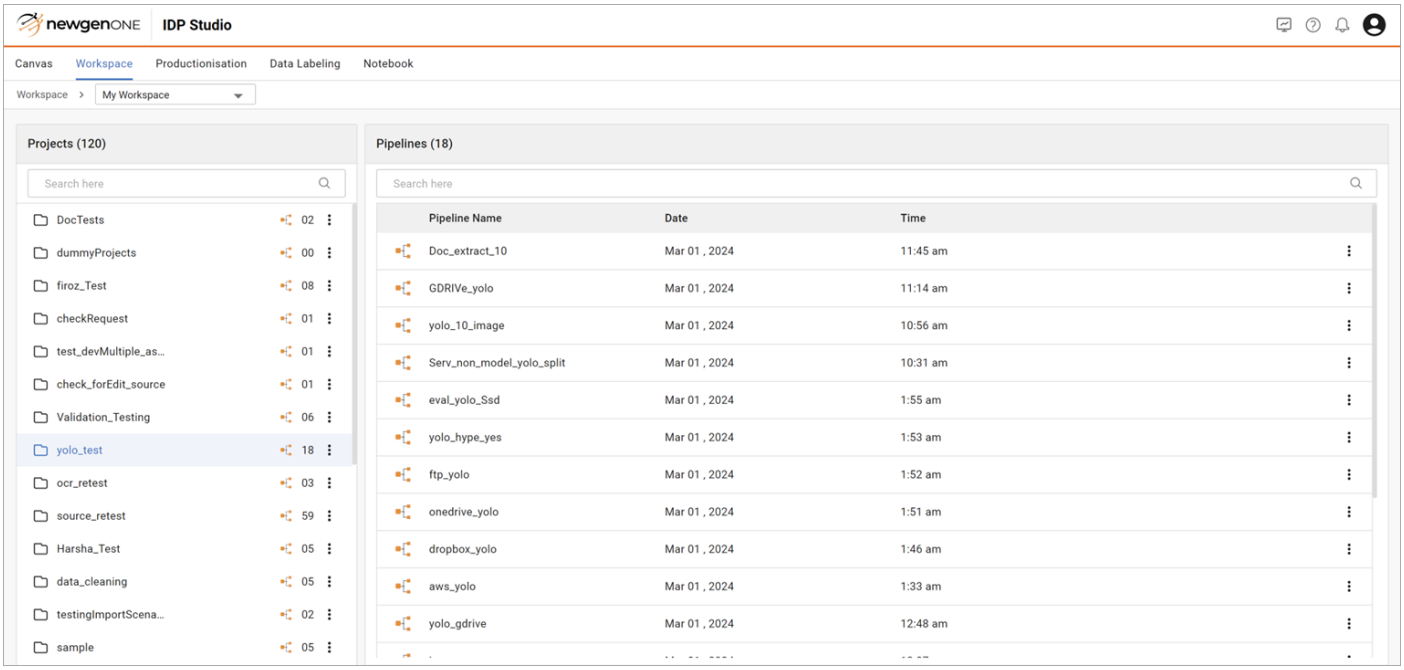
To see the latest status of the pipeline, click the **Refresh** button. On clicking any node, the parameters of that node appear on the right-hand side of the screen. Above the parameters, the Sample Result button appears. By clicking this button, you can view the results produced or produced by the selected node. The Sample Result button is disabled for the following list of nodes:

- All source data nodes
- AI sink nodes
- Evaluator nodes
- Data Exploration nodes
- Train test split
- GeoLayoutLM
- Table Transformer
- Table Detection
- SVM
- Fast Text
- TFIDF
- Mapping
- Inference
- YOLO V8

- When a pipeline is executed successfully, all nodes turn green. You can click on any of the nodes and then click the **Sample Result** button to see the result.
- When a pipeline execution fails, the failed nodes turn yellow. For the failed node the Sample Result button is disabled.

Managing projects and pipelines

The Workspace tab allows you to manage projects and pipelines.



To access the Workspace, perform the following steps:

1. Open the NewgenONE IDP Studio using the credentials. The landing page appears.
2. Click the **Workspace** tab located in the menu at the top.

The Workspace page appears with the following options:

Option	Description
Workspace	Allows you to select any of the following workspaces using the dropdown menu: • **My Workspace** — Comprises the list of projects created by signed-in users. My Workspace appears by default. • **Tenant** — Comprises the list of projects shared with all users in the application. ◦ **Import** — Allows you to import the pipeline into projects in your workspace. For procedural details, see Importing a pipeline. ◦ **Delete** — Allows you to delete a specific pipeline. For procedural details, see Deleting a pipeline. • **Shared With Me** — Comprises the list of all pipelines shared with the signed-in user by other users of the application. ◦ **Import** — Allows you to import the pipeline into projects in your workspace. For procedural details, see Importing a pipeline. ◦ **Delete** — Allows you to delete a specific pipeline. For procedural details, see Deleting a pipeline.
Projects	Shows the projects present in the selected workspace and allows you to perform the following operations: • **Search** — Allows you to search the project by its name. • **Rename** — Allows you to change the project name by clicking the more options icon ⋮ against the project name. • **Delete** — Allows you to delete the project by clicking the More Options icon ⋮ against the project name. The number against each project name denotes the number of pipelines in that project.
Pipelines	Displays the pipelines present in the selected project and allows you to search the pipeline by its name. The Date and the Time columns are for the creation date and time of the pipeline. For procedural details, Operations on the existing pipeline.

Importing a pipeline

To import a pipeline into projects in your workspace, perform the following steps:

1. Click the More Options icon ⋮ against a specific pipeline and select **Import**. The Import dialog appears on the right pane.
2. Enter the **Pipeline Name**.

3. Select the **Project Name** from the list.
   Alternatively, create a new project by clicking **Create New Project** to save the pipeline within a new project.
4. Click the **Confirm** button to import the pipeline.

Deleting a pipeline

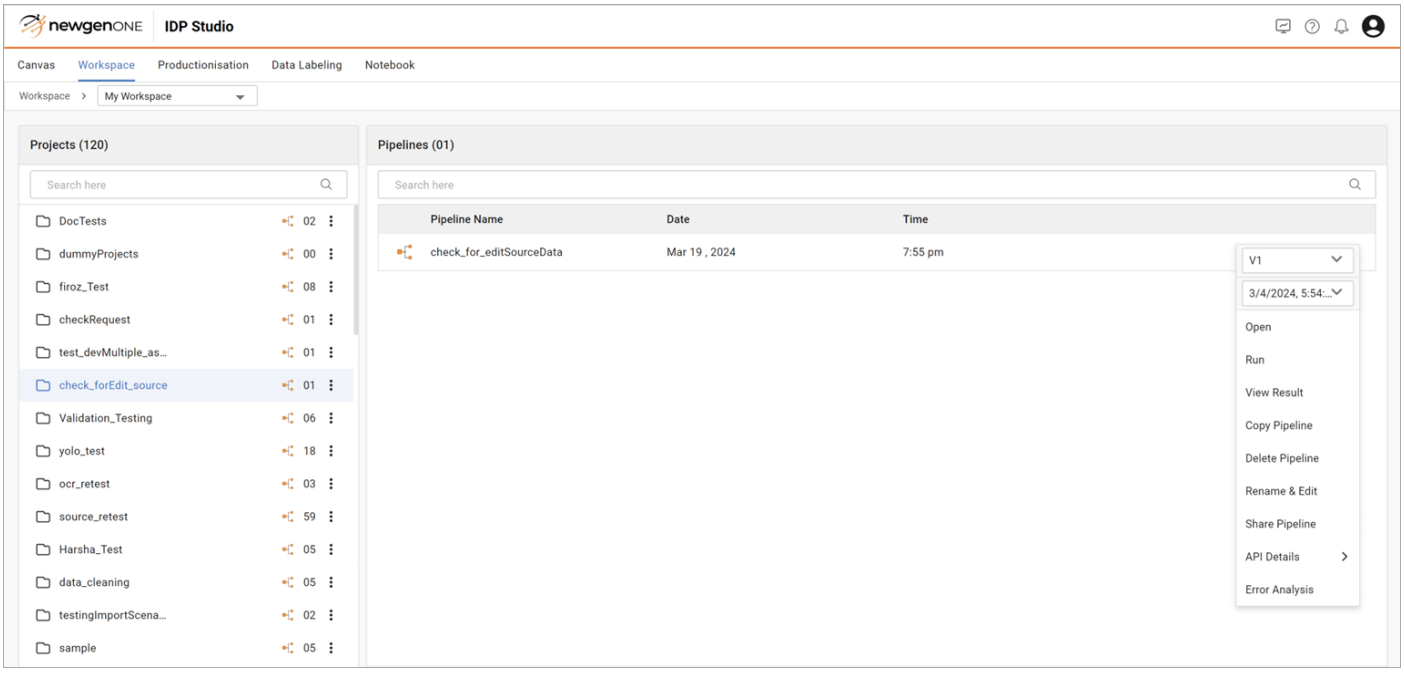
To delete a pipeline, perform the following steps:

1. Click the More Options icon ⋮ against a specific pipeline and select **Delete**. The Delete dialog appears.
2. Click **Delete** to confirm the deletion of the selected pipeline.

Managing pipelines

This section describes the various operations that can be performed on the pipelines on my workspace screen.

Click the More Options icon ⋮ to see more options against a pipeline.



The More Options dropdown comprises the following options:

Option	Description
Version	Allows you to see the version on the pipeline. You can change the version of the pipeline if it has multiple versions.
Date Time	This is the date and time of the successful execution of the pipeline.
Open	Allows you to open the pipeline in the canvas to see more information.
Run	Allows you to execute the selected pipeline.
View Result	Allows you to see the results of the last execution of the selected pipeline version.
Copy Pipeline	Allows you to copy the selected version of the pipeline into another project.
Delete Pipeline	Allows you to delete the selected pipeline from the project.
Rename & Edit	Allows you to edit the pipeline details including the pipeline name.
Share Pipeline	Allows you to share the successful run of the pipeline with a user or tenant (all the users of the application). • The successful pipeline run can be selected by selecting the appropriate datetime using the date time dropdown. • This option appears when the pipeline is executed successfully at least once.
API Details	Allows you to get the API for the pipeline. Select the appropriate successful pipeline run by selecting the respective date time from the dropdown menu. For procedural details, click here.
Error Analysis	Allows you to view the details of the training errors. Every successful pipeline run having a modeling node in it has the error analysis option.

To view the API details, perform the following steps:

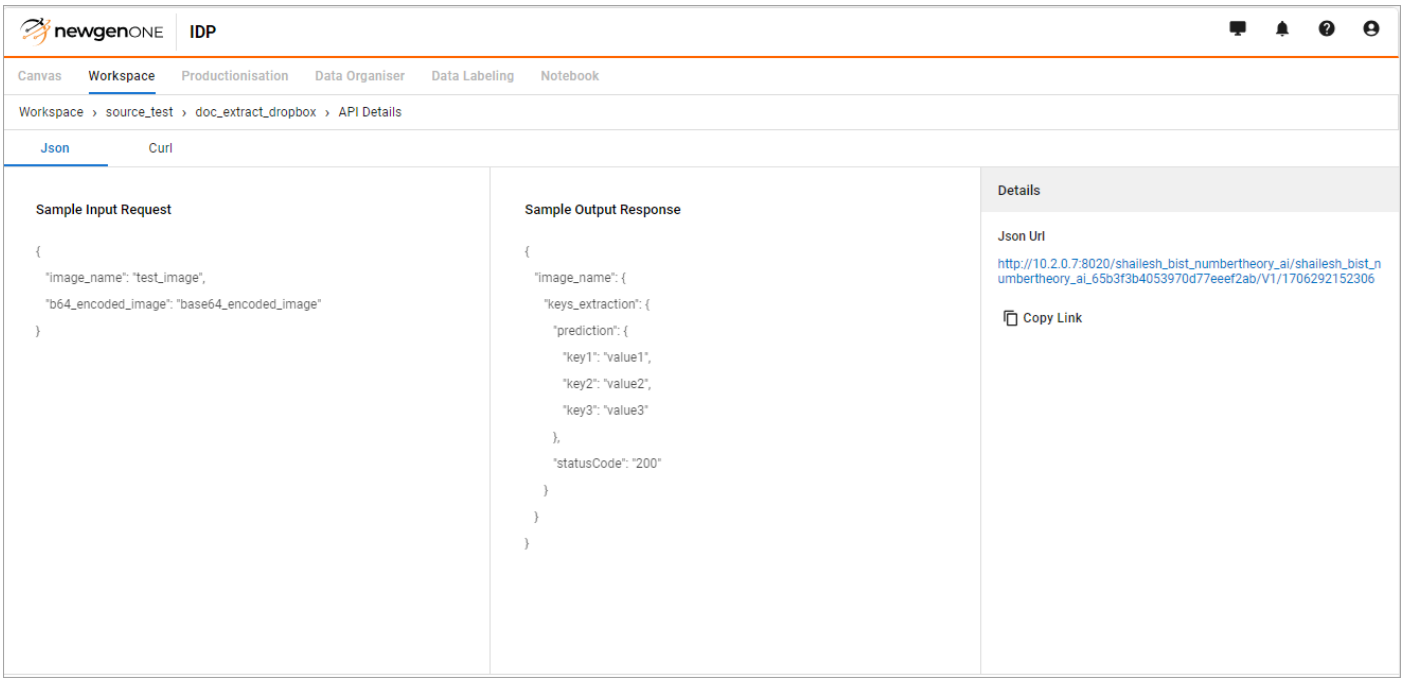
1. On the Workspace page, click the more options icon ⋮ against a required pipeline. The list of operations for the selected pipeline appears.
2. Click **API Details** and select any of the following options:
   - **Get API** — Allows you to generate the API end point from the successfully executed pipeline.

- Successfully executed pipeline having a modeling node in it have the API Details option,
- In the case of the inference pipeline, every successfully executed pipeline has an API Details option.

- **View API** — Displays the API details.

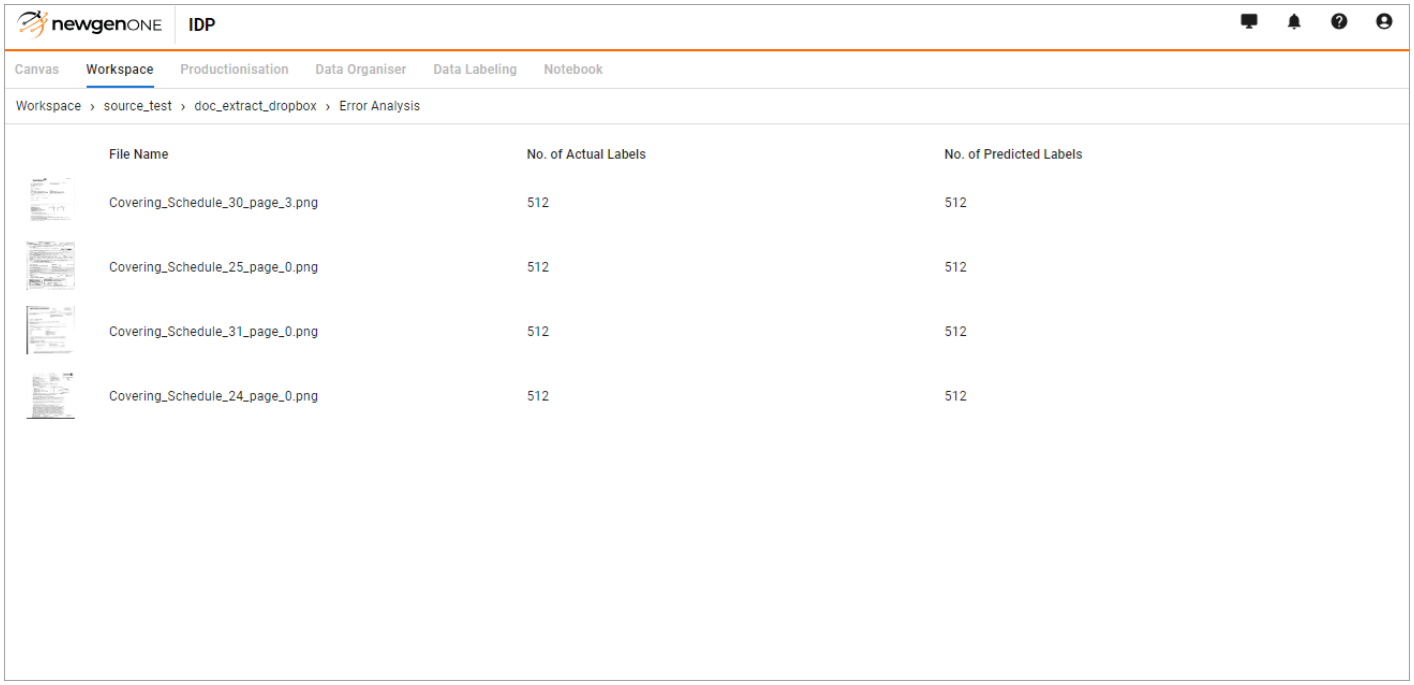
Viewing API details

The API Details page displays detailed information about the API endpoint generated from a successfully executed pipeline with a modeling node.



Analyzing errors

Error Analysis page provides a comprehensive overview of training errors encountered during the execution of a modeling node in a successful pipeline run.



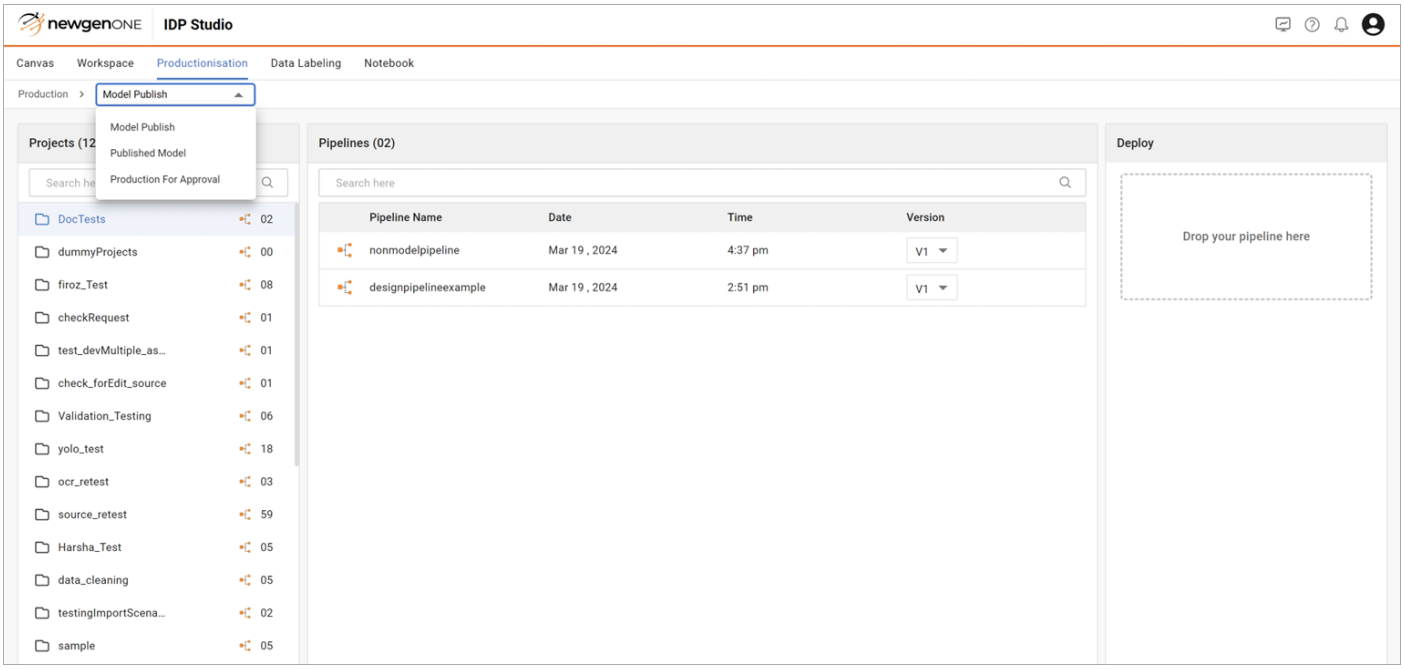
Productionisation

Model retraining is done on production data in the productionisation environment. The productionisation section manages the pipeline scheduling. With the scheduling, the pipeline can be served for production data.

Click the **Productionisation** tab located in the menu at the top.

This section consists of the following topics:

- Model Publish
- Published Model
- Production for Approval



Publishing a model

Model Publish allows you to schedule a single or multiple pipelines within your workspace simultaneously. This is accessed with the **Model Publish** option (which is the default option) under the Productionisation tab.

To publish a model, perform the following steps:

1. Click the **Productionisation** tab and select the Model Publish option. The Model Publish page appears.
2. Select the required project.

   You can search for the project by its name using the search box.

3. Drag the required successful pipeline version with the evaluator node, then drop it into the Deploy section located on the right pane.

   You can search for the pipeline by its name using the search box.

4. Click **Deploy**. The following options appear to set your scheduling parameters:

Option	Description
Project Name	Allows you to enter the project name for production.
Change Frequency	Allows you to set the frequency according to your requirements. Possible options are daily or specific. The default value is daily.
Start Date	Allows you to set the start date using a date picker.
End Date	Allows you to set the end date using a date picker. • The end date must be at least 30 days for monthly scheduling and 7 days for weekly scheduling. • There is no start date in case of specific change frequency.
Select Duration	Allows you to select from weekly or monthly. This option appears if you have changed the frequency to specific.
Run Hourly Basis	Allows you to run the pipeline on an hourly basis.
Admin Users	Allows you to select the administrators using the drop-down menu. The administrators you select here further approve your pipelines for deployment. Any user in the system can be assigned as admin.
Enter Time	Allows you to enter the time when you want to run the pipeline in a day.

5. Click **Deploy**. The pipeline gets deployed.

After deploying the pipelines, all administrators receive notifications for approval. Once approved, the pipelines appear in the Published Models list.

Viewing a published model

The Published Model provides the list of projects and pipelines, along with the details of the project that is deployed to production.

To view the published pipeline or model, perform the following steps:

1. Go to the Productionisation tab.
2. Click the **Published model**. The Published Model page appears with the following tabs:
   • Projects — Comprises the list of all projects.
   • Pipelines — Comprises the list of all pipelines available on selected projects.
3. Select the required project. The deployed pipelines within the selected project appear on the right pane.

   You can search for the project by its name using the search box.

4. Select the required pipeline.

   You can search for the pipeline by its name using the search box.

5. Click the More Options icon ⋮ given next to the pipeline.
6. Click the **View Pipeline** option. The published pipeline page appears with the following details:
   • **Project Name** — Shows the project name of the published pipeline.
   • **Pipeline Name** — Shows the published or deployed pipeline name.
   • **Pipeline Type** — Shows the type of pipeline. such as development.
   • **Created On** — Shows the time and date when the pipeline is created.
   • **Tags** — Displays the tags used in the pipeline.
   • **Comments** — Displays the comments if any.

Viewing approval status

The Production For Approval allows you to view the approval status of projects and pipelines. Production pipelines allow you to view and edit through the platform's production URL, which is different from the platform URL.

To view the pipelines in production, perform the following steps:

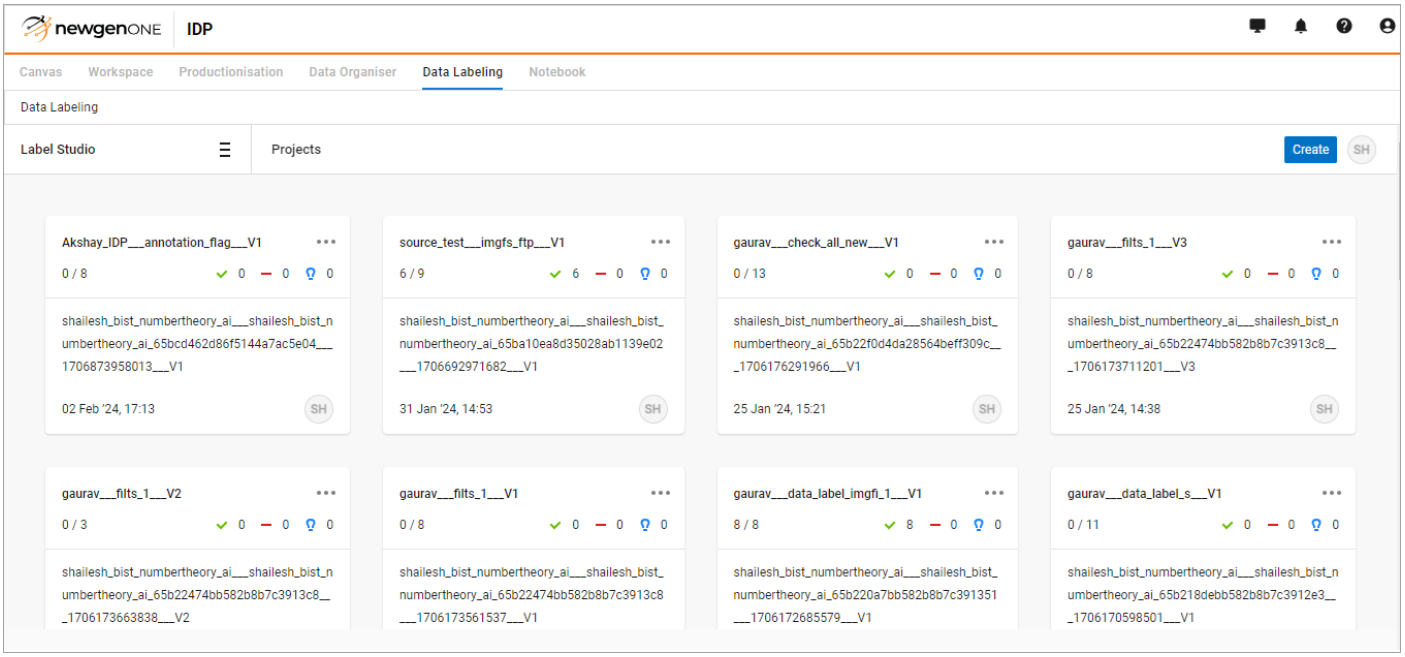
1. Go to the Productionisation tab.
2. Click **Production For Approval**. The Product For Approval page appears.
3. Select the desired project. The pipelines present in the project appear in the Pipelines section.
4. Click the More Options icon ⋮ given next to the pipeline. The following options appear:

Option	Description
View Result	Shows the results of the last successful run of the pipeline.
Open	Opens the pipeline in canvas, but only in view mode. To edit the pipeline, go to the platform and republish the pipeline for production.
Dashboard	Allows you to navigate to the dashboard for the pipeline, where you can see the visual representation of various metrics of the model. You can also view the ROC, LIFT, and GAIN curves of the pipeline model.

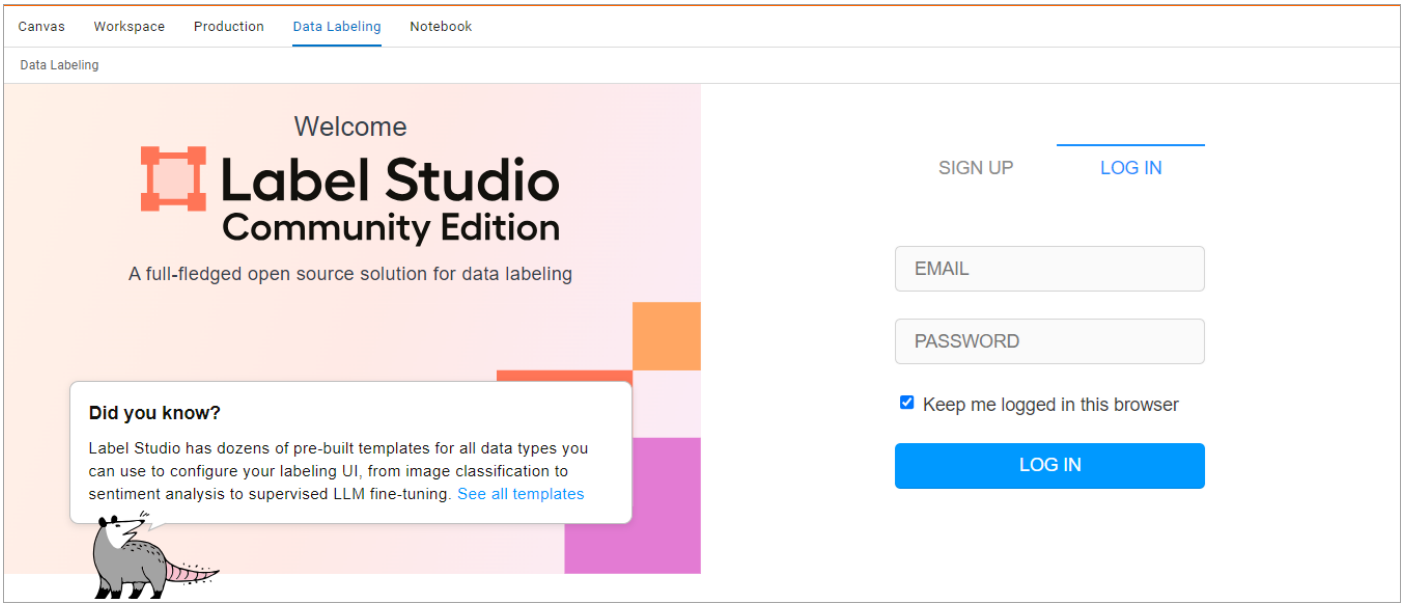
Data labeling

Data Labeling allows you to do labeling in the document images involved in the pipelines. The data labeling projects are available under the Data Labeling tab. These projects are generated when a pipeline is run successfully with the Annotation flag set to *True* in My Workspace.

Although you can create a data labeling project by clicking the *Create* button in the upper right corner of the data labeling page, it's not recommended. This button is merely a feature of the integrated data labeling studio within IDP Studio.
The project name created here has the following naming convention:
Data Labeling Project name = (Project Name in Workspace) _(Pipeline Name)_(Pipeline version).



If you see the below screen asking for credentials to login, then contact your administrator if you do not already have the credentials. Enter the email and password to login.



Creating a data labeling project

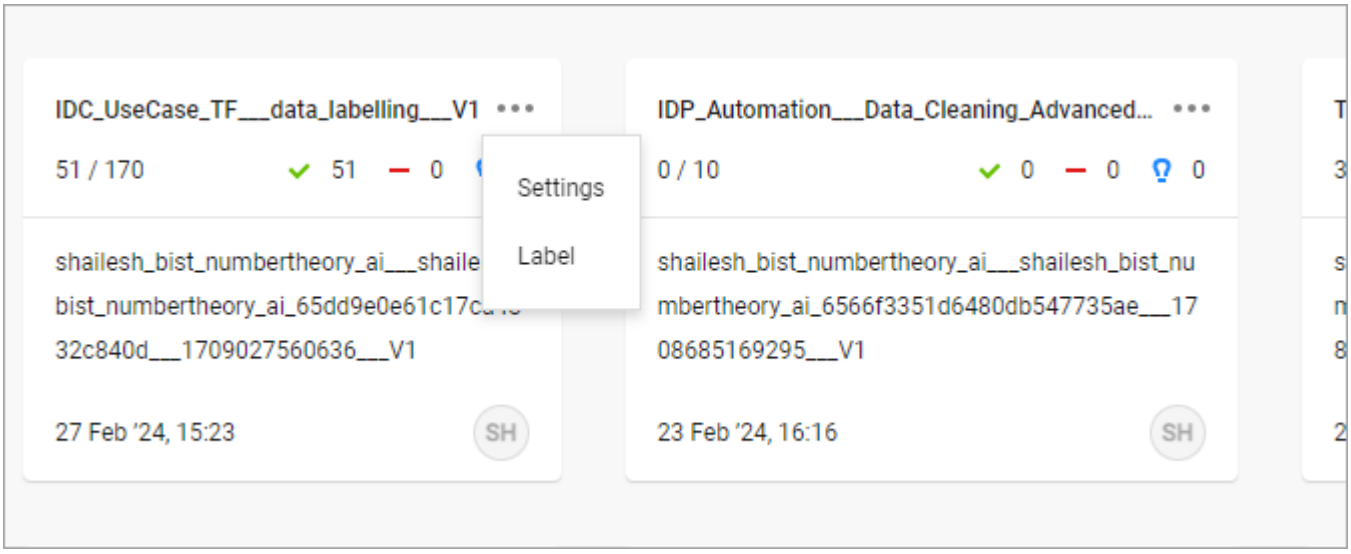
Click a project to open it.

The data labeling project offers various options within the integrated Labeling Studio of IDP Studio. These options are elaborated on in the following subsections.

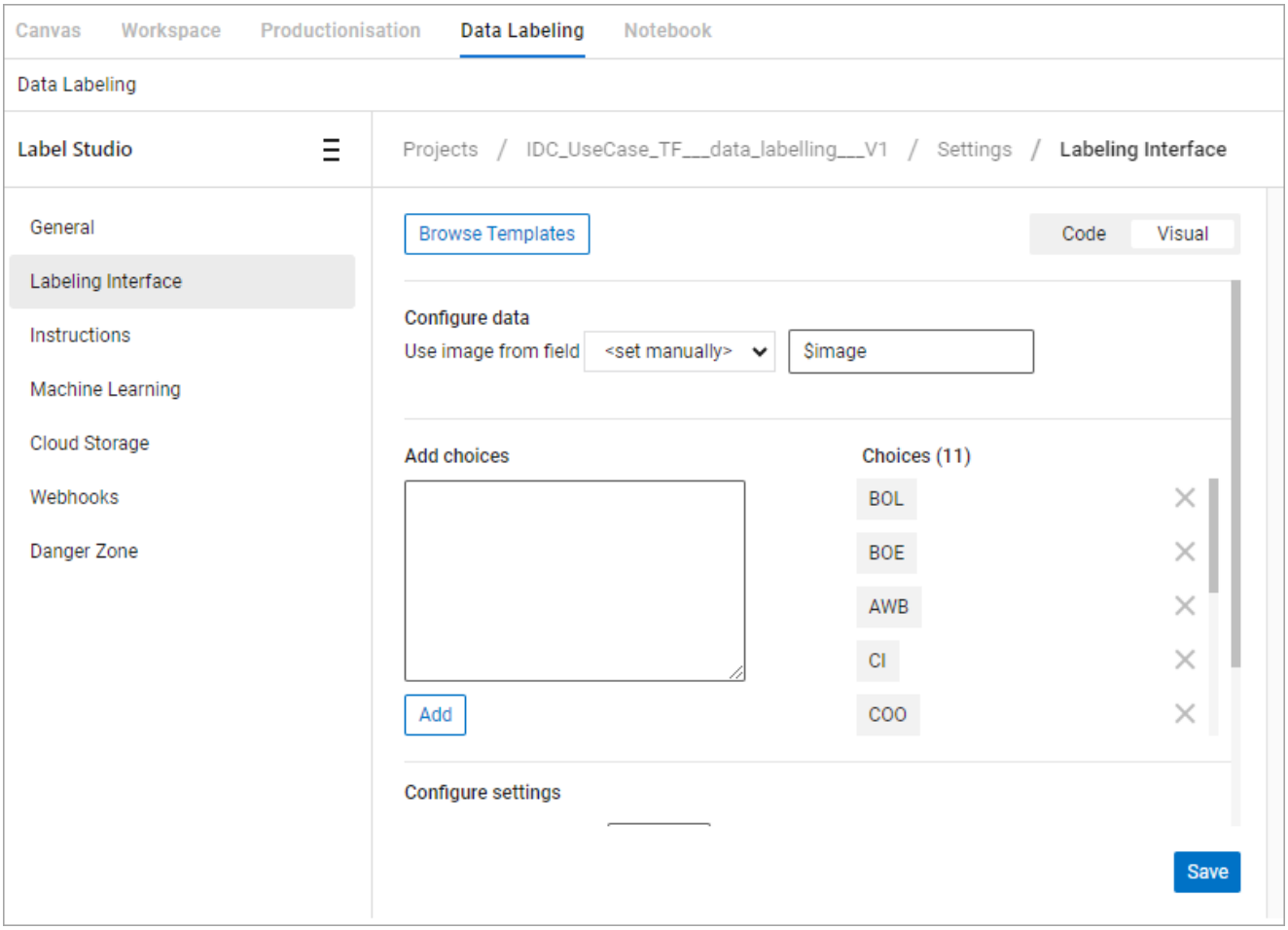
Adding labels to a project

To add new data labels, perform the following steps:

1. Click the More Options icon ••• on the upper right corner of the card of a project.
2. Click the **Settings** option. The screen to define labeling settings appear:



3. The screen to define labeling settings appear:

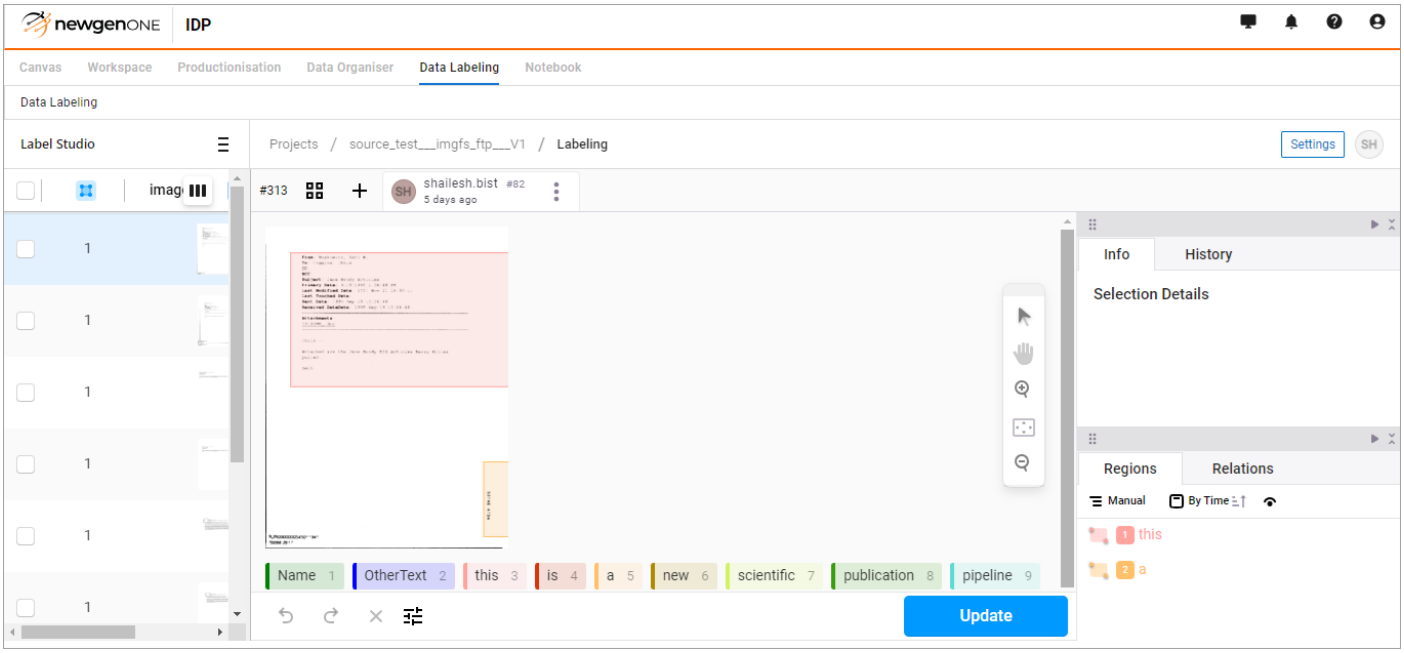


4. Click **Labeling Interface**.
5. Enter the desired label name in the test box under Add choices.
6. Click **save**.

You can remove an existing label by clicking the cross icon ✕.

Adding annotation to a document content

Click an image under the image column. The screen to add annotations to the contents of the image appears:



You can zoom in, zoom out, zoom to fit, and zoom to the actual size of the image using the icon options given on the right side of the image area.

To add an annotation

1. Click the + sign which says Create a new annotation on mouse hover. A new popup appears.
2. Select the annotations given at the bottom.
3. Select the area on the image that you want to annotate.
4. Click the **Save** button.

To update the annotation

1. Select the one from the list on the landing screen of the project under the image column.
2. Click on the annotation from the bottom of the screen.
3. Select the area on the image that you want to annotate.
4. Click the **Update** button.

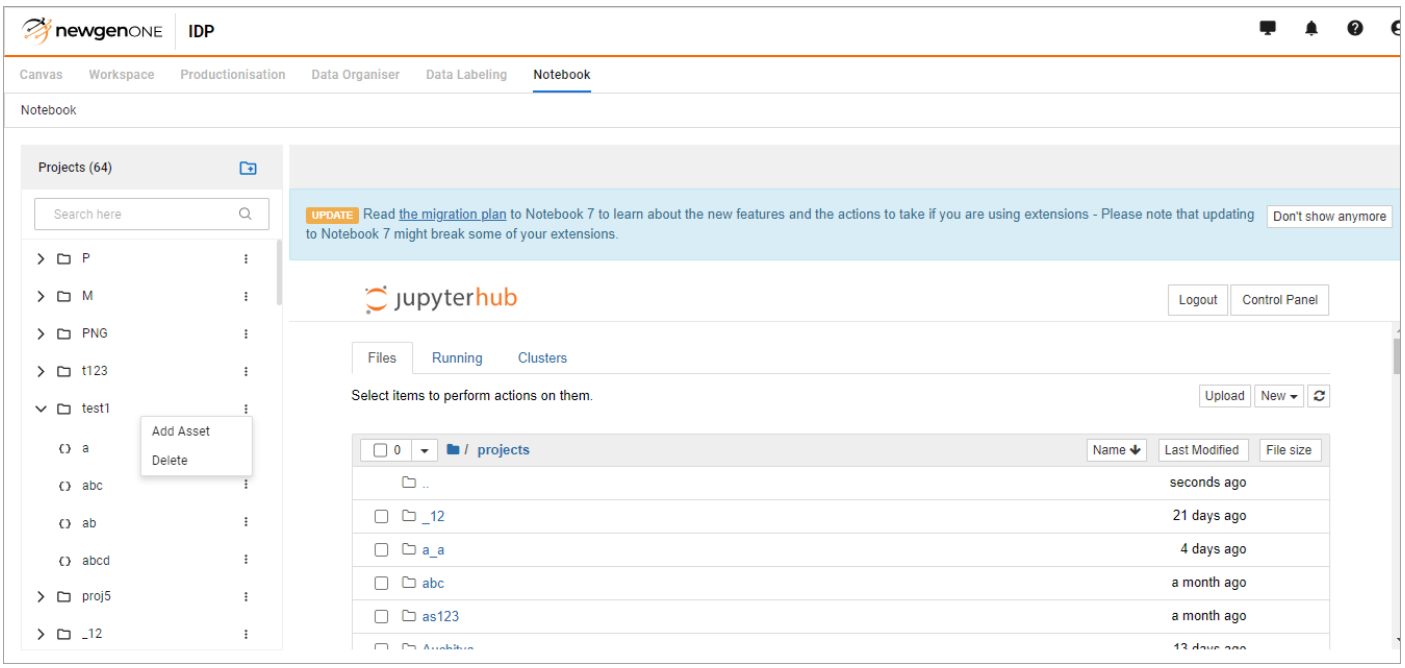
Notebook

The Notebook tab allows you to create or update Jupyter notebooks. You can publish notebooks, for use in a pipeline.

For more procedural details, see Mapping.

Click the **Notebook** tab located in the menu at the top. The Notebook page appears with the following options:

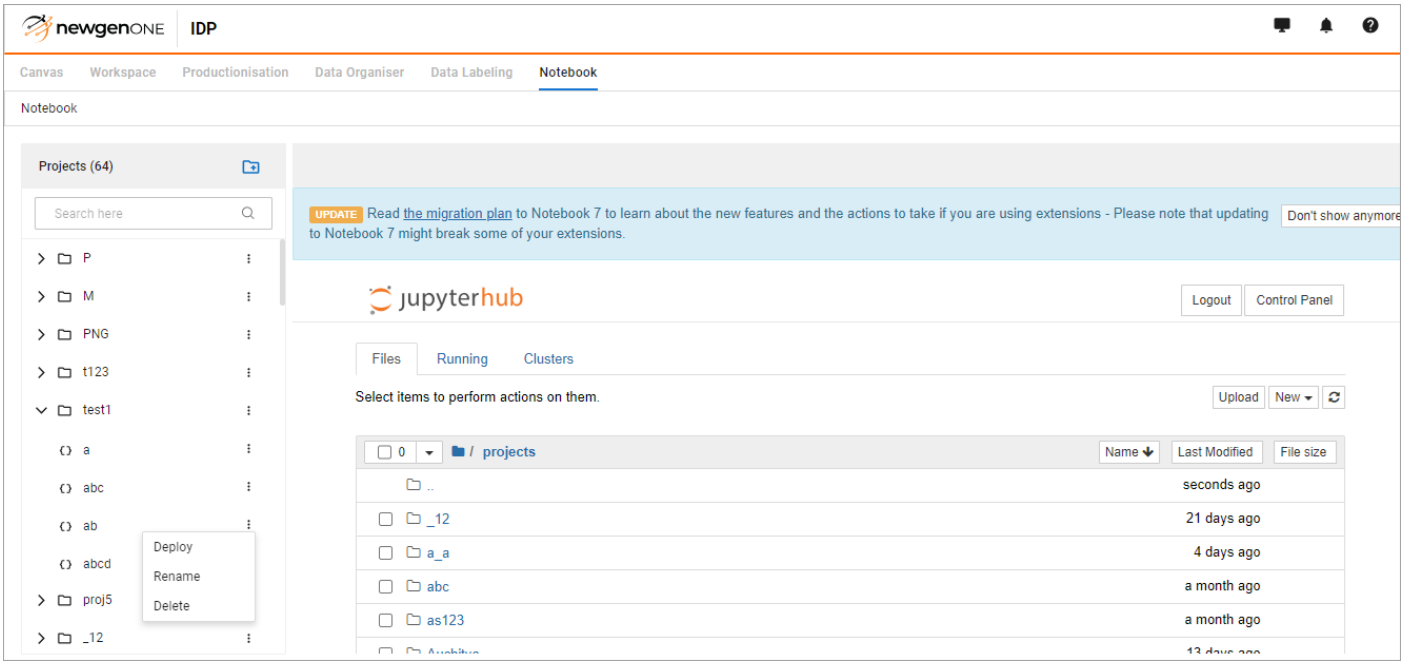
Option	Description
Projects	Shows you the list of available projects.
Create new notebook	Allows you to create a new notebook.
More Options icon	This option against the available project name provides you with the following options: • **Delete** — This allows you to remove the project along with all notebooks in the project. • **Add Asset** — This allows you to add a new notebook to the project.



Clicking a notebook, allows you to view the project section on the right side. You can add or update your code in Python on the notebook as per your requirement and save it.

Click the More Options icon ⋮ against an available pipeline name under a project, this comprises you with the following options:

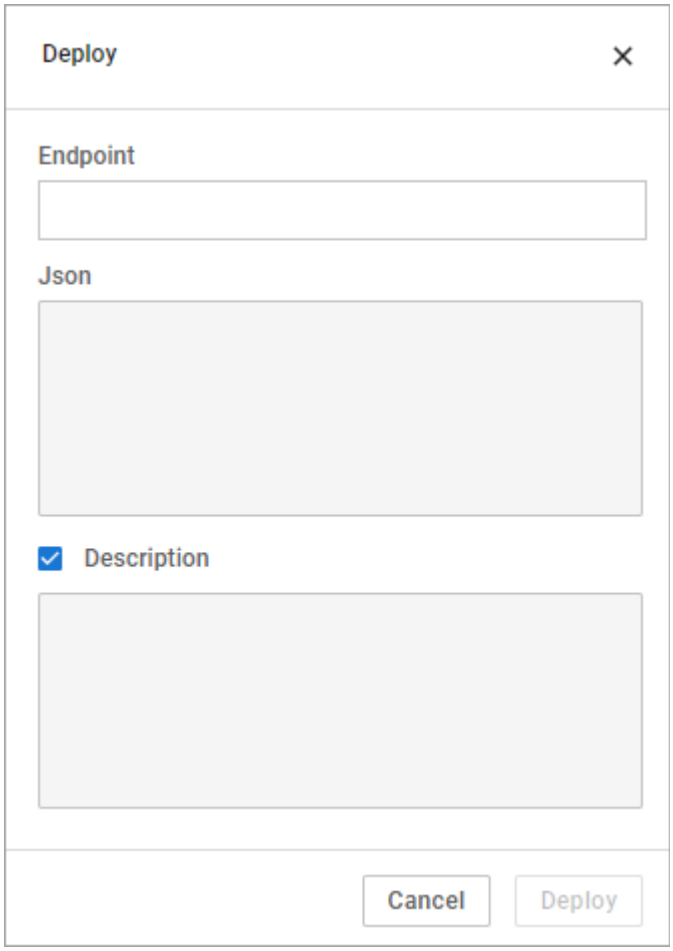
Option	Description
Deploy	Allows you to deploy the notebook to be accessible as an API. For procedural details, see Deploying_a_notebook.
Rename	Allows you to rename the pipeline.
Delete	Allows you to delete the pipeline from the project.



Deploying a notebook

Click **Deploy** shows a dialog with the following options:

- **Endpoint** — Allows you to enter the endpoint name for the API.
- **Input Json**— Allows you to enter the endpoint name for Json (if it is as per your requirement).
- **Description** — Allows you to add a description for your notebook API.

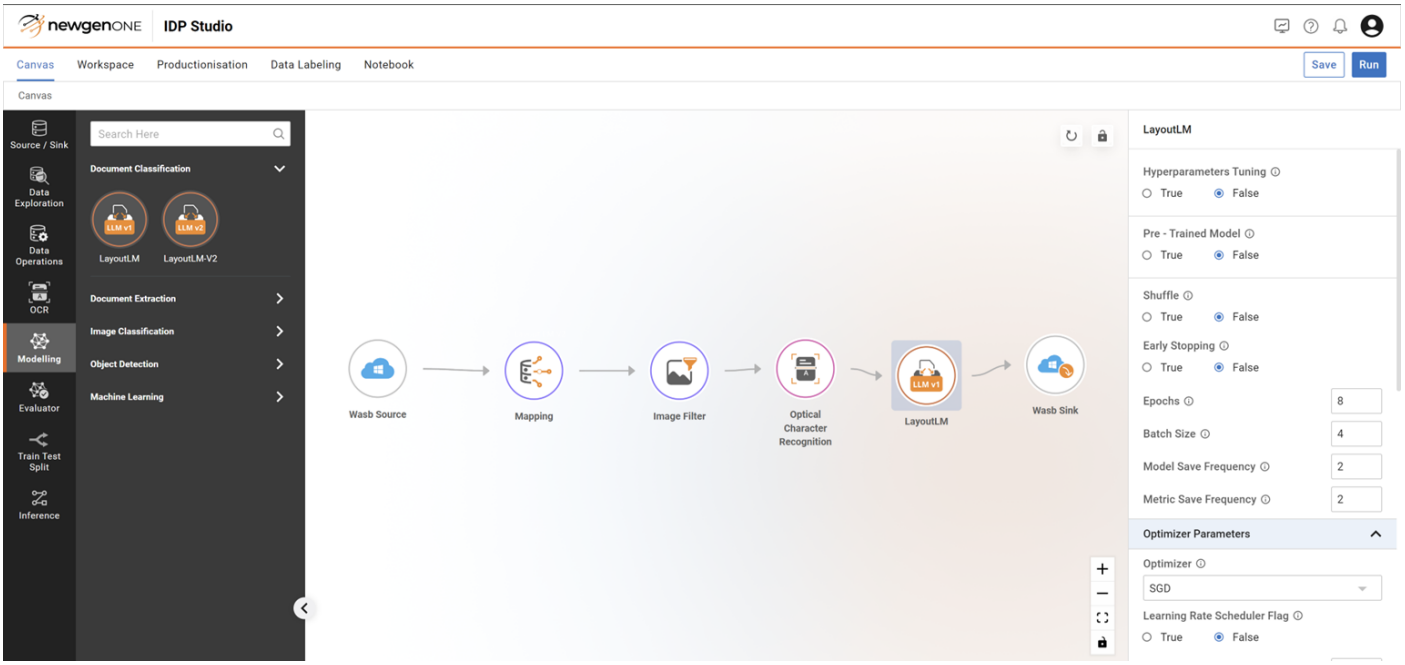


- If you are using the on-prem version of the application, clicking the notebook tab prompts a sign in screen. Enter your username (without @yourdomain) and password into the respective text boxes.
- If you see a message requiring jupyter notebook URL in the notebook area after clicking the Notebook option, then copy that URL hit on a different tab of the browser, and come back to the tab where your application is open. Click on the Notebook option again or refresh the page, you should be able to see your Jupyter notebook space.
- If you are using Microsoft for signing in to the application (for Azure cloud deployment), consider the following points:
    - You are auto-signed in to the jupyterhub with the same credentials you use to sign in to the IDP Studio.
    - If you are signed in to the same browser with any other application (like Outlook) with a different Microsoft sign-in than the one for IDP Studio, then ensure to sign off from that application first.
    - To avoid this situation, use the private or incognito mode of your browser to access the application.
    - Ensure that the browser setting to block the third-party cookies is off.

Creating a pipeline

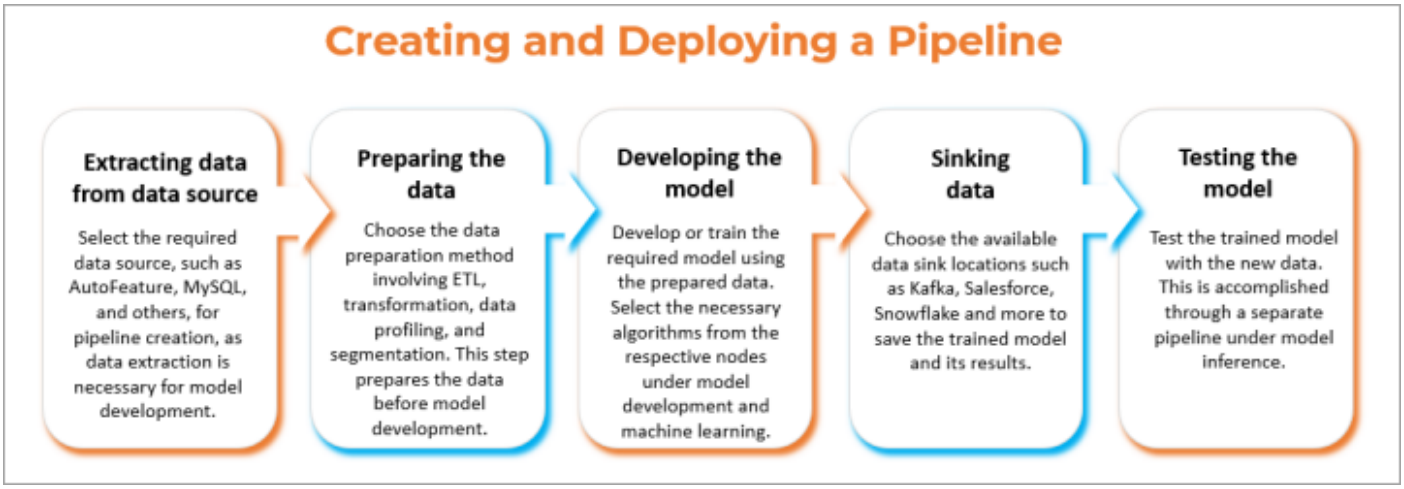
The Canvas enables pipeline creation using pre-built nodes. To add a node, drag and drop it onto the canvas and connect it to the preceding and/or succeeding node. Clicking a node on the canvas displays selection options on the right side of the screen.

While creating a pipeline, you might encounter validation messages. These messages can relate to constraints for adding nodes to the canvas, connecting nodes, or specifying required node types in certain scenarios.



Creating and deploying a pipeline workflow

The following workflow depicts the typical pipeline creation and deployment using the NewgenONE IDP Studio platform



Preparing the data

Data preparation is the first and primary step for model pipeline creation. The data pipeline creation involves the following processes:

- Connecting data
- Exploring data
- Cleaning data

Connecting data

The initial step in creating a pipeline is connecting to the data. First, connect to the data source and then define the data sink (destination) to save the data.

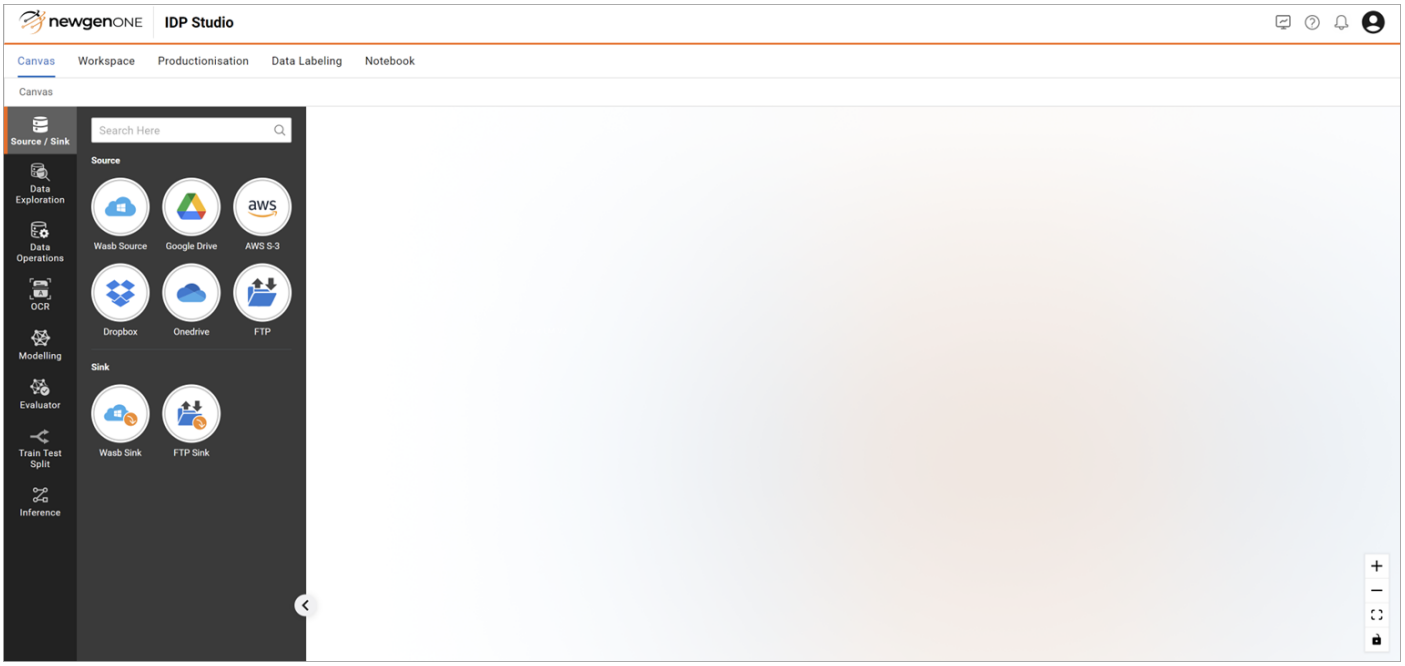
NewgenONE IDP Studio provides a variety of data source and sink connectors (nodes) to choose from and configure. These data source and data sink nodes are available in the node cluster of Sources and Sink on the Canvas under the Database option. The clusters of nodes are:

- Sources
- Sink

The following tables show the availability of the data source and sink nodes in cloud deployment compared to on-premises deployment of the IDP Studio:

Data sources	Cloud deployment	On-Prem deployment
Wasb	Yes	No
Google Drive	Yes	No
AWS S-3	Yes	No
Dropbox	Yes	No
Onedrive	Yes	No
FTP	Yes	Yes

Data Sink	Cloud deployment	On-Prem deployment
Wasb	Yes	No
FTP	Yes	Yes



Reading data source

The Data Source Read cluster offers a range of data connector nodes, enabling you to insert data from various sources into the pipeline.

The application supports images in jpg, png, and bmp formats.

To read the data source, perform the following steps:

1. Go to the Canvas tab.
2. Click Database in the left navigation pane.

The following nodes appear in the Source section:

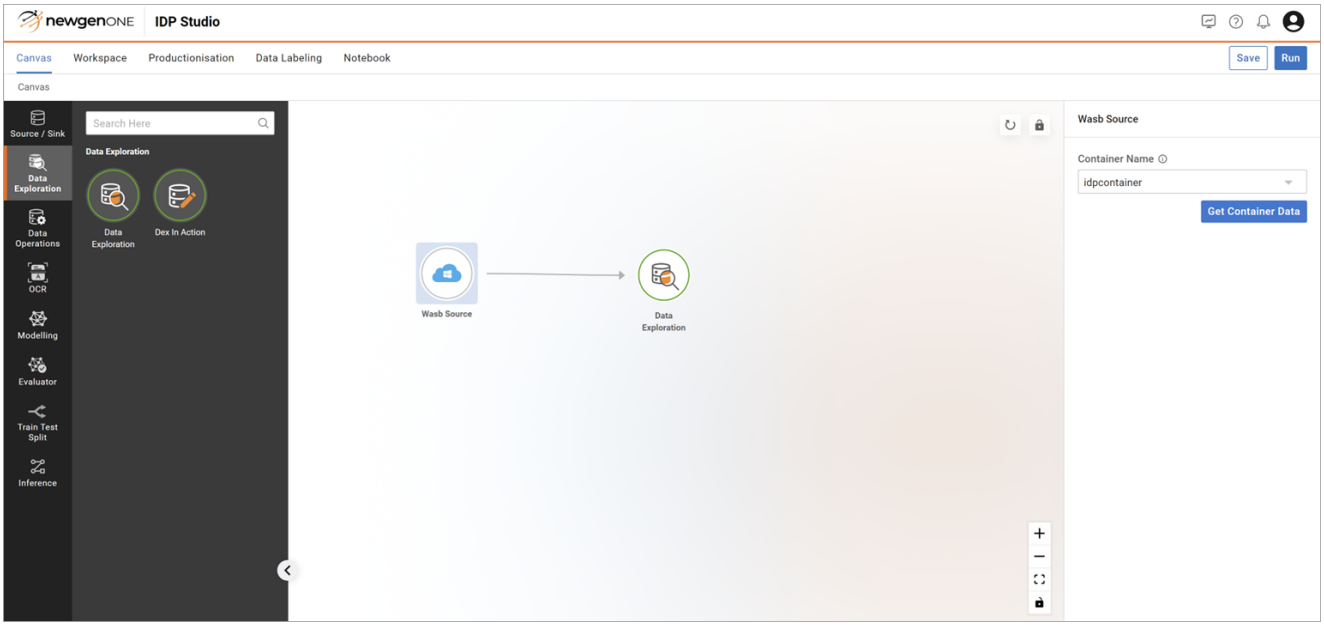
- Wasb Source
- Google Drive
- AWS S-3
- Dropbox
- Onedrive
- FTP

Wasb Source

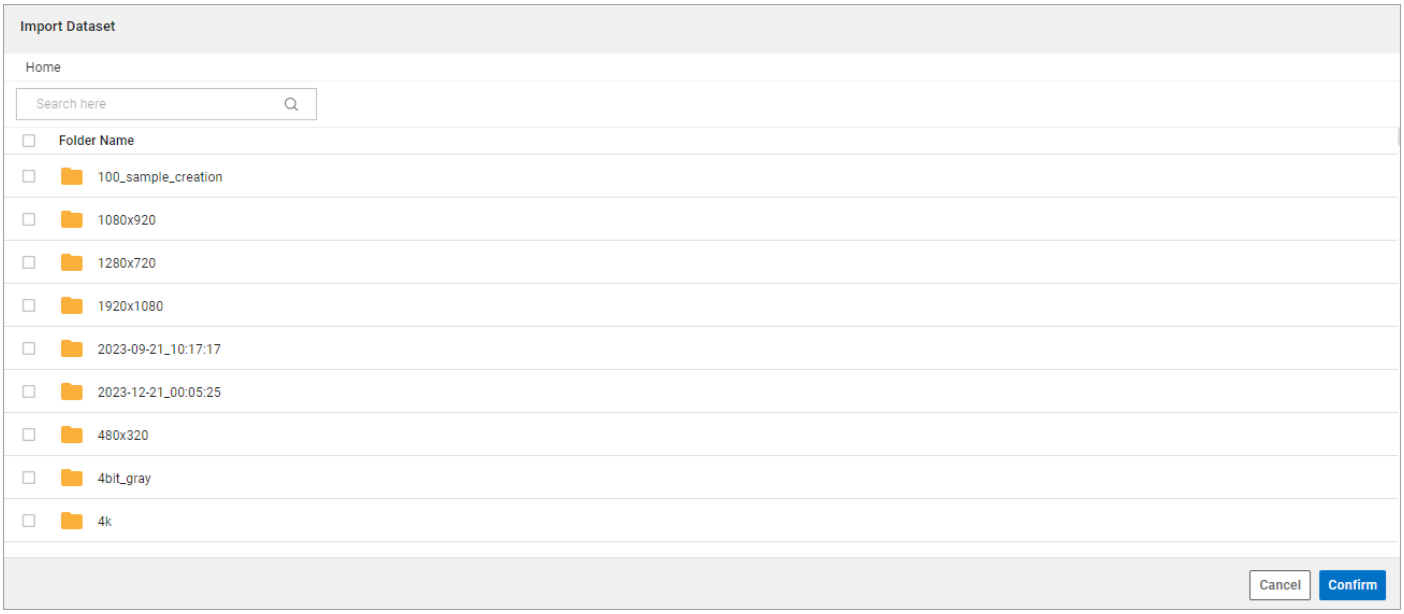
WASB is a file system storage on the Azure cloud, where you can store data irrespective of its structure and schema.

To read the data from the Wasb Source node, perform the following steps:

1. Perform the steps from 1 to 3 mentioned in the Reading data source section.
2. Drag the Wasb Source node and drop it on the Canvas.
3. Connect the Wasb Source node to the succeeding node in your pipeline.



4. Click the Wasb Source node to define its input parameter values. The section appears on the right-hand side of the canvas to provide input parameters.
5. Select the Container Name from the dropdown and click **Get Container Data**. The Import Dataset screen appears.



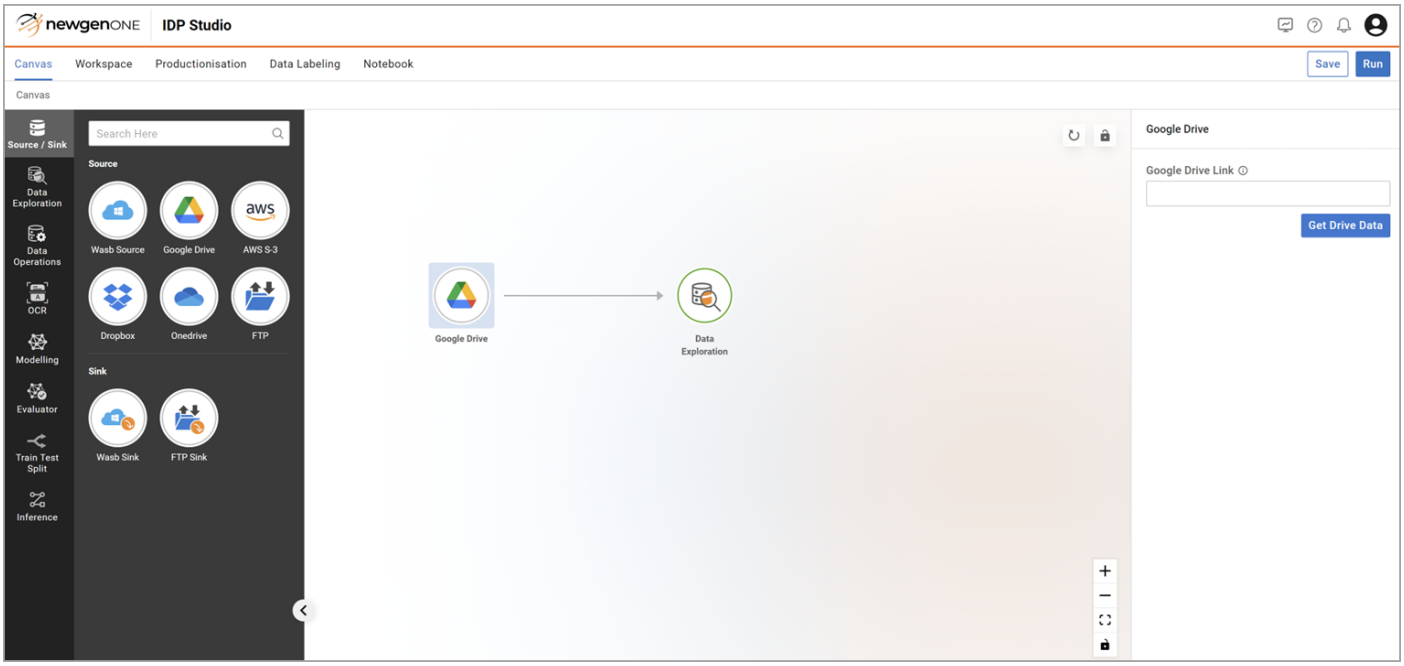
6. Browse through the folders and choose the one you want to use (where your image documents data resides).
7. Click **Confirm**.

Google Drive

Google Drive is a file system storage on Google, where you can store data irrespective of its structure and schema.

To read the data from the Google Drive node, perform the following steps:

1. Perform the steps from 1 to 3 mentioned in the Reading data source section.
2. Drag the Google Drive Source node and drop it on the Canvas.
3. Connect the Google Drive node to the succeeding node in your pipeline.



4. Click the Google Drive node to define its input parameter values. The section appears on the right-hand side of the canvas to provide input parameters.
5. Enter the **Google Drive Link** in the text box and click **Get Drive Data**. The Import Data screen appears.
6. Browse through the folders and choose the one you want to use (where your image documents data resides).
7. Click **Confirm**.

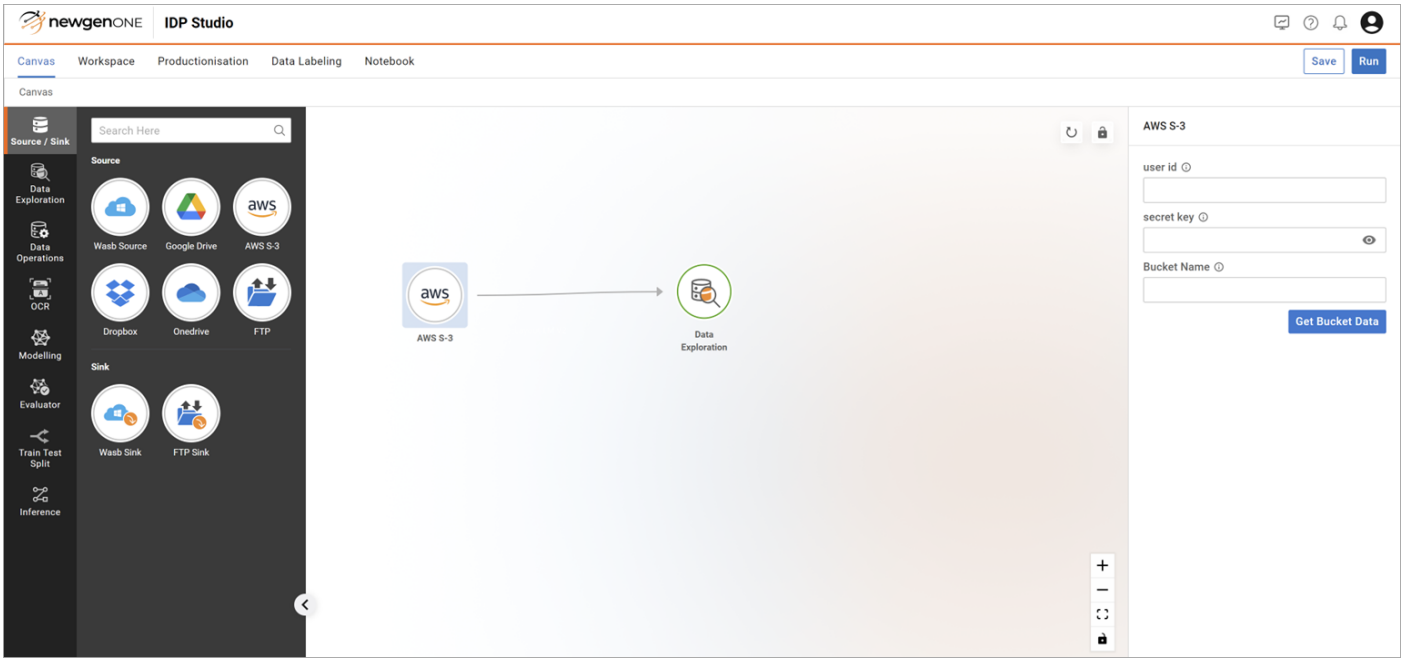
> The Google Source works only with internal Google accounts within the organization where the IDP Studio is deployed.

AWS S-3

AWS S-3 (Amazon S3) is a file system storage on Amazon Web Services, where you can store data irrespective of its structure and schema.

To read the data from the AWS S-3 node, perform the following steps:

1. Perform the steps from 1 to 3 mentioned in the Reading data source section.
2. Drag the AWS S-3 Source node and drop it on the Canvas.
3. Connect the AWS S-3 node to the succeeding node in your pipeline.



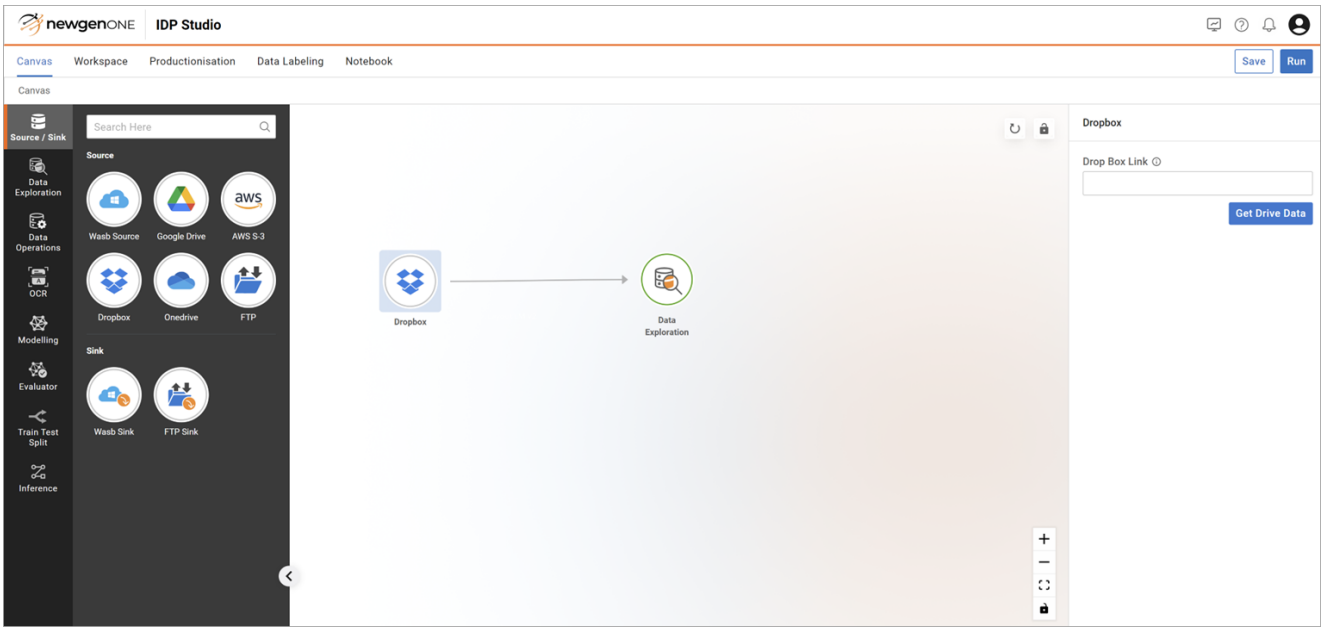
4. Click the AWS S-3 node to define its input parameter values. The section appears on the right-hand side of the canvas to provide input parameters.
5. Enter the **AWS S-3 user ID**, **secret key**, and **Bucket Name** in the respective text boxes and click **Get Bucket Data.**
6. From the Import Data screen navigate to the folder where your image documents are stored. Select that folder.
7. Click **Confirm**.

Dropbox

Dropbox is a file system storage, where you can store data irrespective of its structure and the schema.

To read the data from the Dropbox node, perform the following steps:

1. Perform the steps from 1 to 3 mentioned in the Reading data source section.
2. Drag the Dropbox Source node and drop it on the Canvas.
3. Connect the Dropbox node to the succeeding node in your pipeline.



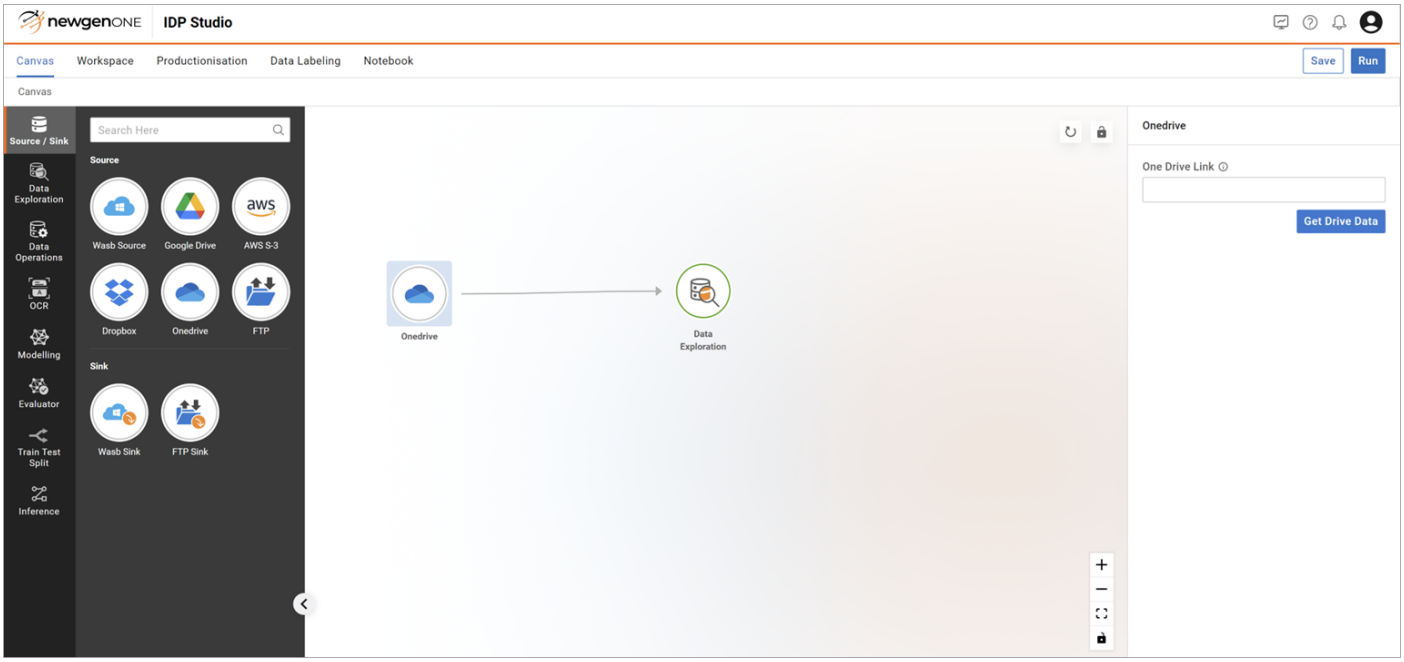
4. Click the Dropbox node to define its input parameter values. The section appears on the right-hand side of the canvas to provide input parameters.
5. Enter the **Dropbox Link** in the text box and click **Get Drive Data**. The Import Data screen appears.
6. Browse through the folders and choose the one you want to use (where your image documents data resides).
7. Click **Confirm**.

Onedrive

Microsoft's One Drive is a file system storage, where you can store data irrespective of its structure and the schema.

To read the data from the One Drive node, perform the following steps:

1. Perform the steps from 1 to 3 mentioned in the Reading data source section.
2. Drag the One Drive Source node and drop it on the Canvas.
3. Connect the One Drive node to the succeeding node in your pipeline.



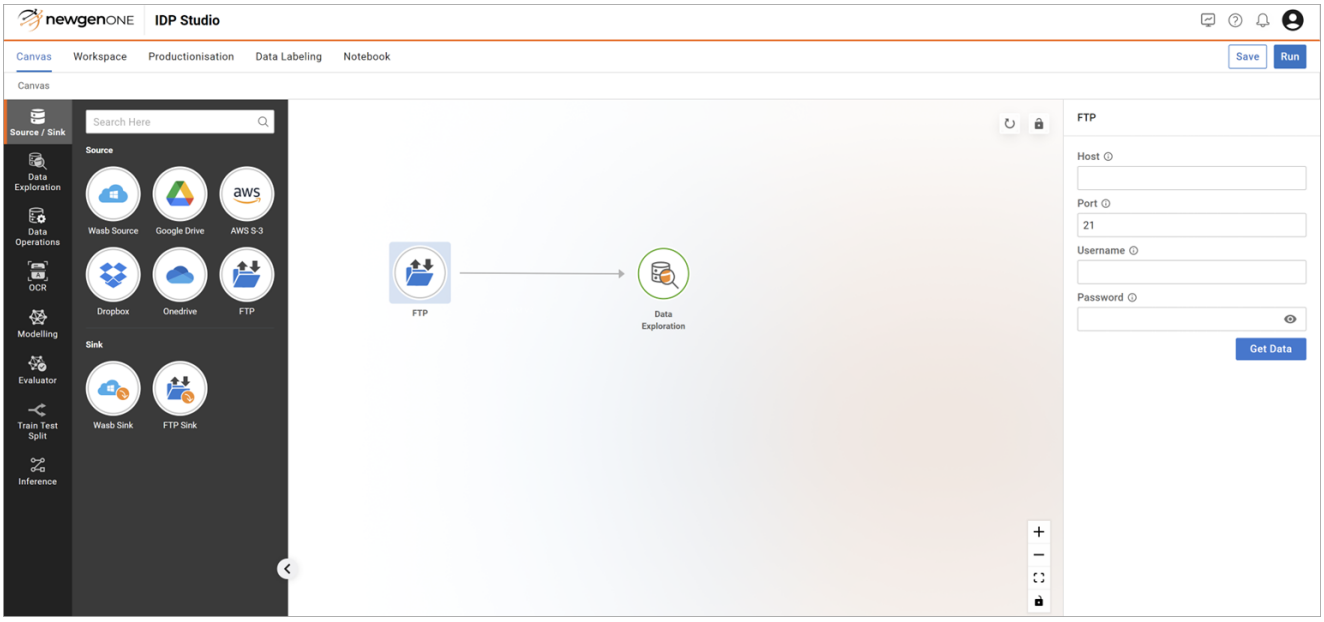
4. Click the One Drive node to define its input parameter values. The section appears on the right-hand side of the canvas to provide input parameters.
5. Enter the **One Drive Link** in the text box and click **Get Drive Data**. The Import Data screen appears.
6. Browse through the folders and choose the one you want to use (where your image documents data resides).
7. Click **Confirm**.

FTP

FTP (File Transfer Protocol) is a network protocol for transmitting files between computers over Transmission Control Protocol or Internet Protocol (TCP/IP) connections. You can store data irrespective of its structure and the schema on a FTP drive location.

To read the data from the FTP node, perform the following steps:

1. Perform the steps from 1 to 3 mentioned in the Reading data source section.
2. Drag the FTP Source node and drop it on Canvas.
3. Connect the FTP node to the succeeding node in your pipeline.



4. Click the FTP node to define its input parameter values. The section appears on the right-hand side of the canvas to provide input parameters.
5. Enter the **FTP Host**, **Port**, **Username**, and **Password** in the respective text boxes and click **Get Data**.
6. From the Import Data screen navigate to the folder where your image documents are stored. Select that folder.
7. Click **Confirm**.

FTP nodes can be connected only to those FTP servers which support FTP over TLS/SSL. Check with your network team regarding the host, port, and credentials required to connect to the FTP server over TLS/SSL.

Sinking data

The Sink section of the Database tab comprises a set of nodes that allow you to save the output data in various sources. To access the supported data sink types, perform the following steps:

1. Go to Canvas.
2. Click Database on the navigation pane on the left side under the canvas menu option.

The following nodes appear in the Sink section:

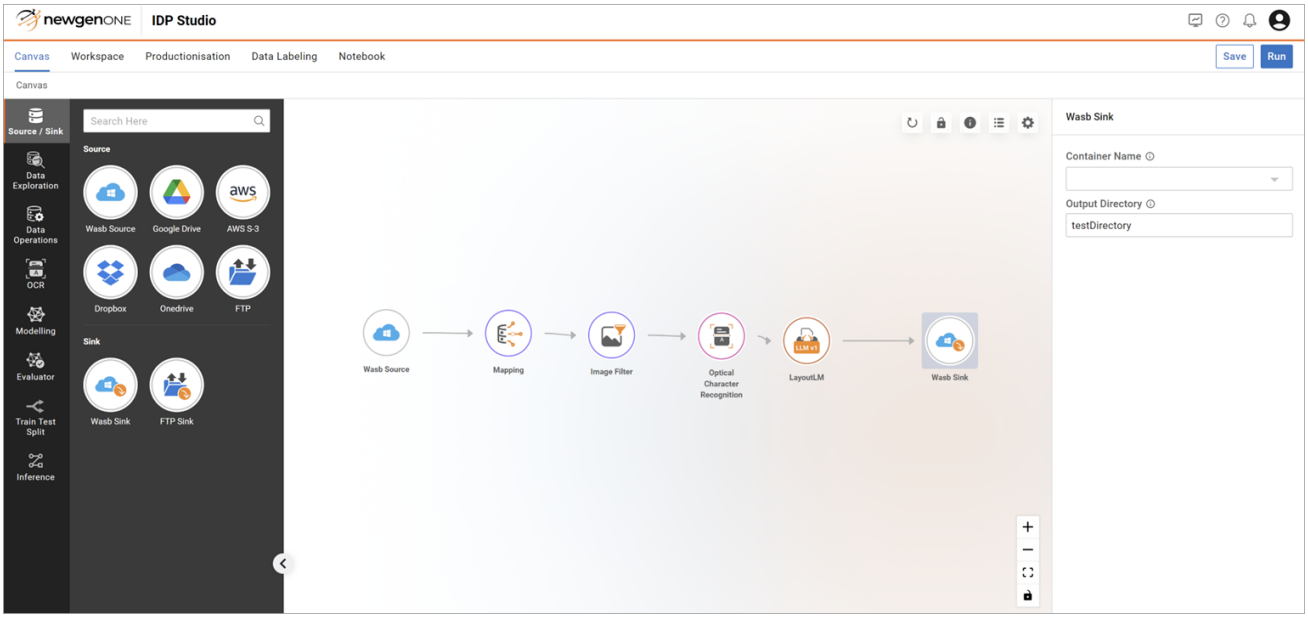
- Wasb Sink
- FTP Sink

Wasb Sink

WASB is a file system storage on the Azure cloud, where you can store data irrespective of its structure and schema.

To save the data on Wasb Sink, perform the following steps:

1. Perform the steps from 1 to 3 mentioned in the Sinking Data section.
2. Drag the Wasb Sink node and drop it on the Canvas.
3. Connect the preceding node to the Wasb Sink node in your pipeline.



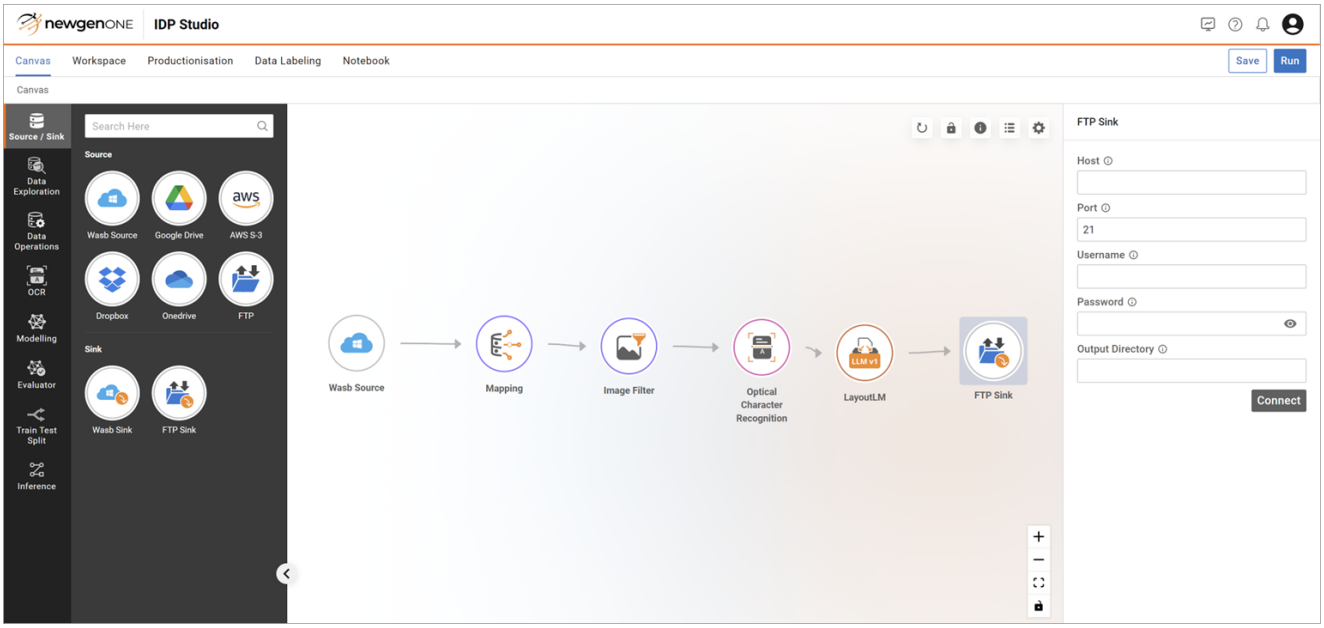
4. Click the **Wasb Sink** node to define its input parameter values. The section appears on the right side of the canvas to provide input parameters.
5. Select the Container Name from the dropdown menu and enter the name of the Output Directory inside the selected container.
6. Save the pipeline.

FTP Sink

FTP sink is a network protocol for transmitting files between computers over Transmission Control Protocol/Internet Protocol (TCP/IP) connections. You can save data irrespective of its structure and the schema on a FTP drive location.

To save the data on FTP Sink, perform the following steps:

1. Perform the steps from 1 to 3 mentioned in the Sinking Data section.
2. Drag the Sink node and drop it on the Canvas.
3. Connect the preceding node to the Sink node in your pipeline.



4. Click the **FTP Sink** node to define its input parameter values. The section appears on the right side of the canvas to provide input parameters.
5. Enter the **FTP Host**, **Port**, **Username**, **Password**, and the **Output Directory** path in the respective text boxes.
6. Save the pipeline.

FTP nodes can be connected only to those FTP servers which support FTP over TLS/SSL. Check with your network team regarding the host, port, and credentials required to connect to the FTP server over TLS/SSL.

Exploring data

Data exploration is one of the critical steps in any machine learning problem. It involves the thorough analysis and understanding of your dataset to gain insights, prepare the data, and make informed decisions about how to proceed with the model-building process.

The Exploration option is accessed from the Canvas menu and contains the following nodes:

- Data Exploration
- Dex in Action

Data Exploration

To perform data exploration, add the Data Exploration node to the data source.

Click the **Data Exploration** node.

You see multiple options on the right side of the screen under the following categories.

- Statistical
- Non-Statistical

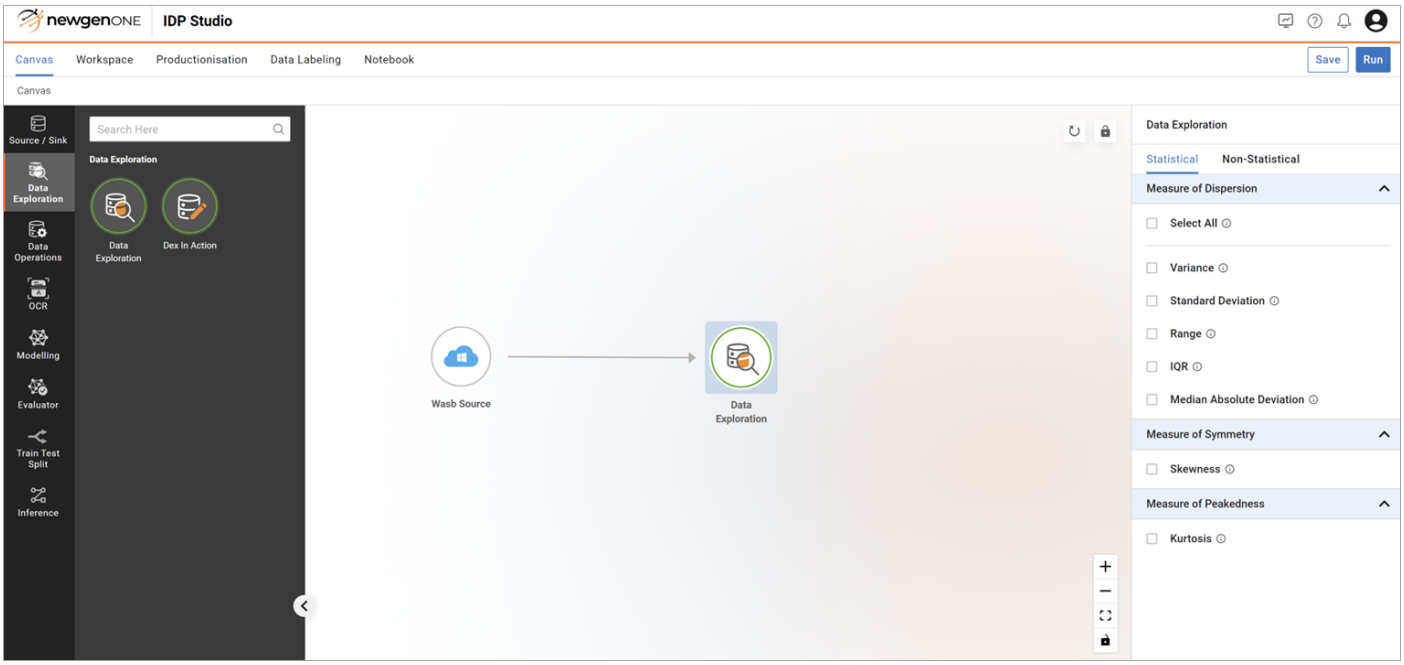
Statistical

The Statistical options are categorized into the following sets:

- **Measure of Dispersion** — It is a statistical concept used in data exploration and analysis that provides valuable insights into the degree of variability or spread of data points from the central tendency. Following are the properties to select under this set:

Properties	Description
Select All	Allows you to apply all settings for Measure of Dispersion in one click.
Variance	Allows you to measure the spread or dispersion of pixel values in an image. It quantifies how much individual pixel values deviate from the mean pixel value. In the context of images, high variance indicates that there are significant variations in brightness or color across the image, while low variance suggests that pixel values are relatively consistent. Variance is commonly used to assess image quality and noise levels.
Standard Deviation	Standard deviation is the square root of the variance and provides a measure of the average deviation of pixel values from the mean. A higher standard deviation implies greater variability in pixel values, while a lower value suggests more uniformity.
Range	The range of pixel values in an image is simply the difference between the maximum and minimum pixel values. It provides a basic measure of the overall spread of values in the image. A wide range suggests a significant difference between the darkest and brightest pixels, while a narrow range indicates a more limited range of pixel values.
IQR	The interquartile range (IQR) is a measure of the spread of pixel values that focuses on the middle 50% of the data. It is calculated as the difference between the third quartile (Q3) and the first quartile (Q1) of the pixel value distribution. IQR is less sensitive to extreme outliers compared to the range and can provide a robust measure of variability in an image.
Median Absolute Deviation	Median absolute deviation (MAD) is calculated as the median of the absolute differences between each pixel value and the median pixel value of the image. MAD is useful for assessing the spread of data when dealing with noisy or non-normally distributed image data and is a robust measure of variability that is less affected by extreme values or outliers in the pixel value distribution.

- **Measure of Symmetry** — The measure of symmetry, also known as skewness, is an important statistical concept used in data exploration in machine learning. It helps in understanding the shape and distribution of data by quantifying asymmetry or lack of symmetry in a dataset. It has the following parameter
    - **Skewness** — The Symmetry class measures image symmetry in Red, Green, and Blue (RGB) format by detecting mirror symmetry. It analyzes pixel patterns across the image, comparing corresponding pixels on opposite sides of a central axis. The skewness of pixel intensity distribution is calculated, and whether it's left-skewed or right-skewed indicates the image's symmetry characteristics.
- **Measure of Peakedness** — The measure under this set provides insight into the shape and peakedness (or flatness) of a probability distribution. It has the following parameter:
    - **Kurtosis** — The Measure of Peakedness calculates kurtosis, a measure of skewness, for the pixel intensities in an image's grayscale representation. It provides three values (one for each color channel), indicating the skewness of the pixel intensity distribution. Higher positive values imply right-skewed distributions, higher negative values imply left-skewed distributions and values closer to zero suggest more symmetrical distributions.



Non-Statistical

The Non-Statistical options are categorized into the following sets:

- **Image Quality Check** — Measuring image quality is an essential step in non-statistical data exploration for machine learning, especially when working with image datasets. While image quality assessment is not a statistical analysis per se, it is a crucial aspect of preprocessing and understanding the data's characteristics. Here are some non-statistical image quality checks for data exploration:

Option	Description
Select All	Allows you to select all options for Image quality checks with a single click.
Quality Metrics	Allows you to select the required metrics from the multiselect dropdown menu. Following are the metrics: • Contrast in an image refers to the difference in intensity or luminance between different parts of the image. • Brightness is a measure of the overall lightness or darkness of an image. It represents the average intensity of all pixels in the image. • Image entropy is a measure of the amount of information or randomness present in an image. Images with high entropy contain a wide range of pixel values and appear more complex, while low-entropy images have more uniform pixel distributions and appear simpler. Image entropy is often used in image analysis and compression. • Brisquare score is a no-reference image quality assessment metric that quantifies the perceived image quality by analyzing statistical features and patterns in the image. It evaluates various aspects of image quality, including sharpness, noise, and contrast. • The blur score measures the degree of blurriness or sharpness in an image.
Image Color Check	Allows you to check if the image is a colored or gray-scale image.

Option	Description
Aspect Ratio Inconsistency Check	Allows you to calculate the aspect ratio of an image and categorize the image based on its aspect ratio as consistent or inconsistent. This function returns the number of images in each category. Documents usually have an aspect ratio of approximately 0.7, so this is set as the default threshold.
Resolution Check	Allows you to calculate the width and height of an input image and returns this resolution information allowing you to assess the image's dimensions. Following are the rules: • < [640, 480]: low • <= [1920, 1080]: medium • rest: high

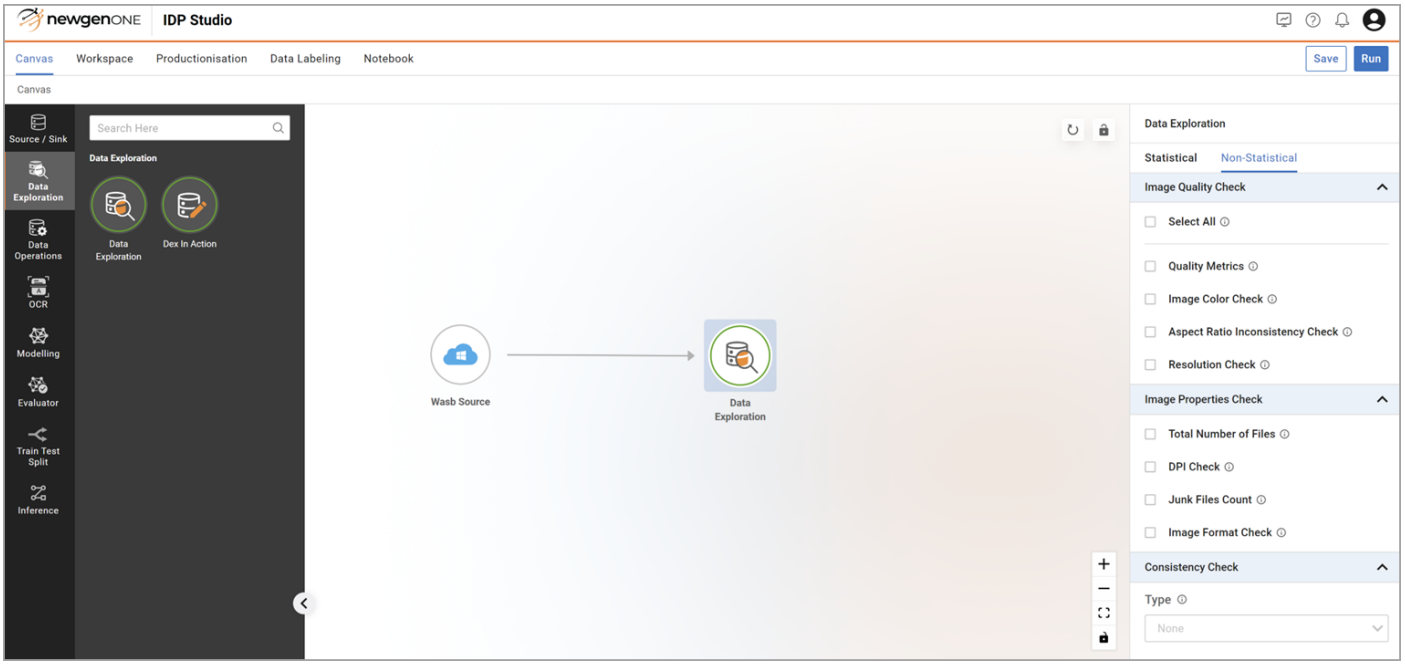
- **Image Properties Check** —This form of exploration involves extracting and analyzing image-specific features and attributes. Here are some non-statistical image properties checks for data exploration:

Option	Description
Total Number of Files	Allows you to count and report the total number of image files (PNG, JPG, JPEG, and BMP) found within a specified folder.
DPI Check	Allows you to check the number of pixels that are present in one inch of an image.
Junk Files Count	Allows you to count the number of junk files present in a folder based on its image size. Files with a size less than 10 bytes and greater than 30MB are considered junk.
Image Format Check	Allows you to provide information about the file format in which the image is stored. Example: PNG or JPEG.

- **Consistency Check** — A measure of consistency check typically refers to assessing the quality and consistency of the data, especially when dealing with non-structured or unstructured data such as text, images, or other types of unformatted information. Following are some non-statistical image consistency checks for data exploration:

- **Type** —It is a multi-selection dropdown with the following options:

Option	Description
Language	Measures the percentage of English text in the image.
Content Area Coverage	Calculates the percentage of text in the document.
Vocabulary Count	Calculates the number of words in the document, including special characters.
Font Height	Calculates the consistency of the font height.
Content Count	Calculates the number of words in the document, excluding special characters and stop words.



Dex in Action

To perform data exploration in action, add the Dex In Action node to the data source.

Click the **Dex In Action** node.

You see multiple options on the right side of the screen under the following categories:

- Statistical
- Non-Statistical

Statistical

The Statistical options are categorized into the following sets:

- **Peakedness** — Kurtosis, or peakedness, measures how much a probability distribution's tails differ from those of a normal distribution. If the image value is not within the threshold, then the image is dropped. Select this option and set the minimum and maximum threshold values if you want to perform dex (data exploration in action) for peakedness.
  - **Min Threshold** — The minimum kurtosis value for the peakedness of the image. It is a numerical value.

    The default value is set to –5.

  - **Max Threshold** — The maximum kurtosis value of the peakedness of the image. It is a numerical value.

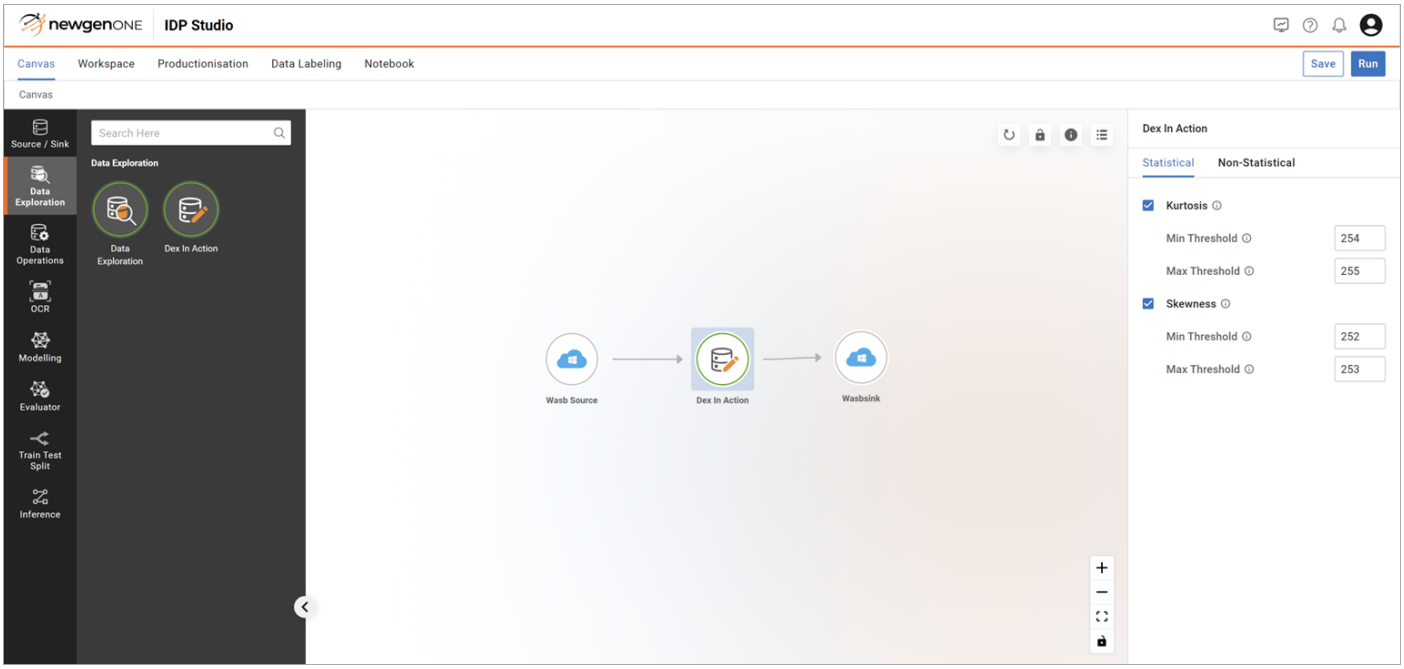
    The default value is set to 5.

- **Symmetry** — Skewness measures the extent to which a probability distribution is stretched or skewed to one side, indicating whether data is more concentrated on the left or right. If the image value is not within the threshold, then the image is dropped. Select this option and set the minimum and maximum threshold values if you want to perform dex (data exploration in action) for Symmetry.
  - **Min Threshold** — The minimum skewness value for the symmetry of the image. It is a numerical value.

The default value is set to –5.

- **Max Threshold** — The maximum skewness value for symmetry of the image. It is a numerical value.

The default value is set to 5.



Non-Statistical

The non-statistical properties that appear on clicking the Filter icon are as follows:

Click the Filter icon for the following properties:

- **Image Quality Metrics** — This option allows you to check the quality of the Image with the following metrics:

Option	Description
Select All	Allows you to select all the Image Quality Metrics in a single click.
Contrast Check	Allows you to remove the images which do not follow the correct criteria of contrast in an image. Set the min and max contrast values. These are numerical values with default values set as 30 and 150.
Brightness Check	Allows you to remove the images which do not follow the correct criteria of brightness in an image. Set the min and maximum brightness values. These are numerical values with default values set as 30 and 150.
Entropy Check	Allows you to remove the images that do not follow the correct criteria of entropy in an image. Set the min and max entropy values. These are the numerical values with default values set to 1 and 7.
Resolution Check	Allows you to drop an image if the resolution of the image is out of the given range. Selecting this option needs to set the min X and min Y allowed coordinates. These are the numeric fields with default values set to 50 for each.
Brisque Score Check	Allows you to remove the image if the value is higher than the threshold. Selecting this option requires you to set the max Brisque Threshold value. This is a numerical value with the default value set to 50.
Blur Score Check	Allows you to remove the image if the value is higher than the threshold. Selecting this option requires you to set the Max Blur Threshold. This is a numerical value with a default value set to 0.5.

- **Image Properties Check** — Allows you to check the image properties by following options:
    - **Image Format Check** — Allows you to select from various image formats and preserve the image according to your selection. To do this, simply pick one or more image formats from the dropdown menu. The possible values are PNG, JPEG/JPG, and BMP.
    - **DPI Image Check** — Allows you to drop an image based on its DPI (drop per inch). To use this feature, specify the minimum DPI value allowed for the images you want to keep.

      This is a numeric field with the default value set to 100.

    - **Junk File Count** — Allows you to drop the junk files based on image size.
- **Consistency Check** — This option allows you to check the documents or image consistency with the following options:

Option	Description
Language Consistency Check	Allows you to drop documents or images that are not consistent in terms of language consistency score for each document. You must set the minimum acceptable consistency score for the image when you enable this option.. This is a numeric field with a default value set to 0.7.
Content Area Coverage Check	Allows you to drop documents or images that are not consistent in terms of content-area-wise-coverage. You must set the minimum acceptable content coverage score for the image when you enable this option.. This is a numeric field with the default value set to 0.5.
Vocab Count Check	Allows you to drop documents that fail to meet the vocab-count threshold, including stop words and extra characters. You must set the minimum acceptable word count when you enable this option. This is a numeric field with a default value set to 50.
Font Height Check	Allows you to drop documents or images that do not satisfy the font-height-consistency criteria. You must set the minimum font height threshold when you enable this option. This is a numeric field with a default value set to 0.7.
Content Count Check	Allows you to drop documents or images that do not satisfy the content-count-score criteria, excluding stop words and extra characters from the documents. You must set the minimum content count score when you enable this option. This is a numeric field with a default value set to 100.

Click the Transform icon for the following properties.

- **Image Quality Transformation** — Allows you to set the following image quality transformation properties:

Option	Description
De-Blur	Allows you to unblur and restore the image. Set the maximum blur value that transforms the image when exceeded. Clicking this option requires you to select the Algorithm Name from the dropdown menu. The possible values in the dropdown are as follows: • Wiener Filter • Richardson Lucy Deconv • Blind De Conv • TV Regularization • Deep De Blur • This is a numeric field with a default value set to 0.5. • The default selection in the Algorithm Name is Wiener Filter.
Adjust Contrast	Allows you to adjust the contrast factor of the image. Setting the minimum contrast value below which the image transforms is necessary. Additionally, setting the Contrast Factor adjusts the contrast factor of the image. Increasing the contrast value makes the text bolder. The default Min Contrast value is set to 50. The default Contrast Factor value is set to 1.
Adjust Brightness	Allows you to adjust the brightness level of the image. Adjust the minimum brightness value that triggers the image transformation. It also requires you to set the Brightness Factor to adjust image brightness. The default Min Brightness value is set to 50. The default Brightness Factor value is set to 1.

Option	Description
Super Resolution	Allows you to improve the resolution of the image using deep learning models. It requires the following parameters to set: • **Min X Coordinate** — Set the minimum width for image transformation. It is a numeric field with a default value set to 100. • **Min Y Coordinate** — Set the minimum height for image transformation. It is a numeric field with a default value set to 100. • **Algorithm Type** — Choose between Basis Interpolation and Basis Deep Learning algorithms. The default selected value is Basis Interpolation. • **Algorithm Name** — Select the deep learning model for upscaling the image. EDSR (Enhanced Deep Residual Networks) and FSRCNN (Fast Super-Resolution Convolutional Neural Network) algorithms. This parameter only appears when you select Basis Deep Learning as the algorithm type. **Resolution Scale** — Set the image resolution scale. This parameter only appears when you select Basis Deep Learning as the algorithm type. This is a numeric field with the default value set to 2.
Convert To Black & White	Allows you to transform the input image into Black and White pixels.

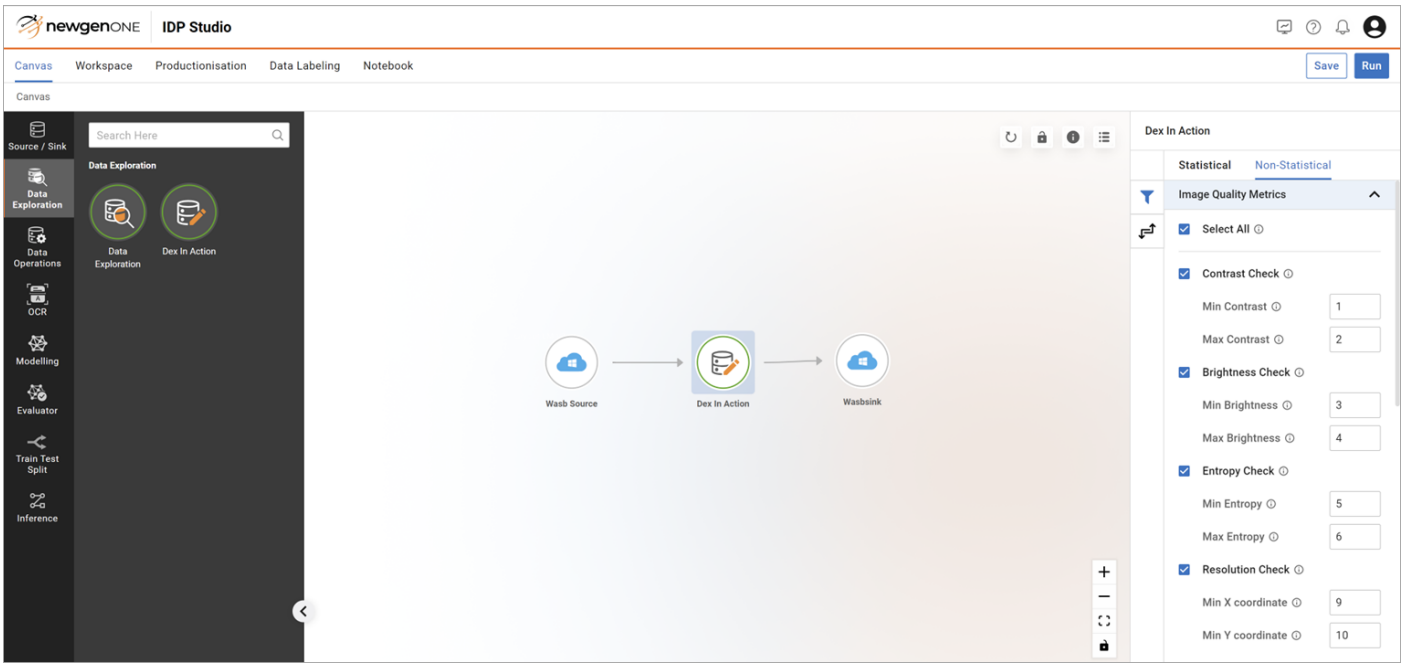
- **Image Properties Transformation** — This option allows you to transform the image properties with the following options:
  - **Change Image Format** — Allows you to change the image format of the image as per the requirement. Select the appropriate image type (PNG, JPEG/JPG, or BMP).

- **Set DPI** — Allows you to set the DPI index based on the following parameters::
  - **Min DPI Threshold** — Set the minimum DPI value that sets the image DPI when it falls below.

    It is a numeric field with a default value set to 120.

  - **DPI Value** — Set the DPI of the image in this field.

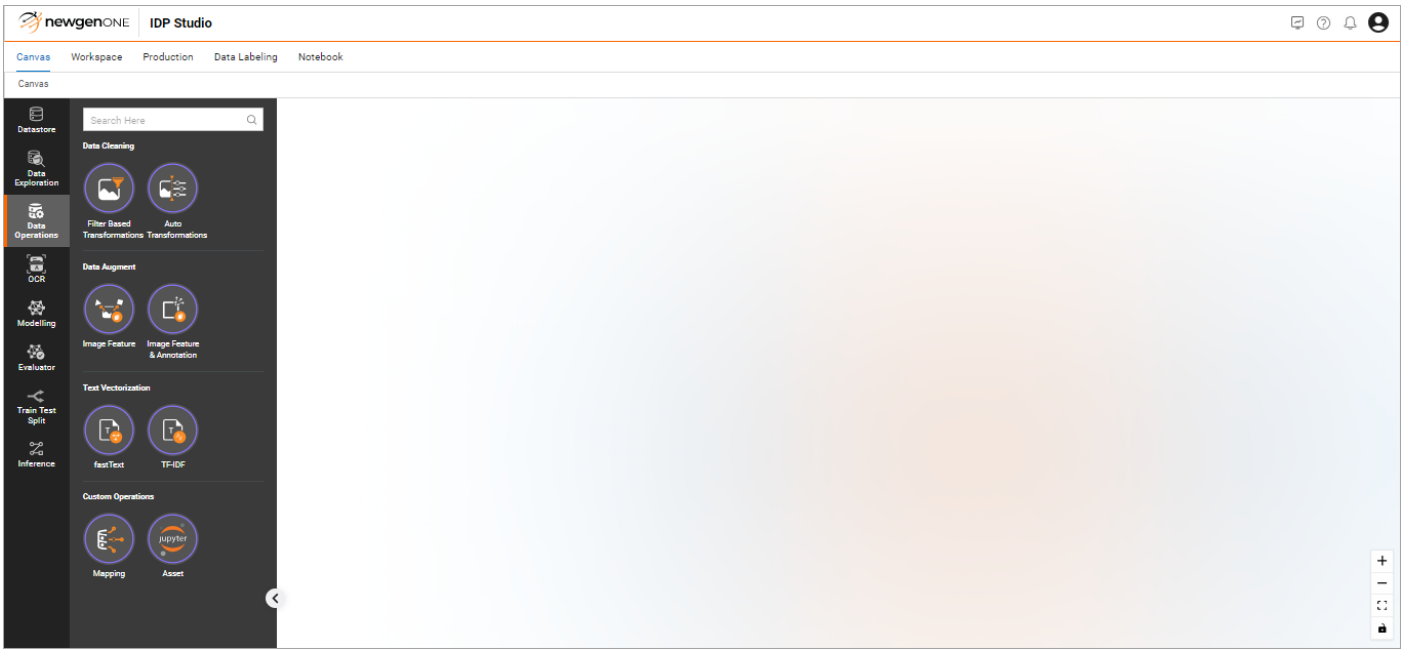
    It is a numeric field with a default value set to 20.



Data operations

This option is available under the canvas tab and has the following node clusters:

- Data Cleaning
- Data Augmentation
- Text Vectorization
- Custom Operations



Data Cleaning

Data cleaning is a fundamental and essential step in model development to ensure data accuracy and consistency.

Following are the data cleaning nodes available in IDP Studio:

- Filter Based Transformations
- Auto Transformation

Filter based transformations

The following table provides the options to select along with the available parameters to set in each option:

Option	Description	Parameters
Despeckle Image	Despeckling an image involves removing speckle noise, a granular noise similar to salt-and-pepper noise in images..	Pixels greater than this Speckle threshold are considered as noise and are removed from the image. It is a numeric field. The default value is set to 25.
Remove Noise Ex	This option allows you to remove the noise (collection of sparse irregular non-geometric silhouettes type) in digitized document images.	Erosion Dimension removes the noise pixels from every dimension of the image. It is a numeric field. The default value is set to 1.
		Median Filter is a boolean value that allows you to apply the median filter to the image. The default value is set to False.
		Distance To Remove is noise components beyond a specified distance from neighboring components in inches are removed. It is a numeric field. The default value is set to 0.075.
Median Filter	This option allows you to remove noise from the image using a non-linear digital filtering technique.	Filter Size set the size to be applied to the image. This is a numeric field. The default value is set to 3.
		Black & White Conversion is a dropdown menu with the following options: • **0** — Converts some filtered white pixels into black and vice-versa. • **1** — Converts certain filtered black pixels to white, unaffected if already white. • **2** — Converts specific filtered white pixels to black; unaffected if already black.

Option	Description	Parameters
		The default value is set to 0.
Adaptive Background Clean	The function is used to clean the background of the input color or gray document image using a local window moving approach.	Window Size is the size of square window in pixels for background removal. It is a numeric field. The default value is set to 3.
		Foreground Enhancement Flag is used to control the color enhancement of the foreground. The default value is set to False.

Auto Transformation

The following table provides the options to select along with the available parameters to set in each option:

Options	Description	Parameters
Auto Correct Orientation	It allows automatic correction of the image orientation.	The Mode defines the decision flag to tell the system in which mode operation is going to be performed. It is a dropdown menu with the following possible values: • **1** — All four-side modes (0,90,180,270 degree orientation) • **2** — For only 0 and 180 degrees of orientation. The default value is set to 1
Auto Deskew	It detects and corrects skewness for any document image within a fixed range with varying levels of accuracy.	Accuracy Level is a dropdown menu that provides four levels of accuracy to select. • The coarse level gives the least accurate skew detection with a maximum deviation of a ±1degree from the actual skew. • Fine and Super Fine, both have higher accuracy levels with maximum deviation of ±0.5 and ±0.1degree respectively. • The Fast & Fine option is a faster approach with a maximum deviation of ±0.5 degrees from the actual skew. The default value is set to Fine.

Options	Description	Parameters
Deduplication	It detects duplicate images in your dataset using the concept of hashing and gives an output folder according to your requirements.	Hash Methods is a dropdown menu that provides four hash methods to select. The following are the possible options: • **CNN** — Convolutional Neural Network (CNN) to extracts feature representations from images, which can then be compared to identify duplicates or similar images. • **WHash** — WHash (Wavelet Hashing) provides functionality for generating and comparing image hashes using wavelet transformations. • **AHash** — AHash (Average Hashing) provides functionality for generating and comparing image hashes using average pixel values. • **PHash** — PHash (Perceptual Hashing) generates and compares perceptual hashes of images. Perceptual hashing focuses on capturing human-perceived similarities between images, making it robust to minor changes. • **DHash** — DHash (Difference Hashing) generates and compares hashes of images by creating hash values by computing the differences between adjacent pixel values in an image.

		• **Threshold Value** — Represents the Hamming distance between two images. Duplicates falling below this threshold are considered valid. It is a numeric field.. **Output Option** — Specifies the desired result format. The available options include Save Unique and Bucket Duplicates.. • The default value is set to CNN. • The default value of the Threshold value is set to 0. • The default value of Output Option is set to Save Unique images.

Augmenting data

Data Augmentation generates realistic synthetic data that preserves the statistical properties and patterns of the original data.

Following are the nodes available in the Data Augmentation cluster of nodes:

- Image Feature
- Image Feature & Annotation

Each of the above nodes has a field called Augmentation Per Image to define the number of augmented images to be generated per input image.

This is a numeric field with the default value set to 5.

Image feature

The options for image classification are given under the following sets:

- Geometric Transforms

- Photometric Transforms
- Transforms
- Image Augmentation

Geometric Transforms

The following table provides the options to select along with the available parameters to set in each option:

Options	Description	Parameters
Padding	Adds extra pixels around the border of the input image.	It has four parameters to set the top, bottom, left, and right paddings. The default value for each of these four parameters is set to 10.
Center Crop	Crops the center portion of the images to the target size.	It has two parameters to set the height and width for the desired (target) image size. The default value is set to 50 for each of these two fields.
Five Crop	Crops the images at the four corners and the center.	It has two parameters to set the height and width for the desired (target) image size. The default value is set to 50 for each of these two fields.
Crop	Crops an image based on the set parameters.	• It has two parameters for the desired height and width of the crop. • It also has four parameters for the top, left, bottom, and right padding of the border of the image. • The default value for the height and width of the crop is set to 128 for each. • The default value for the top, left, bottom, and right padding of the border is set to 10 for each.

Options	Description	Parameters
Resize Crop	Randomly Crops the image in random area and aspect ratio which is resized and scaled to the desired output.	The Resize Crop option includes these parameters for determining the output image: • **Height** — To adjust the height of the image. • **Width** — To adjust the width of the image. • **Lower Bound Scale** — To adjust the lower area of the image. • **Upper Bound Scale** — To adjust the upper area of the image. • **Lower Bound Ratio** — To adjust the lower aspect ratio of the image. • **Upper Bound Ration** — To adjust the upper aspect ratio of the image. • The default value for height and width is set to 32 for each. • The default value for Lower Bound Scale is set to 0.08. • The default value for Upper Bound Scale is set to 1.0. • The default value for Lower Bound Ratio is set to 0.75. • The default value for Upper Bound Ratio is set to 1.3.
Elastic Transform	Elastic transformations are a way to distort an image using random movements, making it look wavy or elastic by generating displacement vectors for all pixels based on random offsets.	The Elastic Transform option includes these parameters for determining the output image: • **Horizontal Magnitude** — To adjust the horizontal magnitude displacement. • **Vertical Magnitude** — To adjust the vertical magnitude displacement. • **Horizontal Smoothness** — To adjust the horizontal smoothness displacement. • **Vertical Smoothness** — To adjust the horizontal smoothness displacement. • The default value of horizontal, and vertical magnitude is set to 50.0 for each. • The default value of horizontal and vertical smoothness is set to 5.0 for each.

Options	Description	Parameters
Resize	Resizes the image to the desired pixel values.	It has three parameters determining the output image: • **Height** — To adjust the height of the image. • **Widht** — To adjust the width of the image. • **Rotation** — To rotate the image randomly within the given range. The default value is set to 360 each.
Rotation	Rotates an image randomly within the given range.	It has the following parameters to determine the output image rotation: • **Min Angle** — To rotate the image for minimum angle. • **Max Angle** — To rotate the image for maximum angle. • **X- Center** — Abscissa of rotation co-ordinates. • **Y - Center** — Ordinate of rotation co-ordinates. • **Expand** — Whether to expands the output to make it large enough to hold the entire rotated image. Expand flag assumes rotation around the center. • The default minimum and maximum values are set to 0 and 180, respectively. • The default value for X-Center and Y-Center is set to 0 for each. • The default value of the Boolean field is set to True.

Photometric Transforms

The following table provides the options to select along with the available parameters to set in each option:

Options	Description	Parameters
Color Jitter	Color Jitter is a data augmentation technique that varies the brightness, contrast, hue, and saturation of the sample images.	It has eight parameters to set for min and max values of brightness, contrast, hue, and saturation factors of the target image. These are the numeric fields. The default value for all minimum values is set to 0. The default value for all maximum values is set to 0.5.
Gaussian Blur	The Gaussian blur is a type of image-blurring filter that uses a Gaussian function.	• It has two parameters to set the minimum and maximum standard deviation to be used for creating a kernel to perform blurring. These parameters are Min Sigma and Max Sigma. • It also has two parameters to set the height and width of the Gaussian kernel. • The default value for the Min Sigma is set to 0.1 and the default value of the Max Sigma is set to 2.0. • The default value of height and width is set to 3 for each.

Options	Description	Parameters
Adjust Sharpness	Randomly increases the sharpness of the image with a probability.	It has a parameter called the Sharpness Factor to adjust the sharpness intensity. It can be any non-negative number. 0 gives a blurred image, 1 gives the original image while 2 increases the sharpness by a factor of 2. The default value of the Sharpness Factor is set to 2.
Posterization	Randomly reduces the number of color or tone levels in an image with a given probability.	It has a parameter to set the number of bits to keep for each channel. The value can be between (0 and 8). The default value of bits is set to 4.
Solarize	Randomly inverts tones in an image with a given probability for a given threshold for pixel values.	It has a parameter to set the Pixel Threshold. All pixels equal to or above this value are inverted. The default value of the Pixel Threshold is set to 50.0.
Invert	Randomly inverts the color or tone of pixels in an image.	Select this option to apply invert to the pixels of the images.
Auto Contrast	Randomly applies autocontrast to an image.	Select this option to apply auto contrast.

Options	Description	Parameters
Equalizer	Equalization enhances image contrast by redistributing the most common intensity values, effectively widening the image's intensity range.	Select this option to apply equalizer.

Transforms

The following table provides the options to select along with the available parameters to set in each option:

Options	Description	Parameters
Horizontal Flip	Randomly flips the image horizontally.	Select this option to apply a horizontal flip.
Vertical Flip	Randomly flips the image vertically.	Select this option to apply a vertical flip.

Image Augmentation

The following table provides the options to select along with the available parameters to set in each option:

Options	Description	Parameters
Auto Augment	Applies auto-augmentation using CIFAR-10, SVHN, and ImageNet schemes.. CIFAR-10 and AutoAugment pick mostly color-based transformations. They are Equalize, AutoContrast, Color, and Brightness. The most commonly picked transformations on SVHN are Invert, Equalize, ShearX/Y, and Rotate. Most of the policies found on ImageNet used color-based transformations.	It has a parameter called Augment Policy. This is a drop-down menu with the possible values are CIFAR-10, SVHN, and ImageNet. The default value is set to CIFAR-10.
Augment	Randomly apply augmentation processing techniques.	It has the following parameters to enter: • Augment Number- It is the number of augmentation transformations to apply sequentially. It is an integer field. The default value is set to 2. • Magnitude- It defines the magnitude for all the transformations. It is an integer field. The default value is set to 9. • Magnitude Bins- It defines the number of different magnitude values. It is an integer value. The default value is set to 31.

Image Feature & Annotation

The options for Document Extraction are given under two sets:

- Photometric Transforms
- Transforms

Photometric Transforms

The following table provides the options to select along with the available parameters to set in each option.

Options	Description	Parameters
Color Jitter	Color Jitter is a data augmentation technique that varies the brightness, contrast, hue, and saturation of the sample images.	It has eight parameters to set for min and max values of Brightness, Contrast, Hue, and Saturation factors of the target image. These are the numeric fields. The default value for all min values is set to 0. The default value for all max values is set to 0.5.
Adjust Sharpness	Randomly increases the image sharpness with a probability.	Sharpness Factor to adjusts the sharpness intensity. It can be any non-negative number. 0 gives a blurred image, 1 preserves the original image, while 2 increases the sharpness by a factor of 2. The default value of the Sharpness Factor is set to 2.
Auto Contrast	Randomly applies autocontrast to an image based on the given probability.	Select this option to apply auto contrast.

Transforms

The following table provides the options to select along with the available parameters to set in each option.

Options	Description	Parameters
Horizontal Flip	Randomly flips the image horizontally.	Select this option to apply a horizontal flip.
Vertical Flip	Randomly flips the image vertically.	Select this option to apply a vertical flip.

Vectorizing text

Text Vectorization comprises the following nodes:

- Fast Text
- TFIDF

Fast Text

A word embedding technique that captures semantic relationships between words by considering character n-grams.

The following table provides the parameters to set in each option under this category:

Options	Description	Parameters
Training Algorithm	Training Algorithm optimizes neural network weights.	Select the algorithm from the dropdown based on the requirement. The available options are CBOW and Skip-Gram. The default selected value is CBOW.
Epochs	It is the number of iterations (epochs) over the complete text corpus.	Enter the appropriate integer value based on your requirement. The default value is set to 5.

Options	Description	Parameters
Learning Rate	It defines the initial learning rate.	Enter the appropriate numerical value between 0 and 1. The default value is set to 0.001.
Vector Size	It defines the dimensionality of the word vectors.	Enter the appropriate integer value. The default value is set to 100.
Min Word Count	It defines the minimum required frequency (count of the occurrence of the word) in the text corpus for a word to be considered for modeling. The model ignores all words with a total frequency lower than this.	Enter the appropriate integer value. The default value is set to 5.
Window	It defines the maximum distance between the current and predicted words within a sentence.	Enter the appropriate integer value. The default value is set to 5.

TF-IDF

TF-IDF stands for Term Frequency-Inverse Document frequency. A statistical embedding technique that reflects the importance of a word in a document compared to the entire corpus. Words appearing frequently in a document but rarely overall get a higher tf-idf weight, indicating their relevance to that specific document.

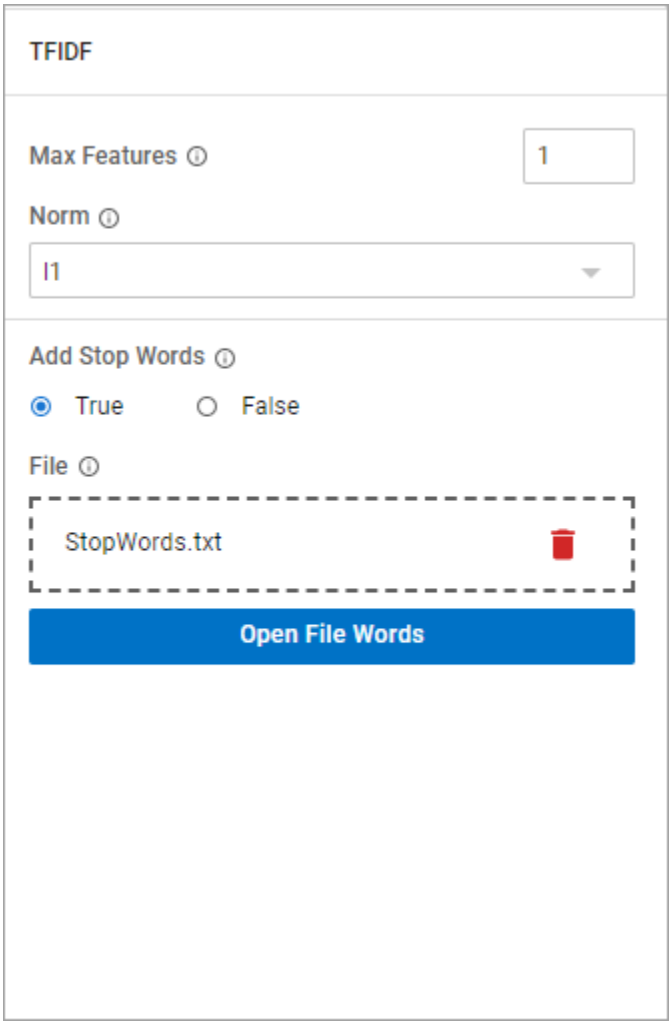
The following table provides the parameters to set in each option under this category:

Options	Description	Parameters
Norm	Each output row possesses a unit norm, which can be one of the following: • l2: The sum of squares of vector elements is 1. The cosine similarity between two vectors is their dot product when the l2 norm has been applied. • l1: Sum of absolute values of vector elements is 1.	Select the appropriate norm from the dropdown. The available options are: l1 and l2. The default value is set to l1.

Options	Description	Parameters
Custom Stop Words	Include files or words containing terms that are considered stop words and must be removed during tokenization.	Select True and import the file with stop words. For procedural details see Adding stop words

Adding stop words

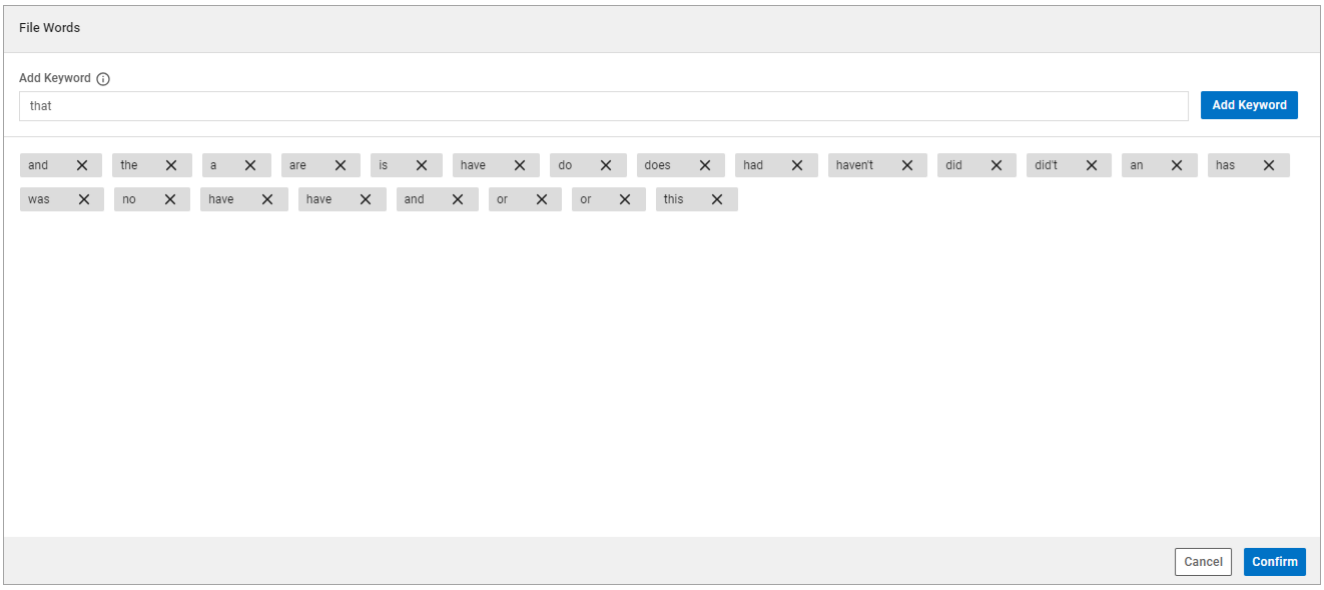
1. Drag and drop the TFIDF node in the canvas area. You can connect to the preceding and succeeding nodes your requirement.
2. Click the **TFIDF** node. The parameter options appear on the right side of the canvas area.
3. Select **True** for the Add Stop Words option.
4. Drag and drop your text file which has the stop words from your computer to the area in the box with a dotted -- perimeter. Alternatively, you can click the **browse** link and import your stop word file.



You can remove the imported stop word file by clicking the icon .
A stop word text file format is with one word per line as shown in the sample file:

```
and
the
a
are
is
have
do
does
had
haven't
did
did't
An
has
was
no
HAVE
Have
And
OR
or
```

5. Click the **Open File Words** button. The File Words dialog appears:



6. To add more stop words, enter the stop word and click **Add Keyword**. Keep adding as many stop words as you need.
7. Click **Confirm** to save.

Mapping images

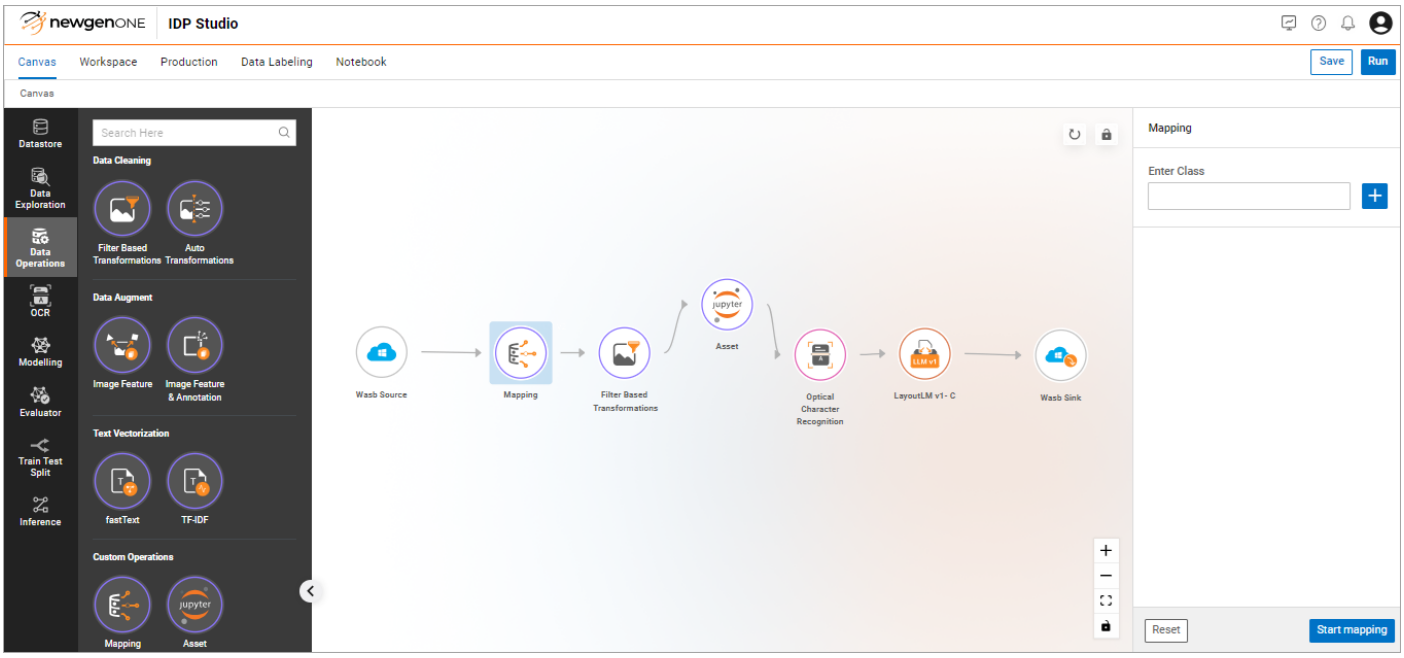
Custom operation allows you to map document images to the defined classes and for select one of the deployed notebooks. To see how to deploy a notebook, refer to the section for the Notebook option on the menu bar.

This section contains the following nodes:

- Mapping
- Asset

Mapping

Drag and drop this node to the canvas area and connect with the required preceding and succeeding nodes. By clicking the node in the canvas area, the Mapping parameters appear on the right side of the screen.

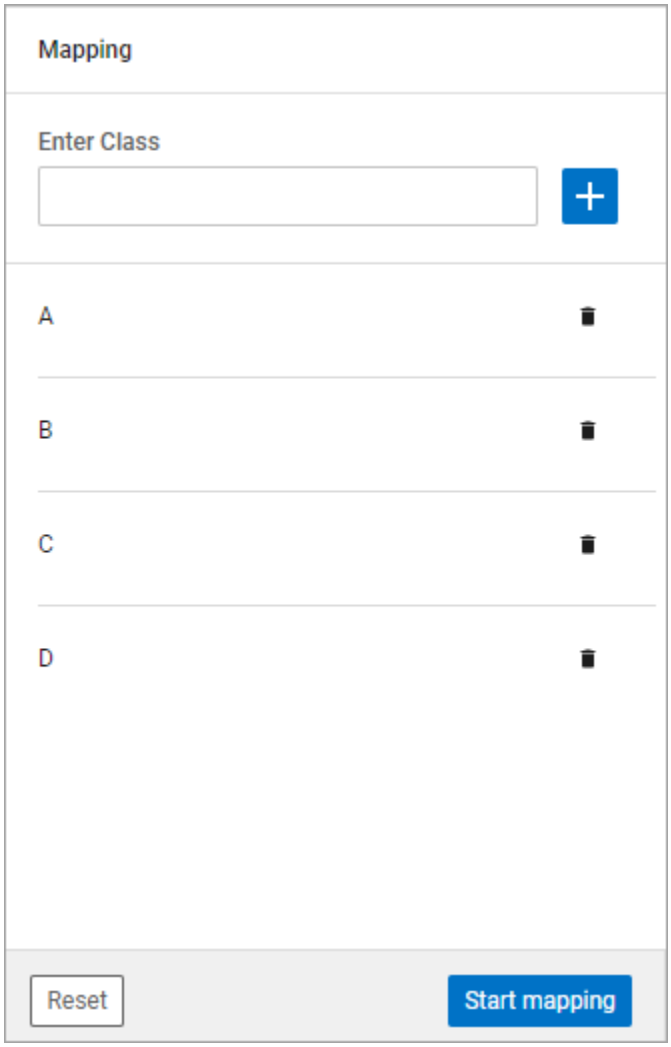


Following are the steps to setup this node:

1. Add custom classes
2. Mapping of folders or directories with the classes

Add custom classes

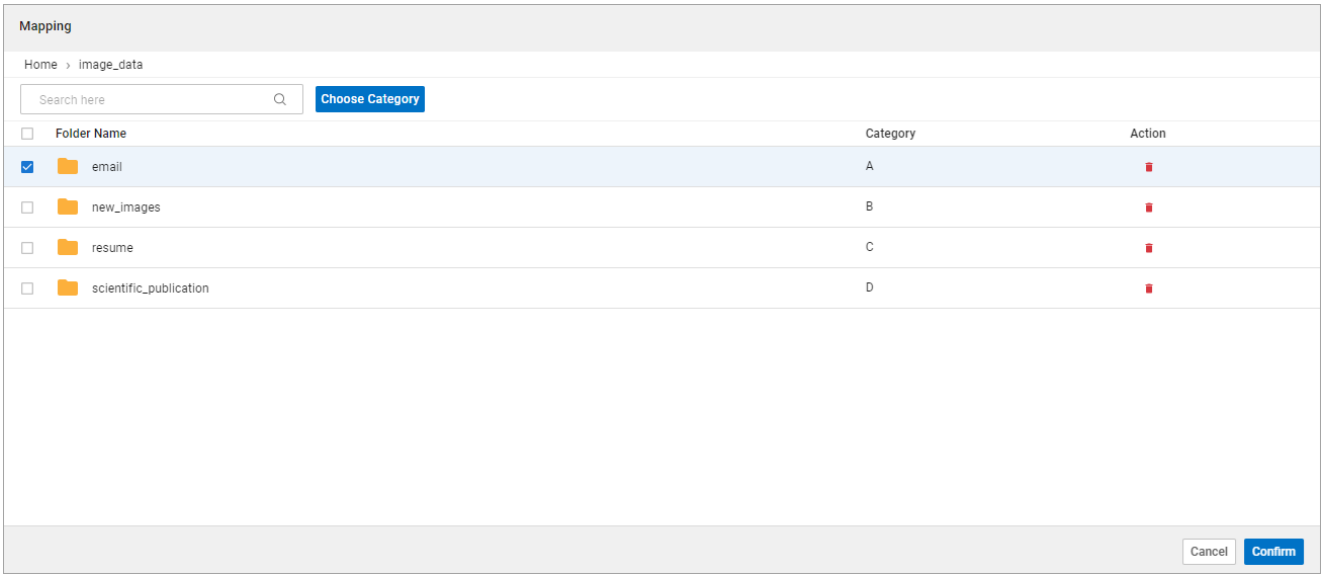
1. Select the mapping node.
2. On the right side you see the options to add classes.



3. Enter your class name in the text box and click the plus sign to include it in the list.
   You can name your class as required. For example, you have the images of the document types - invoices, bills, contracts, terms, and conditions under separate respective folders or directories. Then you can add the class names as invoices, bills, contracts, terms, and conditions.
4. Click **Start mapping**. The Mapping dialog appears.
   You can reset the list by clicking the **Reset** button.

Mapping of folders or directories with the classes

1. In the Mapping dialog, navigate to the folder or directory that has data as per any of the classes.



2. Select the folder or directory and click **Choose Category**. The Edit Class dialog appears.

   On this screen, you can also remove an existing mapping by clicking the icon.

3. Select one of the predefined custom classes or create a new class by clicking the **Create Custom** link.

4. Click **Save** to save your selection and close the dialog.
5. Click **Confirm** to save your mappings.

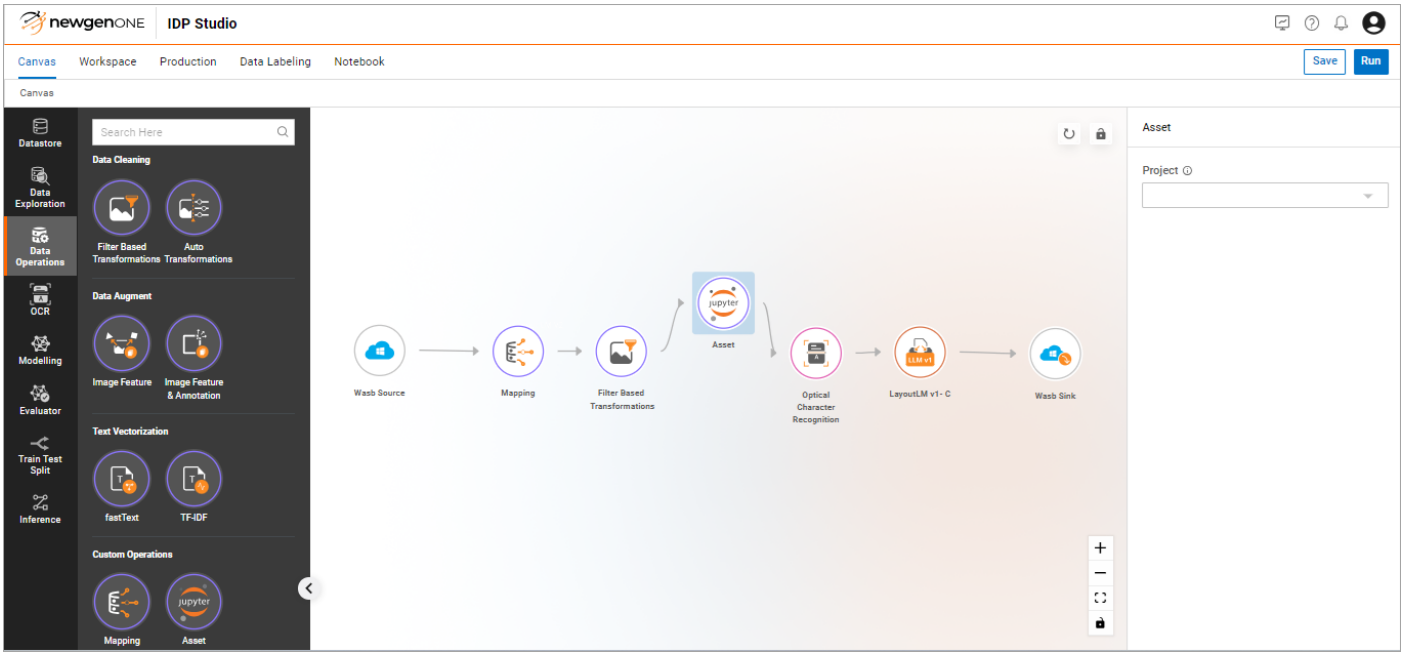
Asset

Drag and drop this node to the canvas area and connect with the required preceding and succeeding nodes. By clicking the node in the canvas area the Notebook parameters appear on the right side of the screen.

To define the properties of the Asset node, perform the following steps:

1. Select the project from the dropdown menu.
2. This project was created under the Notebook tab available on the upper screen menu.

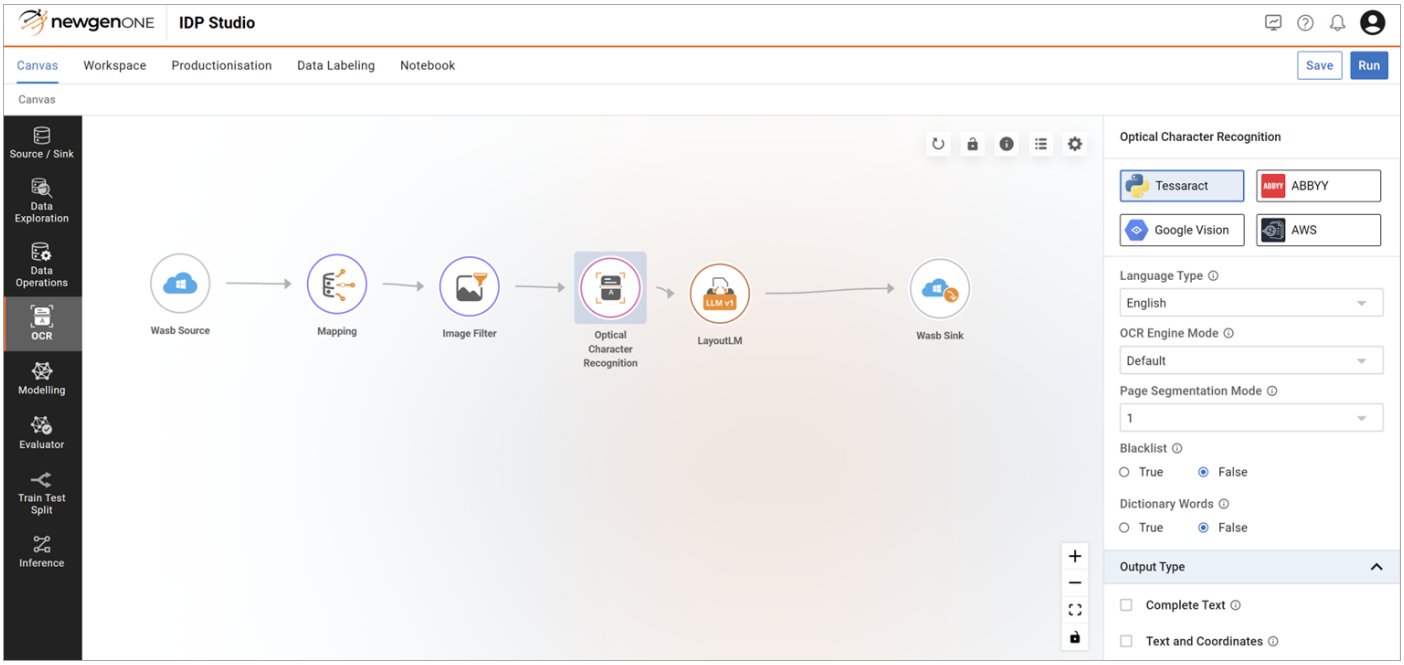
3. Steps to create a new project, creating and deploying a notebook are given in the Notebook section.
4. Select the asset from the dropdown menu under the selected project. This dropdown field has the list of notebook names that are under the selected project and were deployed.



Optical character recognition

Optical character recognition (OCR) is a technology that allows machines, typically computers, to interpret and recognize text characters from images or scanned documents. It plays a crucial role in converting printed or handwritten text into machine-readable text data.

To add an OCR node in the pipeline, drag and drop it to the canvas area and connect it with the preceding and succeeding nodes in the pipeline.



The options to select for OCR are given under the following sets:

- OCR Type
- Output Type

OCR Type

There are the following options to select and configure the parameters:

- **Tesseract** — Allows you to configure the following parameters:
  - **Language Type** — This is a dropdown menu that allows you to select the language of the document or text in the image. The available options are English, German, French, Spanish, and Hindi.

The default selected value is English.

- **OCR Engine Mode** — This is a dropdown menu that allows you to select the required OCR engine mode from the available options: Legacy engine, LSTM engine, Both Legacy + LSTM engine, Default.

  The default value is selected to Default.

- **Page Segmentation Mode** — This is a dropdown menu with available options as 1,2,3,4. Following are the meanings for each of these options:
  - **1** —Fully automatic page segmentation, but no OSD.
  - **2** —Assume a single uniform block of text.
  - **3** —Sparse text. Find as much text as possible in no order.
  - **4** —Sparse text with OSD.

    The default value is selected as 1.

- **Blacklist** — Select **True** and provide the Blacklist Characters in the text box. To exclude them from recognition by Tesseract. It is case-sensitive. For example, abcdefghijklmnopqrstuvwxyz to not detect any lower-case alphabets in OCR.

  The default value is selected as False.

- **Dictionary Words** — Select **True** to browse and upload the text file having a dictionary of words. This file can have words as well as special characters where each line in the file has a single word or special character.

  The default value is selected as False.

  The image shows the data format in the dictionary file:

```
ANHOI
HELL
INDUSTRIAL
VEHICLE
IMP
AND
EXP
CO
.
LTD
15
WEST
WANG
ANG
ROAD
,
HEFEI
ANHUI
CHINA
TO
ORDER
OF
ICICI
BANK
MUMBAI
+
GEMINI
EQUIPMENT
RENTALS
PVT
HOUSE
```

- **Google Vision** — To use the Google Vision OCR, configure the following parameters:
  - **Language Type** — Allows you to select the language of the document or text in the image. The options include English, German, French, Spanish, and Hindi.

    The default value selected is English.

  - **Feature Type** — Allows you to select the feature type. The possible values are TEXT DETECTION, and DOCUMENT TEXT DETECTION.

    The default value selected is TEXT DETECTION.

- **AWS** — To use the Amazon Textract as OCR, configure the following parameters:
  - **Feature Type** — Allows you to select the feature types. The possible values are Tables, Forms, and Signatures. You can select one or more options.

- **ABBYY** — It provides a set of tools to allow the conversion of different types of documents, such as scanned images, PDF files, or photographs, into editable and searchable text.
  To use ABBYY, configure the following parameters:
  - **Language Type** — Select the language of the document or text in the image. The options include English, German, French, and Spanish.

    The default value selected is English.

  - **Recognition Quality** — Select the mode to optimize the recognition of the image. Possible options are:
    - **Thorough** — Recognition runs in a full mode that is optimized for quality.
    - **Balanced** — Recognition runs in a balanced mode in terms of speed and quality.
    - **Fast** — Recognition runs in a fast mode at the cost of a moderately increased error rate.

      The default selected value is Thorough.

- The following OCR tools are available only on cloud implementation of deployment of the IDP Studio platform:
  - AWS Textract
  - Google Vision
- The following OCR tools are available on both on-premises and cloud implementation of deployment of the IDP Studio platform:
  - Tesseract
  - ABBYY
- Following are the prerequisites that each customer of IDP Studio must fulfill to have respective OCR tool in the platform:
  - ABBYY server engine is required to enable it.
  - AWS app clientid and secret is required for using AWS Textract.
  - It is required to enable Cloud Vision API corresponding to a particular Google account. This Google account must be an organization account and not a Gmail account. After enabling the Cloud Vision API, It is required to get the credentials in JSON format in the API/Service Details section of the same API.

Output Type

Depending on your requirement, select an output type from the following options:

- **Complete Text** — Select this option if you want Complete Text as output. There are no additional parameters to configure.

- **Text and Coordinates** — Select this option if you want Text along with coordinates as output. There are no additional parameters to configure.

Splitting model

This option under the canvas is for splitting the data into training, testing, and validation sets. In machine learning, model training is done on the training data while the validation set is used for the evaluation of the models. The final model is tested on the testing data.

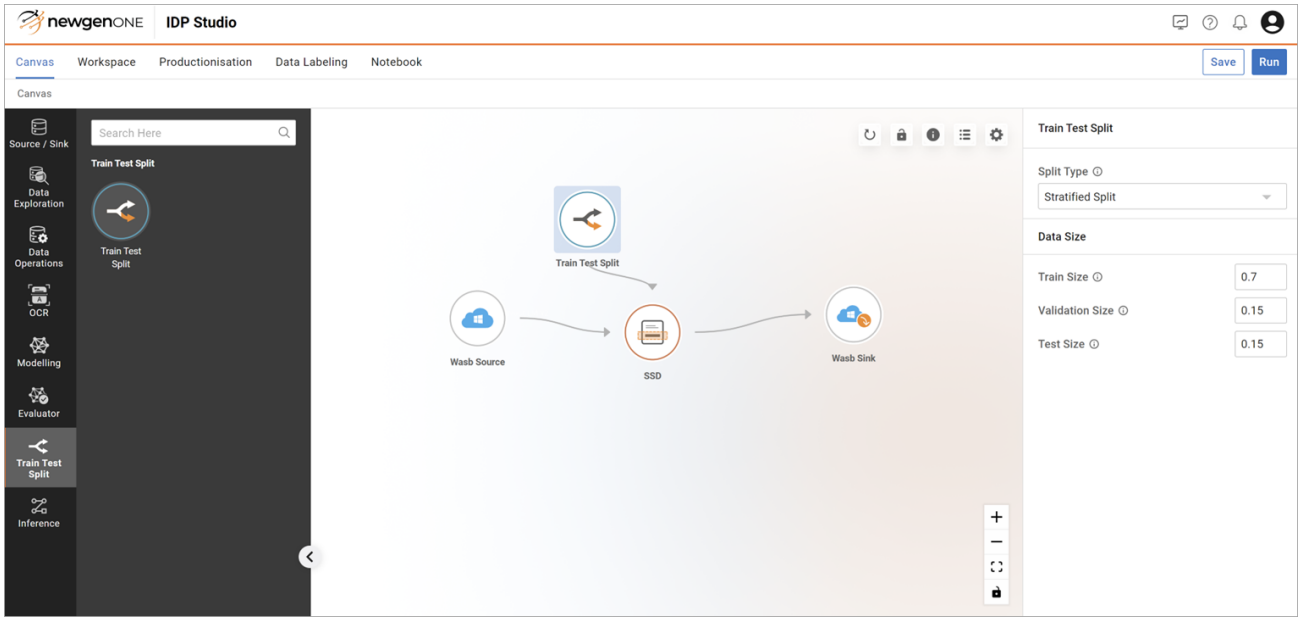
The Model Split option includes the Train Test Split node.

Train Test Split

To add a Train Test Split node in the pipeline, drag and drop it to the canvas area and connect it with the preceding and succeeding nodes in the pipeline.

Following are the constraints for this node:

- Can connect to nodes in the Data Augmentation and modelling nodes only.
- There is only one Train Test Split node on the canvas for a pipeline.
- No node can connect before this node; it is considered a special node.



It has the following parameters to configure:

- Split Type: . The default value is selected as Random. The other available option is Stratified Split.
- Data Size: There are the following parameters to set under this category:
    - Train Size
    - Validation Size
    - Test Size

These parameters accept numerical values between 0 to 1, excluding 0 and 1. The default values are set as 0.70, 0.15, and 0.15, respectively, for Train Size, Validation Size, and Test Size.

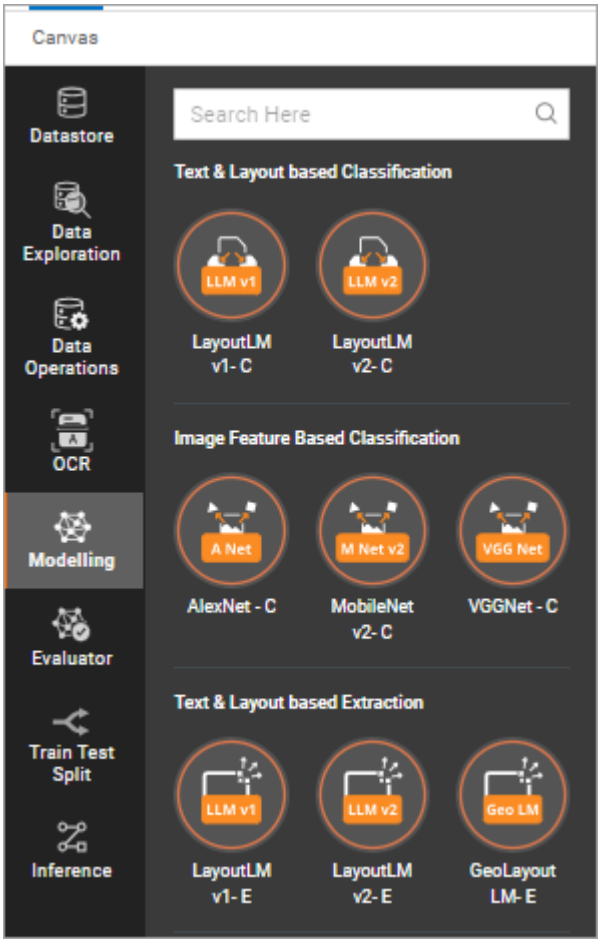
If the train test node is not used in a model development pipeline, the following default values are used by the IDP Studio application:

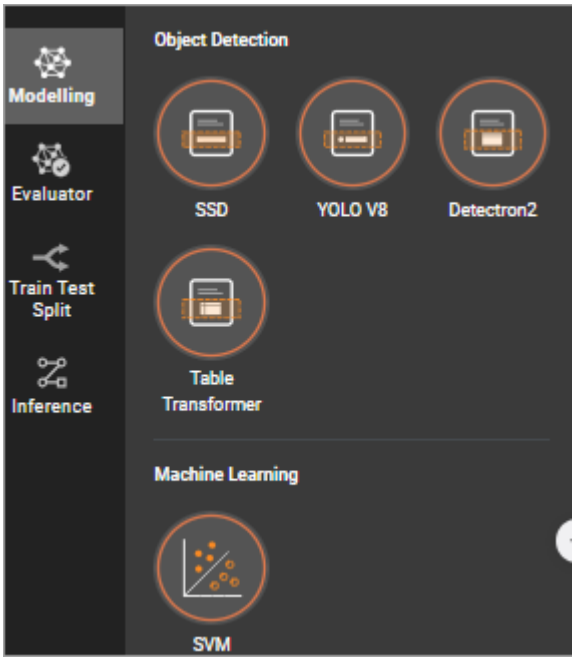
- Train Size: 0.70
- Validation Size: 0.20
- Test Size: 0.10

Modelling

This canvas option organizes nodes associated with modeling algorithms. It categorizes them into different algorithm types, each containing a specific cluster of nodes:

- Text & Layout based Classification
- Text & Layout based Extraction
- Image Feature Based Classification
- Object Detection
- Machine Learning





Consider the following points for Hyperparameter tuning:

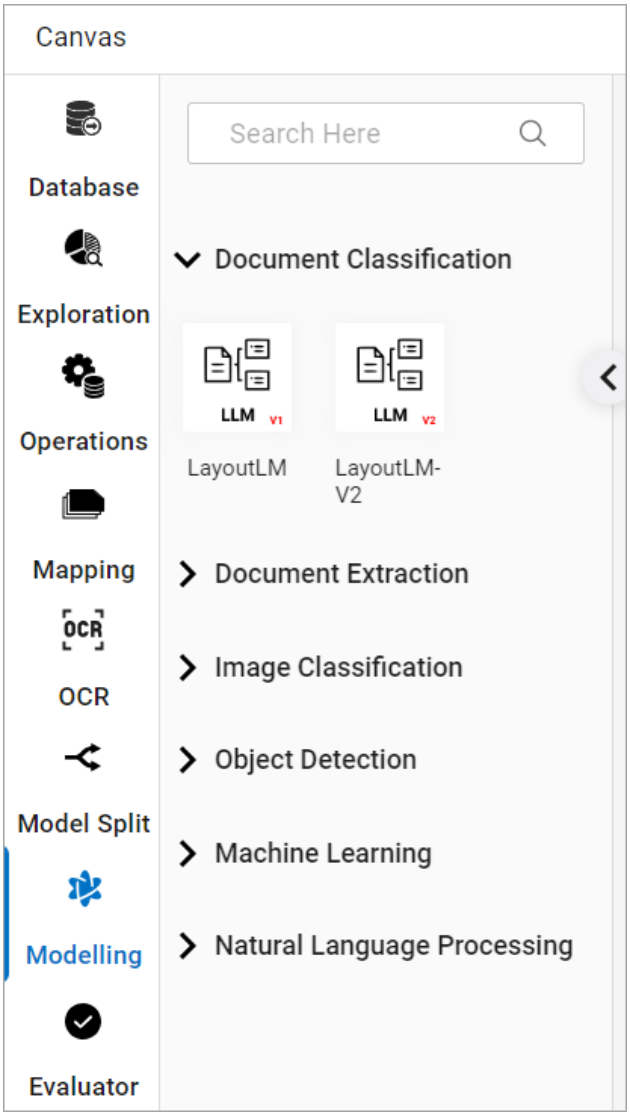
- Any of the algorithm modeling nodes under the above categories might contain hyperparameter options.
- These parameters act as hyperparameters when the flag to enable hyperparameter tuning is True. In this case, the label against the input field for the parameter is appended with (H).
- Multiple values separated by a comma are allowed for such parameter options.

Text and layout-based classification

Text & Layout based Classification models are transformer models that are used to classify documents by jointly analyzing text and layout information in the document.

This category comprises the following algorithms:

- LayoutLM v1-C
- LayoutLM v2-C



LayoutLM v1-C

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters
- Model Layer Parameters
- Attention Mechanism Parameters
- Embedding Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameters configurations.	Select True to enable hyperparameter tuning. • The default value is set to False.
Pre - Trained Model	This option is a flag to enable pre-trained models to train the current model.	Select True to use pre-trained models in the training. • The default value is set to False.
Shuffle	This option is a flag to enable data shuffling during the training process.	Select True to shuffle the data. • The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. • The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. • The default value is set to 8.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. • The default value is set to 4.

Option	Description	Parameter
Evaluation Metrics	Metric that is used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics. The possible options are: • Accuracy • Precision Micro Average • Precision Macro Average • Recall Micro Average • Recall Macro Average • Fmeasure Micro Average • Fmeasure Macro Average The default selected value is Accuracy.
Model Save Frequency	Models get saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.
Metric Save Frequency	Model Metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics get saved. The default value is set to 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is SGD.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select True to enable the learning rate scheduler. The default value is set to False.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate scheduler. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set to 0.001.

Option	Description	Parameter
Momentum	Momentum is a method that helps to accelerate gradient vectors in the right directions, thus leading to faster convergence by accumulating the gradient of the past steps to determine the right direction.	Enter the appropriate number. The default value is set to 0.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

Model Layer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hidden Activation Function	The non-linear activation function (function or string) in the encoder and pooler.	Select the appropriate activation function. The available values are: • relu • gelu • gelu_new The default selected value is gelu.
Hidden Dropout Probability	The dropout probability for all fully connected layers in the embeddings, encoder, and pooler.	Enter the probability. It is a numerical value between 0 and 1. The default value is set to 0.1.
Layer Norm Epsilon	The epsilon is used by the layer normalization layers.	Enter the appropriate value. The default value is set as 1e-12.

Option	Description	Parameter
Initializer Range	The standard deviation of the truncated_normal_initializer for initializing all weight matrices.	Enter the appropriate value. The default value is set to 0.02.

Attention Mechanism Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Attention Dropout Probability	The dropout ratio for the attention probabilities.	Enter an appropriate float number between 0 and 1. The default value is set to 0.1.

Embedding Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Position Embedding Type	Type of position embedding.	Select the appropriate value. The available values are: • Absolute • Relative Key The default selected value is Absolute.
Max Position Embeddings	The maximum sequence length that this model might ever be used with. This option is available only when the Pre-Trained Model is set to False.	Enter the appropriate number. The default value is set to 512.
Max 2D Position Embeddings	The maximum value that the 2D position embedding might ever be used with.	Enter the appropriate number. The default value is set to 1024.
Vocab Size	Vocabulary size of the LayoutLMv2 model. Defines the number of different tokens that can be represented by the inputs_ids passed when calling LayoutLMv2Model.	Enter the appropriate number. The default value is set to 30522.

Option	Description	Parameter
Type Vocab Size	The vocabulary size of the token_type_ids passed when calling LayoutMv2Model.	Enter the appropriate number. The default value is set to 2.

LayoutLM v2-C

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters
- Model Layer Parameters
- Attention Mechanism Parameters
- Embedding Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.
Pre - Trained Model	This option is a flag, if set to True then pre-trained models can be used to train the current model.	Select True to use pre-trained models in the training. The default value is set to False.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. The default value is set to False.

Option	Description	Parameter
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. The default value is set to 8.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. The default value is set to 4.
Convert Sync Batchnorm	Synchronize the batch normalization layers of the visual backbone. This option is a flag, if set to True the batch normalization layers get synchronized.	Select True to enable the synchronization of the batch normalization layers. The default value is set to False.
Evaluation Metrics	Metric that is used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • Accuracy • Precision Micro Average • Precision Macro Average • Recall Micro Average • Recall Macro Average • Fmeasure Micro Average • Fmeasure Macro Average The default selected value is Accuracy.

Option	Description	Parameter
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics must be saved. The default value is set to 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is SGD.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select True to enable the learning rate scheduler. The default value is set to False.

Option	Description	Parameter
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set to 0.001.
Momentum	Momentum is a method that helps to accelerate gradient vectors in the right directions, thus leading to faster convergence by accumulating the gradient of the past steps to determine the right direction.	Enter the appropriate number. The default value is set to 0.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

Model Layer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hidden Activation Function	The non-linear activation function (function or string) in the encoder and pooler.	Select the appropriate activation function. The available values are: • relu • gelu • gelu_new The default selected value is gelu.
Number of Hidden Layer	Several hidden layers in the Transformer encoder.	Enter the suitable number. The default value is set to 12.
Hidden Dropout Probability	The dropout probability for all fully connected layers in the embeddings, encoder, and pooler.	Enter the probability. It is a numerical value between 0 and 1. The default value is set to 0.1.
Intermediate Size	The dimensionality of the intermediate (feed-forward) layer in the Transformer encoder.	Enter the appropriate value. The default value is set to 3072.
Layer Norm Epsilon	The epsilon is used by the layer normalization layers.	Enter the appropriate value. The default value is set as 1e-12.
Initializer Range	The standard deviation of the truncated_normal_initializer for initializing all weight matrices.	Enter the appropriate value. The default value is set to 0.02.

Option	Description	Parameter
Image Feature Pool Height	The Height of the average-pooled feature map.	Enter the appropriate value. • The default value is set to 7.
Image Feature Pool Width	The Width of the average-pooled feature map.	Enter the appropriate value. • The default value is set to 7.

Attention Mechanism Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Number of Attention Heads	Number of attention heads for each attention layer in the Transformer encoder.	Enter the appropriate number. • The default value is sett to 12.
Attention Dropout Probability	The dropout ratio for the attention probabilities.	Enter an appropriate float number between 0 and 1. • The default value is set to 0.1.
Max Relative Positions	The maximum number of relative positions to be used in the self-attention mechanism.	Enter the appropriate number. • The default value is set to 128.
Relative Positions Bins	The number of relative position bins to be used in the self-attention mechanism.	Enter the appropriate number. • The default value is set to 32.
Max Relative 2D Position	The maximum number of relative 2D positions in the self-attention mechanism.	Enter the appropriate number. • The default value is set to 256.
Relative 2D Position Bins	The number of 2D relative position bins in the self-attention mechanism.	Enter the appropriate number. • The default value is set to 64.

Option	Description	Parameter
Relative Attention Bias	It is a flag to indicate whether to use a relative attention bias in the self-attention mechanism.	Set the value to True to use a relative attention bias. The default value is set as False.
Spatial Attention Bias	It is a flag to indicate whether to use a spatial attention bias in the self-attention mechanism.	Set the value to True to use a spatial attention bias. The default value is set as False.

Embedding Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Visual Segment Embedding	It is a flag to indicate whether or not to add visual segment embeddings.	Set the value to True to add visual segment embeddings. The default value is set to False.
Fast Query Key Value Optimization	It is a flag to indicate whether to use a single matrix for the queries, keys, and values in the self-attention layers.	Set the value to True to use a single matrix for the queries, keys, and values. The default value is set to True.
Max Position Embeddings	The maximum sequence length that this model might ever be used with.	Enter the appropriate number. The default value is set to 512.
Max 2D Position Embeddings	The maximum value that the 2D position embedding might ever be used with.	Enter the appropriate number. The default value is set to 1024.

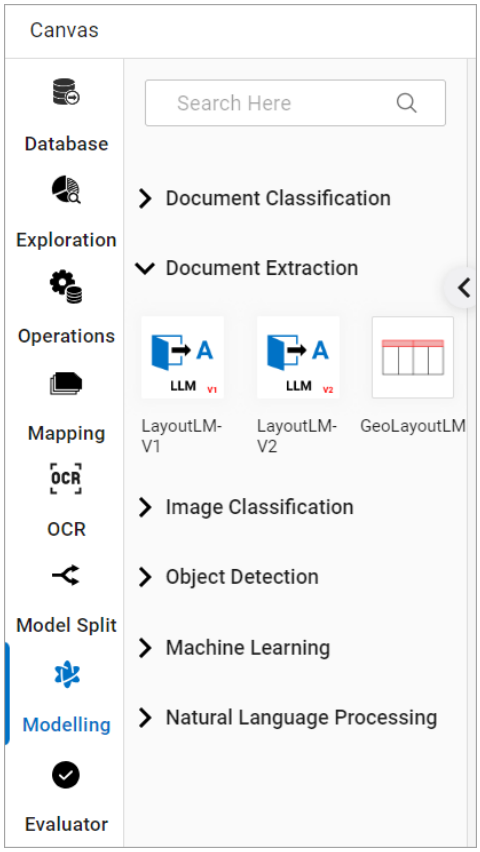
Option	Description	Parameter
Coordinate Size	Dimension of the coordinate embeddings.	Enter the appropriate number. The default value is set to 128.
Shape Size	Dimension of width and height embeddings.	Enter the appropriate number. The default value is set to 128.
Vocab Size	Vocabulary size of the LayoutLMv2 model. Defines the number of different tokens that can be represented by the inputs_ids passed when calling LayoutLMv2Model.	Enter the appropriate number. The default value is set to 30522.
Type Vocab Size	The vocabulary size of the token_type_ids passed when calling LayoutMv2Model.	Enter the appropriate number. The default value is set to 2.

Text and layout-based extraction

Text & Layout based Extraction models are transformer models used for information extraction from documents by combining the textual embeddings with positional information and layout information.

This category comprises the following algorithms:

- LayoutLM v1-E
- LayoutLM v2-E
- GeoLayoutLM-E



LayoutLM v1-E

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters
- Model Layer Parameters
- Attention Mechanism Parameters
- Embedding Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select False to not use pre-trained models in the training. The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. The default value is set to 40.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. The default value is set to 4.
Evaluation Metrics	Metric that is used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • Accuracy • Precision Micro Average • Precision Macro Average • Recall Micro Average • Recall Macro Average • Fmeasure Micro Average • Fmeasure Macro Average • The default selected value is Accuracy.

Option	Description	Parameter
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set to 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is AdamW.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select False to disable the learning rate scheduler. The default value is set to True.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.

Option	Description	Parameter
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set to 0.001.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

Model Layer Parameters

The following table provides the parameters to set in each option under this set

Option	Description	Parameter
Hidden Activation Function	The non-linear activation function (function or string) in the encoder and pooler.	Select the appropriate activation function. The available values are: • relu • gelu • gelu_new The default selected value is gelu.
Hidden Dropout Probability	The dropout probability for all fully connected layers in the embeddings, encoder, and pooler.	Enter the probability. It is a numerical value between 0 and 1. The default value is set to 0.1.
Layer Norm Epsilon	The epsilon is used by the layer normalization layers.	Enter the appropriate value. The default value is set to 1e-12.
Initializer Range	The standard deviation of the truncated_normal_initializer for initializing all weight matrices.	Enter the appropriate value. The default value is set to 0.02.

Attention Mechanism Parameters

The following table provides the parameters to set in each option under this set.

Option	Description	Parameter
Attention Dropout Probability	The dropout ratio for the attention probabilities.	Enter an appropriate float number between 0 and 1. The default value is set to 0.1.

Embedding Parameters

The following table provides the parameters to set in each option under this set.

Option	Description	Parameter
Position Embedding Type	Type of position embedding.	Select the appropriate value. The available values are: • Absolute • Relative Key The default selected value is Absolute.
Max Position Embeddings	The maximum sequence length that this model might ever be used with. This option is available only when the Pre-Trained Model is set to False.	Enter the appropriate number. The default value is set to 512.
Max 2D Position Embeddings	The maximum value that the 2D position embedding might ever be used with.	Enter the appropriate number. The default value is set to 1024.
Vocab Size	Vocabulary size of the LayoutLMv2 model. Defines the number of different tokens that can be represented by the inputs_ids passed when calling LayoutLMv2Model.	Enter the appropriate number. The default value is set to 30522.
Type Vocab Size	The vocabulary size of the token_type_ids passed when calling LayoutMv2Model.	Enter the appropriate number. The default value is set to 2.

LayoutLM v2-E

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters
- Model Layer Parameters
- Attention Mechanism Parameters
- Embedding Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select True to use pre-trained models in the training. The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. The default value is set to 8.

Option	Description	Parameter
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. The default value is set to 4.
Convert Sync Batchnorm	Synchronize the batch normalization layers of the visual backbone. This option is a flag, if set to True the batch normalization layers get synchronized.	Select True to enable the synchronization of the batch normalization layers. The default value is set to False.
Evaluation Metrics	Metric that is used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirement. The possible options are: • Accuracy • Precision Micro Average • Precision Macro Average • Recall Micro Average • Recall Macro Average • Fmeasure Micro Average • Fmeasure Macro Average The default selected value is Accuracy.
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set to 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is AdamW.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select True to enable the learning rate scheduler. The default value is set to False.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set to 0.001.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.

Option	Description	Parameter
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

Model Layer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hidden Activation Function	The non-linear activation function (function or string) in the encoder and pooler.	Select the appropriate activation function. The available values are: • relu • gelu • gelu_new The default selected value is gelu.
Hidden Dropout Probability	The dropout probability for all fully connected layers in the embeddings, encoder, and pooler.	Enter the probability. It is a numerical value between 0 and 1. The default value is set to 0.1.
Layer Norm Epsilon	The epsilon is used by the layer normalization layers.	Enter the appropriate value. The default value is set as 1e-12.
Initializer Range	The standard deviation of the truncated_normal_initializer for initializing all weight matrices.	Enter the appropriate value. The default value is set to 0.02.

Attention Mechanism Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Attention Dropout Probability	The dropout ratio for the attention probabilities.	Enter an appropriate float number between 0 and 1. The default value is set to 0.1.
Max Relative Positions	The maximum number of relative positions to be used in the self-attention mechanism.	Enter the appropriate number. The default value is set to 128.
Max Relative 2D Position	The maximum number of relative 2D positions in the self-attention mechanism.	Enter the appropriate number. The default value is set to 256.
Relative Attention Bias	It is a flag to indicate whether to use a relative attention bias in the self-attention mechanism.	Set the value to False to not use a relative attention bias. The default value is set to True.
Spatial Attention Bias	It is a flag to indicate whether to use a spatial attention bias in the self-attention mechanism.	Set the value to False to not use a spatial attention bias. The default value is set to True.

Embedding Parameters

Set the following parameter:

- **Fast Query Key Value Optimization** — It is a flag to indicate whether to use a single matrix for the queries, keys, and values in the self-attention layers. Set the value to False if you do not want to use a single matrix for the queries, keys, and values.

  It is set to True by default.

GeoLayoutLM-E

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters
- Model Layer Parameters
- Attention Mechanism Parameters
- Embedding Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select False to not use pre-trained models in the training. The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. The default value is set to 10.

Option	Description	Parameter
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. The default value is set to 4.
Evaluation Metrics	Metrics that are used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirement. The possible options are: • Accuracy • Precision • Recall • F1-Score The default selected value is Accuracy.
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set to 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is AdamW.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select False to disable the learning rate scheduler. The default value is set to True.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set to 0.005.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate. The default value is set to 0.1.

Model Layer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hidden Activation Function	The non-linear activation function (function or string) in the encoder and pooler.	Select the appropriate activation function. The available values are: • relu • gelu • gelu_new The default selected value is gelu.
Hidden Dropout Probability	The dropout probability for all fully connected layers in the embeddings, encoder, and pooler.	Enter the probability. It is a numerical value between 0 and 1. The default value is set to 0.1.
Layer Norm Epsilon	The epsilon is used by the layer normalization layers.	Enter the appropriate value. The default value is set to 0.005
Initializer Range	The standard deviation of the truncated_normal_initializer for initializing all weight matrices.	Enter the appropriate value. The default value is set to 0.02.

Embedding Parameters

Specify the following parameter:

- **Positional Encoding Type** — It is a type of position embedding. Position embeddings are added to input data to provide information about the relative positions of elements in a sequence. It is a drop-down field. The available values are crel and sincos.

  The default value is set as crel.

Attention Mechanism Parameters

Specify the following parameter:

- **Attention Dropout Probability** — Dropout probability for attention layers. This parameter controls the dropout rate specifically for attention mechanisms. Enter an appropriate float number between 0 and 1 in the text box.

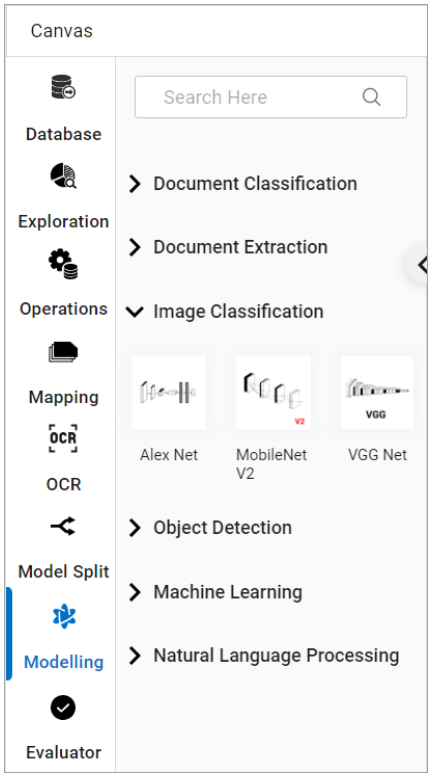
  The default value is set to 0.1.

Image feature-based classification

Image Feature based Classification models are CNNs trained to extract informative features (like edges and shapes) from images and then classify documents based on these features.

This category comprises the following algorithms:

- AlexNet-C
- MobileNet V2-C
- VGGNet-C



AlexNet-C

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. • The default value is set to False.
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select False to not use pre-trained models in the training. • The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. • The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. • The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. • The default value is set to 8.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. • The default value is set to 4.

Option	Description	Parameter
Evaluation Metrics	Metrics that are used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • Accuracy • Precision Micro Average • Precision Macro Average • Recall Micro Average • Recall Macro Average • F1-Score • Fmeasure Micro Average • Fmeasure Macro Average The default selected value is Accuracy.
Dropout	Randomly makes some of the elements of the input tensor to zeros (in other words it drops some of the randomly chosen elements) with probability p using samples from a Bernoulli distribution.	Enter a suitable positive float value less than 1. The default value is set to 0.5.
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set to 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is SGD.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select True to enable the learning rate scheduler. The default value is set to False.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set to 0.001.
Momentum	Momentum is a method that helps to accelerate gradient vectors in the right directions, thus leading to faster convergence by accumulating the gradient of the past steps to determine the right direction.	Enter the appropriate number. The default value is set to 0.

Option	Description	Parameter
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

MobileNet V2-C

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select False to not use pre-trained models in the training. The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. The default value is set to False.

Option	Description	Parameter
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. The default value is set to 8.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. The default value is set to 4.
Evaluation Metrics	Metrics that are used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • Precision Micro Average • Precision Macro Average • Recall Micro Average • Recall Macro Average • Fmeasure Micro Average • Fmeasure Macro Average • The default selected value is Accuracy.
Dropout	Randomly makes some of the elements of the input tensor to zeros (in other words it drops some of the randomly chosen elements) with probability p using samples from a Bernoulli distribution.	Enter a suitable positive float value less than 1. The default value is set to 0.5.
Width Multiplier	Adjusts the number of channels in each layer by this amount.	Enter the appropriate decimal value. The default value is set to 1.0.

Option	Description	Parameter
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set to 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is SGD.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select True to enable the learning rate scheduler. The default value is set to false.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.

Option	Description	Parameter
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate. The default value is set to 0.001.
Momentum	Momentum is a method that helps to accelerate gradient vectors in the right directions, thus leading to faster convergence by accumulating the gradient of the past steps to determine the right direction.	Enter the appropriate number. The default value is set to 0.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

VGGNet-C

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.

Option	Description	Parameter
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select False to not use pre-trained models in the training. • The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. • The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. • The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. • The default value is set to 8.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. • The default value is set to 4.
Batch Normalizer	It is a flag to use Batch Normalization during the model training process. Batch Normalization is typically used in Neural Networks to improve the training speed, stability, and performance. It normalizes the input of each layer across the mini-batch during training by adjusting and scaling the activations.	Set it to True to enable Batch Normalization. • The default value is set to False

Option	Description	Parameter
Evaluation Metrics	Metrics that are used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • Precision Micro Average • Precision Macro Average • Recall Micro Average • Recall Macro Average • Fmeasure Micro Average • Fmeasure Macro Average The default selected value is Accuracy.
Variant	This option requires you to select an appropriate variant of the VGG Net.	Select the appropriate value. The available values are: VGG- 11, VGG- 13, VGG- 16, VGG- 19. The default value is set to VGG- 19.
Dropout	Randomly makes some of the elements of the input tensor to zeros (in other words it drops some of the randomly chosen elements) with probability p using samples from a Bernoulli distribution.	Enter a suitable positive float value less than 1. The default value is set to 0.5.
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.

Option	Description	Parameter
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set to 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is SGD.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select True to enable the learning rate scheduler. The default value is set to False.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.

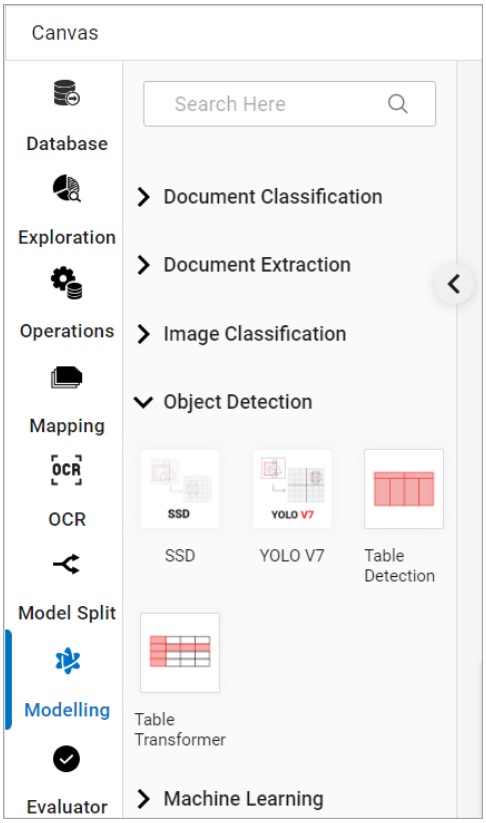
Option	Description	Parameter
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set to 0.001.
Momentum	Momentum is a method that helps to accelerate gradient vectors in the right directions, thus leading to faster convergence by accumulating the gradient of the past steps to determine the right direction.	Enter the appropriate number. The default value is set to 0.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

Object Detection

Object detection models are used for finding objects within an image. These models provide the box coordinates as well as the labels associated with that detected object.

This category comprises the following algorithms:

- SSD
- YOLO V8
- Detectron2
- Table Transformer



SSD

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Model Configuration
- Optimizer Parameters

General Parameters

The following table provides the parameters to set in each option under this set

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.

Option	Description	Parameter
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select False to not use pre-trained models in the training. • The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. • The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. • The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. • The default value is set to 8.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. • The default value is set to 4.
Evaluation Metrics	Metrics that are used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • MAP IOU 50 • MAP IOU 75 • MAP IOU 50_95 • The default selected value is MAP IOU 50.
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. • The default value is set to 2.

Option	Description	Parameter
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. • The default value is set to 2.

Model Configuration

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Trainable Backbone Layers	Number of trainable (non-frozen) layers starting from the final block. Valid values are between 0 and 5 meaning all backbone layers are trainable.	Enter the suitable integer value. • The default value is set to 4.
Score Thresh	Score threshold used for postprocessing the detections.	Enter the suitable numeric value. • The default value is set to 0.01.
NMS Thresh	NMS threshold used for postprocessing detections.	Enter the suitable numeric value. • The default value is set to 0.45.
IOU Thresh	The minimum IoU threshold between the anchor and the GT box for positive consideration during training.	Enter the suitable numeric value. • The default value is set to 0.5.
TopK Candidates	Number of best detections to keep before NMS.	Enter the suitable integer value. • The default value is set to 400.
Detections Per Image	Number of best detections to keep after NMS.	Enter the suitable integer value. • The default value is set to 200.

Option	Description	Parameter
Positive Fraction	A number between 0 and 1 that indicates the proportion of positive proposals used during the training of the classification head. It is used to estimate the negative to positive ratio.	Enter the suitable numeric value between 0 and 1. The default value is set to 0.25.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is SGD.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	SelectTrue to enable the learning rate scheduler. The default value is set to False.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set as 0.001.

Option	Description	Parameter
Momentum	Momentum is a method that helps to accelerate gradient vectors in the right directions, thus leading to faster convergence by accumulating the gradient of the past steps to determine the right direction.	Enter the appropriate number. The default value is set as 0.
Weight Decay	Weight decay for the optimizer.	Enter the suitable numeric value. The default value is set as 5e-4.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set as 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate numbe. The default value is set as 0.1.

YOLO V8

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.

Option	Description	Parameter
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select True to not use pre-trained models in the training. The default value is set to False.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. The default value is set to 8.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. The default value is set to 32.
Evaluation Metrics	Metrics that are used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • MAP IOU 50 • MAP IOU 75 • MAP IOU 50_95 The default selected value is MAP IOU 50.
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set as 2.

Option	Description	Parameter
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set as 2.
Variant	It is the variant of the algorithms being used.	Select the variant of the algorithm to be used. The possible values are Nano, Small, Medium, Large, and Extra Large.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is SGD.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select False to disable the learning rate scheduler. The default value is set to True.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.

Option	Description	Parameter
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. • The default value is set to 0.001.
Momentum	Momentum is a method that helps to accelerate gradient vectors in the right directions, thus leading to faster convergence by accumulating the gradient of the past steps to determine the right direction.	Enter the appropriate number. • The default value is set to 0.937
Weight Decay	Weight decay for the optimizer.	Enter the suitable numeric value. • The default value is set to 0.0005.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. • The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. • The default value is set to 0.1.

Detectron2

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. • The default value is set to False.
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select False to not use pre-trained models in the training. • The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. • The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. • The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. • The default value is set to 20.
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. • The default value is set to 4.

Option	Description	Parameter
Evaluation Metrics	Metrics that are used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • MAP IOU 50 • MAP IOU 75 • MAP IOU 50_75 The default selected value is MAP IOU 50.
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set to 2.
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set to 2.
ROI IoU Threshold	IoU Threshold is used during the process of non-maximum suppression (NMS).	Enter the appropriate numeric value. The default value is set to 0.5.
RPN NMS Threshold	During NMS, if the confidence score of a bounding box is below this threshold, it might be suppressed, meaning it won't be considered as a final detection.	Enter the appropriate numeric value. The default value is set to 0.7.
Proposal Generator Min Size	The minimum size of the region proposal generator.	Enter the appropriate numeric value. The default value is set to 0.
Fpn Fuse Type	The method used for fusing (combining) features from different scales within the Feature Pyramid Networks (FPN).	Select the suitable value. The available values are sum, avg, max, concat. The default selected option is sum.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is AdamW.
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select True to enable the learning rate scheduler. The default value is set to False.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate value. The default value is set to 0.00005 when the hyperparameter tuning flag is True else 0.001.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

Table Transformer

Click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

- General Parameters
- Optimizer Parameters
- Model Layer Parameters
- Embedding Parameters

General Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Hyperparameters Tuning	This option is a flag to enable hyperparameter tuning to optimize the model based on different hyperparameter configurations.	Select True to enable hyperparameter tuning. The default value is set to False.
Pre - Trained Model	This option is a flag, If set to True then pre-trained models can be used to train the current model.	Select False to not use pre-trained models in the training. The default value is set to True.
Shuffle	This option is a flag, if set to True then data is shuffled during the training process.	Select True to shuffle the data. The default value is set to False.
Early Stopping	A technique where loss is monitored every epoch and training is stopped once loss is not decreasing. This option is a flag; if set to True, this technique becomes enabled.	Select True to check the training loss in each epoch. The default value is set to False.
Epochs	An epoch refers to one complete pass through the entire training dataset during the training of a model.	Enter the number of epochs to be used. The default value is set to 20.

Option	Description	Parameter
Batch Size	The batch size represents the number of data samples that are processed in one forward and backward pass through a neural network during training.	Enter the number of batches to be used. The default value is set to 4.
Evaluation Metrics	Metric that are used to evaluate the best hyperparameter configuration. This option is enabled only when the hyperparameter tuning flag is set to True.	Select the appropriate metrics as per your requirements. The possible options are: • Accuracy • Precision • Recall • F1-Score The default selected value is Accuracy.
Model Save Frequency	Models are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the models are saved. The default value is set as 2.
Metric Save Frequency	Model metrics are saved after this many alternative epoch iterations.	Enter the number of alternative epochs after which the model metrics are saved. The default value is set as 2.

Optimizer Parameters

The following table provides the parameters to set in each option under this set:

Option	Description	Parameter
Optimizer	Optimizers are optimization techniques to improve the model's performance by minimizing the loss functions.	Select the appropriate optimizer. The available options are: • AdamW • SGD • RMS Prop The default selected value is AdamW.

Option	Description	Parameter
Learning Rate Scheduler Flag	Flag to enable Learning Rate Scheduler used to change Learning Rate during training.	Select True to enable the learning rate scheduler. The default value is set to False.
Learning Rate Scheduler Type	This option is enabled only when the Learning Rate Scheduler Flag is set as True. This option sets the learning rate scheduler type.	Select the appropriate value. The available values are: • Step LR • Cyclic LR • Exponential LR • Polynomial LR • Reduce LR on Plateau The default selected value is Step LR.
Learning Rate	It defines the learning rate for the model training.	It is a value in fractions. Enter the appropriate. The default value is set to 0.005.
Step Size	It denotes the period of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 32.
Gamma	It denotes the multiplicative factor of learning rate decay. This option is available only when the Learning Rate Scheduler Flag is set to True.	Enter the appropriate number. The default value is set to 0.1.

Model Layer Parameters

The following table provides the parameters to set in each option under this set.

Option	Description	Parameter
Backbone Network	This is the Neural Network (NN) architecture serving as the backbone for the Table Transformer model. This option is available only when the Pre-Trained Model parameter under General parameters is set to False.	This is the dropdown field with the default value set to resnet18. The other possible dropdown value is resnet50.

Option	Description	Parameter
Number of Encoder Layers	Indicates the number of encoder layers in the Table Transformer model. Encoder layers process the input data and capture hierarchical features. This option is available only when the Pre-Trained Model parameter under General parameters is set to False.	Enter an integer value. • The default value is set to 6.
Number of Decoder Layers	Specifies the number of decoder layers in the model. Decoder layers generate the output sequence based on the encoded information from the input	Enter an integer value. • The default value is set to 6.
FeedForward Network Dimension	The dimensionality of the feedforward neural network within each transformer layer. The feedforward network processes information within each attention head. This option is available only when the Pre-Trained Model parameter under General parameters is set to False.	Enter an integer value. • The default value is set to 2048.
Hidden Layer Dimension	The dimensionality of the hidden layers in both the encoder and decoder. These layers play a crucial role in learning complex representations of the input data. This option is available only when the Pre-Trained Model parameter under General parameters is set to False.	Enter an integer value. • The default value is set to 256.
Number Of Attention Heads	Specifies the number of attention heads in the multi-head attention mechanism. A higher number allows the model to focus on different aspects of the input simultaneously. This option is available only when the Pre-Trained Model parameter under General parameters is set to False.	Enter an integer value. • The default value is set to 8.
Dropout	It is a regularization technique. This parameter sets the probability of dropping out (ignoring) a unit during training, preventing overfitting and promoting model generalization.	Enter the appropriate numeric value between 0 and 1. • The default value is set to 0.1.

Option	Description	Parameter
Number of Queries	Number of queries in the attention mechanism. Queries are used to determine which parts of the input sequence are most relevant for generating the output. This option is available only when the Pre-Trained Model parameter under General parameters is set to False.	Enter an integer value. The default value is set to 125.

Embedding Parameters

These are the following parameters under this set is:

- **Positional Embedding Type** — Indicates the type of position embeddings used in the model. Position embeddings convey information about the position of elements in the input sequence. The other available option is 'learned'. This option is available only when the Pre-Trained Model parameter under General parameters is set to False.

  The default selected value is sine.

- **Embedding Dimension** — It is the dimensionality of the embeddings used in the model. Embeddings are the vector representations of tokens in the input sequence, and this parameter determines their size. Enter the appropriate integer value.

  The default value is set to 64.

Machine Learning

Machine Learning models like SVM are supervised machine learning algorithms that classify data by finding an optimal decision boundary in high-dimensional space, predicting classes based on their text embedding vectors, which capture semantic information about the image content.

This category comprises the following algorithm:

- SVM (Support Vector Machine)

Support Vector Machine

The SVM (Support Vector Machine) is a supervised learning algorithm and is particularly effective for classification tasks in high-dimensional spaces.

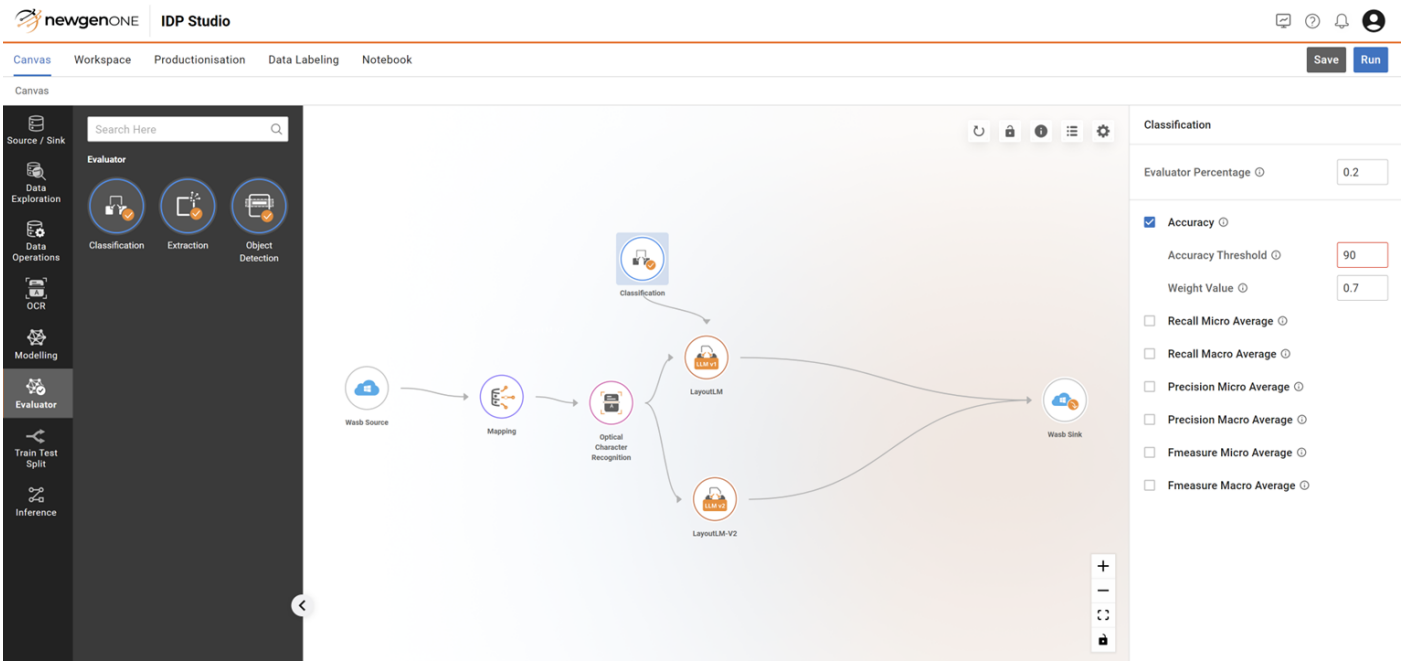
The following table provides the parameters to set in each option under this category:

Option	Description	Parameter
Kernel	Specifies the algorithm's kernel type.	Select the appropriate kernel. The available options are Linear, Polynomial, RBF, and Sigmoid. The default selected value is rbf.
Tolerance	It defines the tolerance for stopping criterion.	Enter the appropriate numerical value. The default value is set to 0.0003.
Max Iteration	It defines the maximum number of iterations.	Enter the appropriate integer value. The default value is set to 1.
Regularization	It defines the regularization value.	Enter the appropriate integer value. The default value is set to 1.

Evaluator

The evaluator component measures the performance or effectiveness of a trained model. Nodes under this category are usually used when you add multiple algorithm models in the pipeline to evaluate the best preconforming one among these models.

The evaluator node is applied to any one of the multiple models.



In the pipeline, the Document Classification evaluator determines the best model between LayoutLM and LayoutLM-V2, both belonging to the Document Classification category.

This option under the canvas keeps the nodes related to modelling algorithms.

It is divided into the following types of algorithms, each containing the appropriate cluster of nodes:

- Classification
- Extraction
- Object Detection

To access the parameters for the above nodes, click the node after adding it on the canvas, and the parameter options appear on the right side of the screen. These parameters are grouped into the following sets:

Each of the evaluator nodes has the following common properties:

- **Evaluator Percentage** — It is the percentage change threshold after which the updated model gets saved in production. Enter the appropriate evaluator percentage value between 0 and 1.

  The default value is set to 0.2.

The rest of the parameter options for classification and extraction evaluator are given in the following table:

Option	Description	Parameters
Accuracy	If this option is selected then the model is evaluated based on accuracy.	Enter the threshold value below which the model does not get saved. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.
Precision Micro Average	If this option is selected then the model is evaluated based on the precision micro average.	Enter the threshold value below which the model does not get saved. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.
Precision Macro Average	If this option is selected then the model is evaluated based on the precision macro average.	Enter the threshold value below which the model does not get saved. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.

Option	Description	Parameters
Recall Micro Average	If this option is selected then the model is evaluated based on the recall micro average.	Enter the threshold value below which the model does not get saved. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.
Recall Macro Average	If this option is selected then the model is evaluated based on the recall macro average.	Enter the threshold value below which the model does not get saved. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.
Fmeasure Micro Average	If this option is selected then the model is evaluated based on the Fmeasure micro average.	Enter the threshold value below which the model does not get saved. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.
Fmeasure Macro Average	If this option is selected then the model is evaluated based on the Fmeasure Macro Average.	Enter the threshold value below which the model does not get saved. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.

The parameters for the object detection evaluator option are given in the following table.

Option	Description	Parameter
MAP IOU 50	If this option is selected then the model is evaluated based on MAP IOU 50 value.	Enter the MAP IOU 50 value. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.
MAP IOU 75	If this option is selected then the model is evaluated based on the MAP IOU 75 value.	Enter the MAP IOU 75 value. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.
MAP IOU 50-95	If this option is selected then the model is evaluated based on MAP IOU 50-95 value.	Enter the MAP IOU 50-95 value. Enter the percentage weightage (in float), which must be divided accordingly if more than one model is used. The default value is set to 0.

Inference

Model inference refers to the process of using a trained machine learning or deep learning model to make predictions or decisions on new, unseen data.

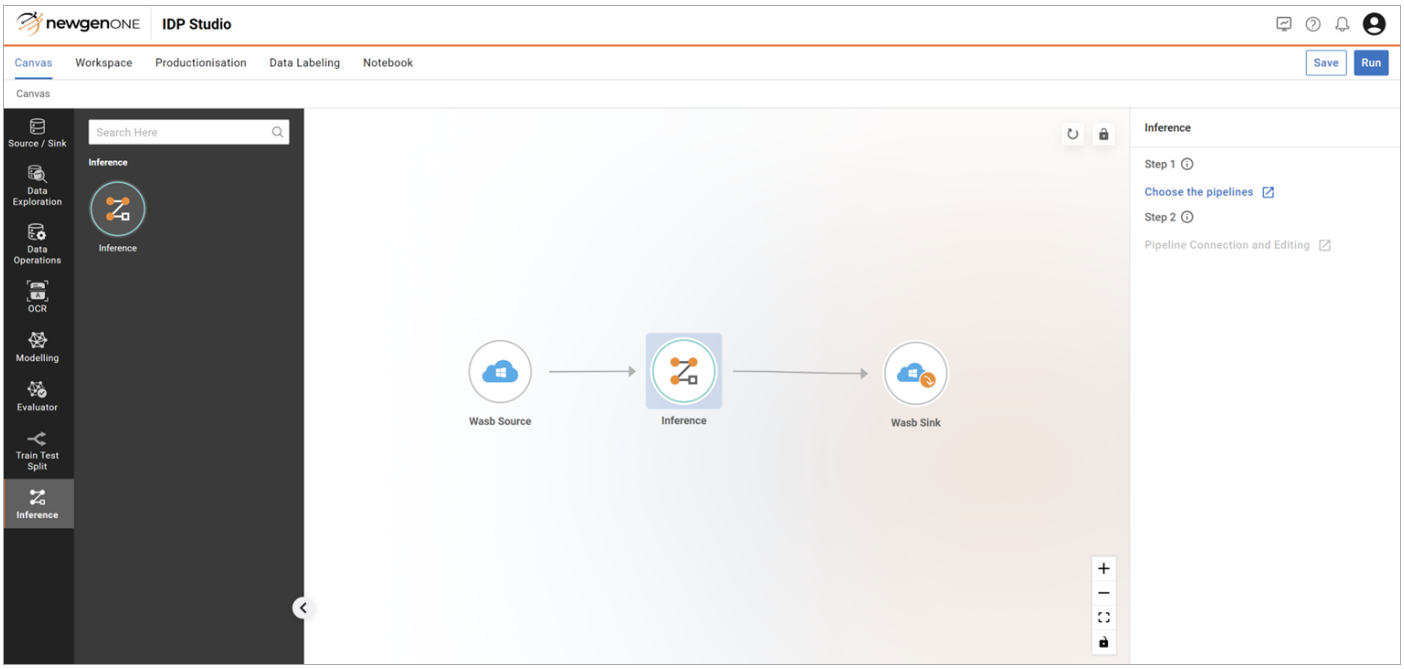
Creating Inference pipeline

To create an inference pipeline, perform the following steps:

1. Connect required nodes on canvas area
2. Setup source data
3. Setup sink location
4. Setup inference node

Connect required nodes on canvas area

1. From the Canvas menu's Database section, drag and drop the required source and sink nodes onto the canvas area.
2. From the Canvas menu's Inference section, drag and drop the inference node onto the canvas area.
3. Connect from the source node to the inference node and from the inference node to the sink node.



Setup source data

To set up source data refer Reading data source.

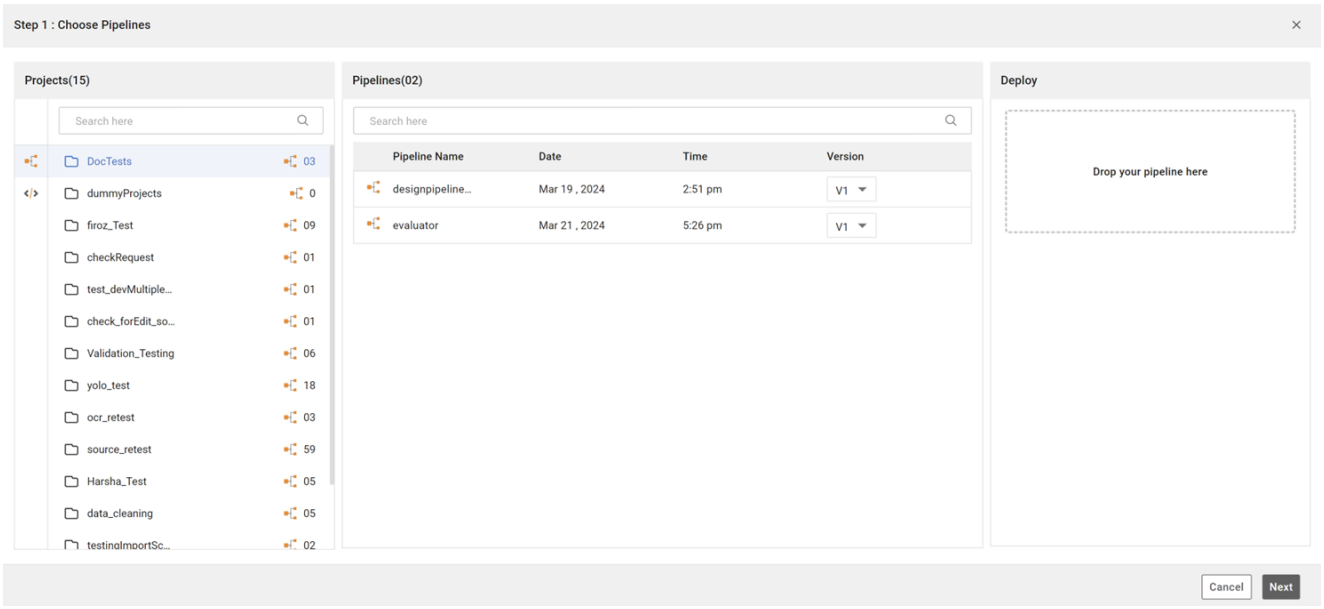
Setup sink location

To set up the sink location refer Sinking data.

Setup inference node

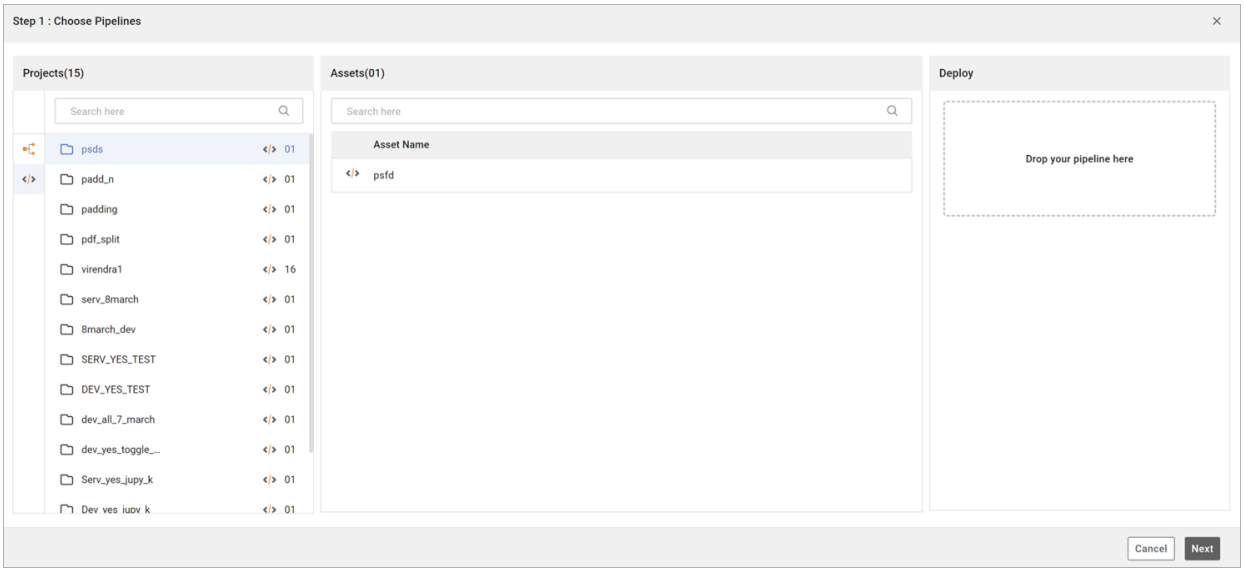
To setup the inference node, perform the following steps:

1. Click the **inference** node on the canvas area.
2. Click the link **Choose the pipelines** under Step 1 on the dialog that appears on the right side of the canvas area. The following screen appears:

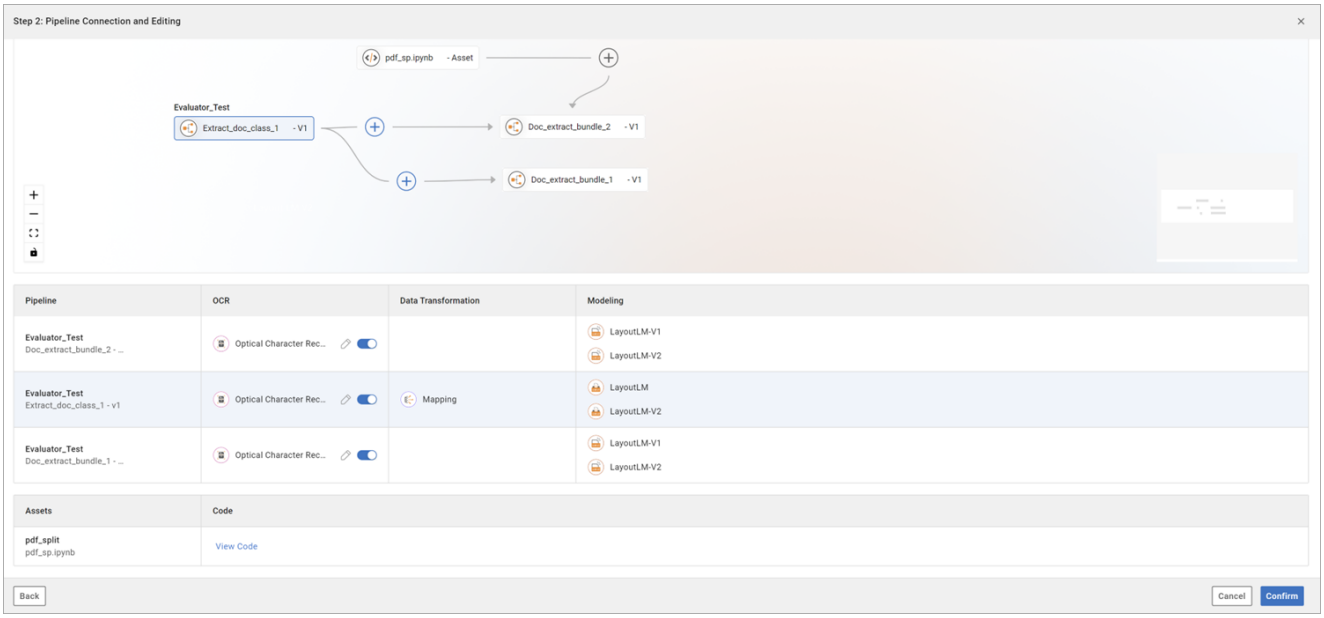


3. Search for your project and pipeline using the search box. Select the version of the pipeline you want to do the inference on.
4. Drag and drop the pipeline into the box given under the deploy section on the right side of the screen.
5. If you need to run some custom Python code before or after the pipeline during inference, then do the following:
    a. Click the icon </> on the left side of the project lists. It shows the list of projects and assets for the jupyter notebook. These projects and assets were created under the Notebook menu tab which is available on the upper side of the landing page of the IDP Studio.

b. Search for your project and asset name using the search boxes and drag and drop the required asset into the box given under the deploy section on the right side of the screen.

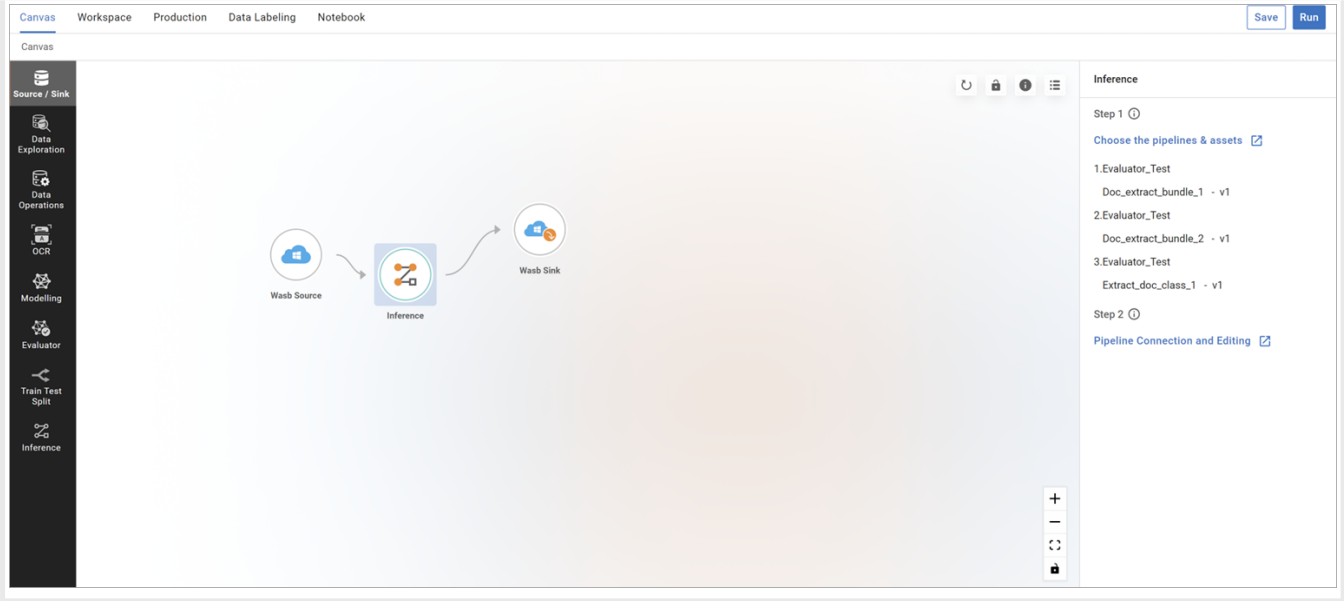


6. Click the **Next** button. The Step 2: Pipeline Connection and Editing dialog appears:



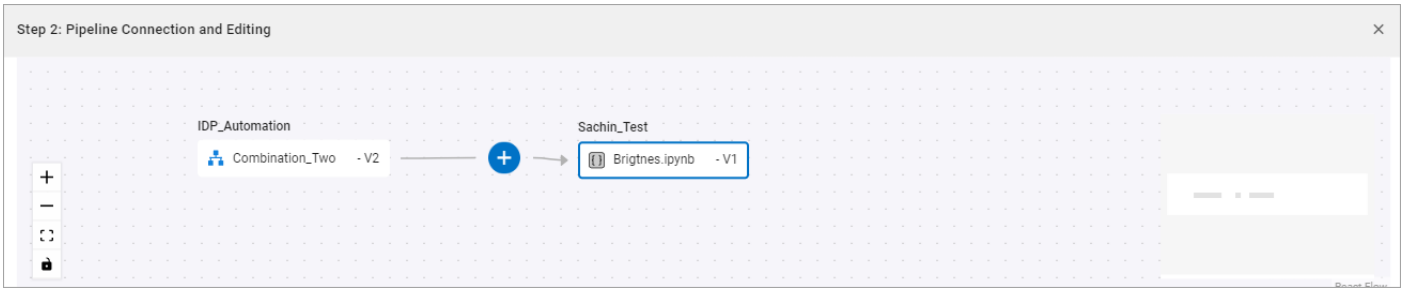
7. Configure the following levels:
   a. Setup on canvas area
   b. Model Inference Table

You can directly go to Step 1 or Step 2 of an existing inference pipeline by clicking the respective link from the Inference setup dialog that appears on the right side of the canvas area when you click the Inference node of an Inference pipeline.



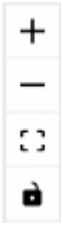
Setup on canvas area

The selected pipeline and asset appear here on a canvas inside the dialog. You can connect the pipeline and the jupyter notebook asset with a connector. For example, if you want to execute the Jupyter Notebook after running the pipeline, you must use the connector from the pipeline node to the asset node on the inference canvas area.



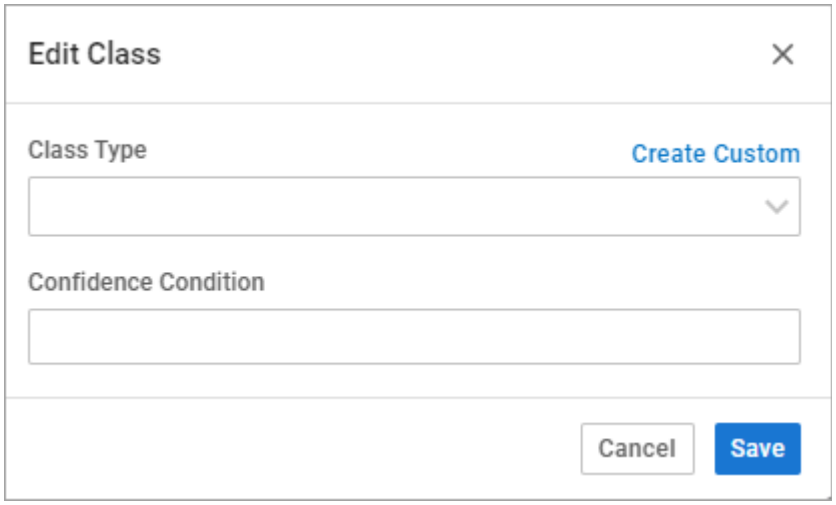
On the other hand, if you needed to run the Jupyter Notebook before executing the pipeline, you must have used the connector from the asset node to the pipeline node on the inference canvas area.

You can zoom in, zoom out, fit to view, or lock the nodes with connectors by clicking toggle interactively from the options present at the left of the canvas area here. Every connector has the option to add image mapping classes.



To edit a class, perform the following steps:

1. The plus icon allows you to select or create image classes used during inference.
2. Clicking the plus icon opens the Edit Class dialog:



If your pipeline doesn't have the mapping node, the Class Type dropdown menu is empty; otherwise, it displays the classes from your mapping node.

3. To create a new custom class type, click the **Create Custom** link, enter the name of the new class type, and then click **Save**.
4. The custom class gets added and appears in the dropdown menu. You usually create a custom class when you've implemented logic in your Jupyter Notebook to categorize certain documents under this class.
5. Select one or more class types from the dropdown menu and assign the **Confidence Condition**.
6. Click **Save** to save the class mapping. This completes the setup on the canvas area.

Understanding confidence condition

In the Confidence Condition, enter the condition. If it holds against the actual confidence score received on the test data for any of the selected classes, the predicted data moves forward to the next pipeline or asset node. Otherwise, execution stops.

For example, when you select the classes email and cheque from the dropdown menu, the confidence condition is > 0.23, < 0.23, >= 0.23, or <= 0.23, applying to both classes. Typically, the condition is > n, < n, >= n, or <= n, where n represents the confidence score, ranging between 0 and 1.

Understanding a use case

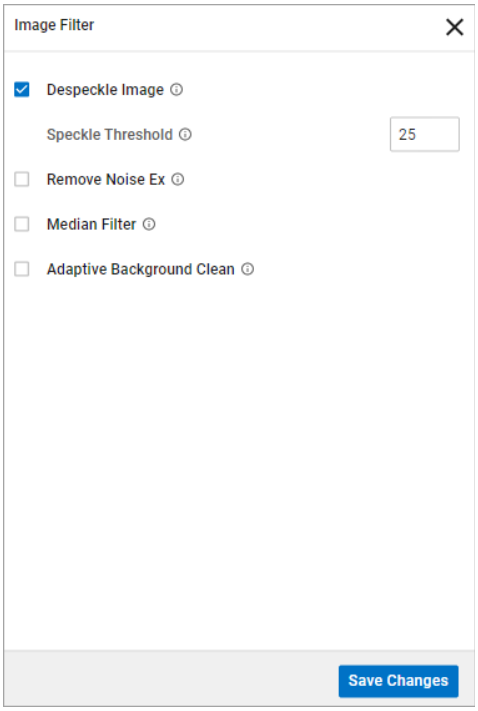
Now, let's say you set the confidence condition as >0.23 and have chosen email and cheque as the classes from the dropdown. If the first model development pipeline node predicts that the testing data belongs to the class email with a confidence score of 0.50. Then as the predicted class (email) is one of the classes you have set the confidence condition for, and the confidence score you had in the confidence condition was 0.23, which is less than the predicted score (0.50), the predicted data is passed to the next pipeline or asset node.

On the other hand, let's say you had used the confidence score as 0.70 in the confidence condition for classes email and cheque. However, if the class predicted by the model pipeline is an email with a confidence score of 0.50, then the predicted data is not passed to the next pipeline or asset node. In this case, the execution stops for the rest of the nodes.

Model inference table

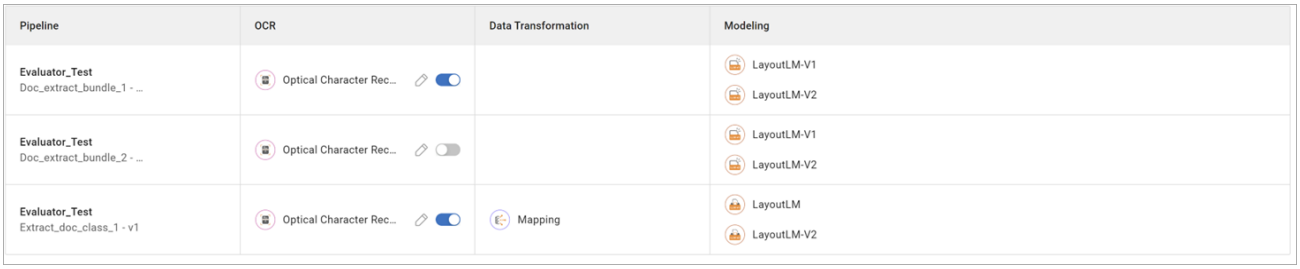
At the bottom table with the name Model Inference to the canvas, you see the following details of the model pipeline:

- Pipeline name and version along with its project name in the first column of the table.
- From the second column onwards, you see the various nodes available in the pipeline. Parameters of some of these nodes can be edited by clicking the icon  . For example, if you need to edit the parameters of the Image Filter, click the icon  against it. The Image Filter dialog appears:



Update the parameters as required and click **Save Changes** to save the updated values.

- If you want to exclude a node from your pipeline, then toggle with the option . In the current example, if you need to exclude one of the OCR nodes from the original pipeline, click the option next to it to deactivate it in the pipeline. In this case, the node is bypassed during pipeline execution in inference.



- You can't edit or toggle the following node categories: Mapping (mapping node only), Modelling, Operations (Data Augment nodes only), and Evaluator.
- You cannot edit the Mapping (Notebook node only) node category, but you can toggle it.

Appendix

The following table lists the constraints for the IDP Studio:

S. No.	Platform Element	Constraints
1.	Source Nodes	This must be the first node in any pipeline. No node can be connected before this node.
		Multiple nodes are not yet supported. Only single nodes are allowed on the canvas.
2.	Sink Nodes	This must be the last node of the pipeline.
		The sink node should be connected to every pipeline except (Src -> Dex).
		Only a single Sink node is allowed in the canvas.
3.	Data Exploration Nodes	Only a single Data Exploration node is allowed in the canvas.
		After Dex Exploration, no next node is required. Dex Exploration serves as the final node, so nothing connects to it afterward.
		Both Dex and Dex action cannot be present on the canvas at the same time.
4.	Data Augment Nodes	Only a single Data Augment node is allowed in the canvas.
		Image Classification Augment only connects to Document Classification, Image Classification, and Machine Learning(SVM). This is when both the argument and model are present in the pipeline.
		Document Extraction Augment only connects to Document Extraction algorithms but not for geolayouts. This is when both argument and model are present in the pipeline.
		Object Detection Augment only connects to Object Detection algorithms but not to Table Transformer and Table Detection. This is when both argument and model are present in the pipeline.
5.	Mapping Nodes	Must be attached with Sink only.
		Remove Max features from TFIDF.

S. No.	Platform Element	Constraints
		The SVM node can only connect from either the TFIDF/FastText node or the evaluator.
		Doc Class and Doc Extract both require OCR, so the pipeline must contain OCR when these models are present.
6.	Train Test Split Node	Can connect to augment or model node only.
		Only a single Train Test Split node is allowed in the canvas.
		No node can connect before this node, is considered a special node.
7.	Modelling Nodes	Must be attached only to the Sink node.
		Remove Max features from TFIDF.
		SVM node can only connect to either TFIDF/ FastText node.
		Doc Class and Doc Extract both require OCR, so the pipeline must contain OCR when these models are present.
		Attach Evaluator when two or more than two models are connected in parallel.
		The modelling pipeline can only go in serving.
		The model pipeline cannot go into data labelling for annotations.
8.	Get API / View API	Modelling or Dev pipeline which is run successfully and its result present in ES.
		Not needed for Serving or Inference engine pipeline with only a single Dev pipeline and no other Dev node or Jupyter.
		Needed for Serving or Inference engine pipeline with multiple combinations of Dev and Jupiter: Bundle API
9.	Jupyter notebook Assets	Only Deployed assets must be shown in Jupyter NoteBook.
		Endpoints cannot be the same.
		Must be shared within users.
		Deployed assets which are being used must give a warning before deleting.
		Pipeline, where assets are deleted, must show pipeline but cannot run due to deleted asset. Some warning in UI or logs: Asset not found.

S. No.	Platform Element	Constraints
10.	Inference	For the inference pipeline, annotation flag, and model monitoring must not be present.
		Only the source or notebook can connect before the model inference node.
		Only the sink or notebook can connect after the model inference node.
		An annotation flag must not be displayed in the case of an inference pipeline.
		For the engine pipeline, the User must have to select at least one Dev pipeline in the engine node.
11.	Evaluator	The evaluator node connect only to model nodes.

Glossary

Accuracy

Accuracy is determined by dividing the proportion of accurate predictions by the overall number of predictions. It provides an overall assessment of model correctness but might not be suitable for imbalanced datasets.

Precision

Precision is a crucial metric that assesses the ratio of accurate positive predictions among all model-generated positive predictions. This metric is required in scenarios where the consequences of false positives can be expensive.

Recall

Recall measures the ratio of accurate positive predictions among all actual positive instances present in the dataset.

F1-Score

The F1 metric F1 score is required to assess the overall performance of a classification model during the imbalance between the classes. It provides a balance between precision (ability to minimize false positives) and recall (ability to capture true positives) of a classification model. It is commonly used in binary classification tasks where there are two classes (positive and negative).

Weighted metrics

Weighted Metrics include F1, Precision, and Recall. Each class is assigned a weight based on its importance or prevalence in the dataset. These weights reflect the relative significance of each class. Then, it calculates the weighted average of the respective metric.

Standard Deviation

The standard deviation is a numeric value that measures the spread of data around its mean. A lower standard deviation value indicates that most of the data is closer to the mean, while a larger value indicates that the data is widely spread.

Skewness

Skewness is a statistical measure that describes the asymmetry of the probability distribution of a real-valued random variable about its mean.

Image Channel

An image channel refers to a component of an image that represents the intensity values of pixels along a particular dimension or property. For example, In a grayscale image, there is only one channel. The intensity value ranges from 0 (black) to 255 (white), with shades of gray in between. In an RGB (Red, Green, Blue) image, there are three channels: one for each color component (red, green, blue). Each channel represents the intensity of its corresponding color at each pixel.

Learning rate

Learning rate is a parameter required for model optimization. It controls the step size or rate at which a model's parameters are updated during training.

Mean

The mean is the arithmetic average of a set of numbers, calculated by summing all the numbers in the dataset and then dividing by the count of numbers.

Median

The median is the middle number in a sorted dataset. If the dataset has an odd number of values, the median is the middle number. If the dataset has an even number of values, the median is the average of the two middle numbers.

Mode

The mode is the value that appears most frequently in a dataset. A dataset can have one mode, more than one mode, or no mode at all.

ReLU

ReLU stands for Rectified Linear Unit, and it is one of the most commonly used activation functions in neural networks, particularly in deep learning models. It is a simple yet powerful function that introduces non-linearity into the network, allowing it to learn complex patterns and relationships in the data. The ReLU. ReLU function returns the input value if it is positive and returns zero otherwise.

GeLU

GeLU stands for Gaussian Error Linear Unit, is an activation function that aims to capture non-linearities in neural networks while maintaining some of the desirable properties of linear units. It was introduced as an activation function in deep learning models, particularly in transformer architectures like BERT (Bidirectional Encoder Representations from Transformers).In simpler terms, GeLU is a smooth approximation of the rectifier (ReLU) function and the Gaussian cumulative distribution function (CDF).

Optimizer

An optimizer is an algorithm or method used to adjust the model parameters (such as weights and biases) to minimize the error or loss function during training. The goal of optimization is to find the set of parameters that best fit the training data and generalize well to unseen data.

SGD

SGD, which stands for Stochastic Gradient Descent, is one of the fundamental optimization algorithms. It is a variant of the standard gradient descent algorithm and is particularly well-suited for large-scale and high-dimensional optimization problems.

AdamW

AdamW is an optimization algorithm derived from Adam (Adaptive Moment Estimation) optimizer. AdamW introduces a weight decay term to Adam, which helps in stabilizing training and improving generalization performance.

RMS Prop

RMS Prop, which stands for Root Mean Square Propagation, is an adaptive learning rate method that adjusts the learning rate for each parameter based on the magnitude of the gradients.

Stop Words

These are the commonly occurring words in a language. These words are usually removed from the text corpus before processing or analysis. Examples of stop words in English include "the," "is," "and," "of," "in," "to," etc.