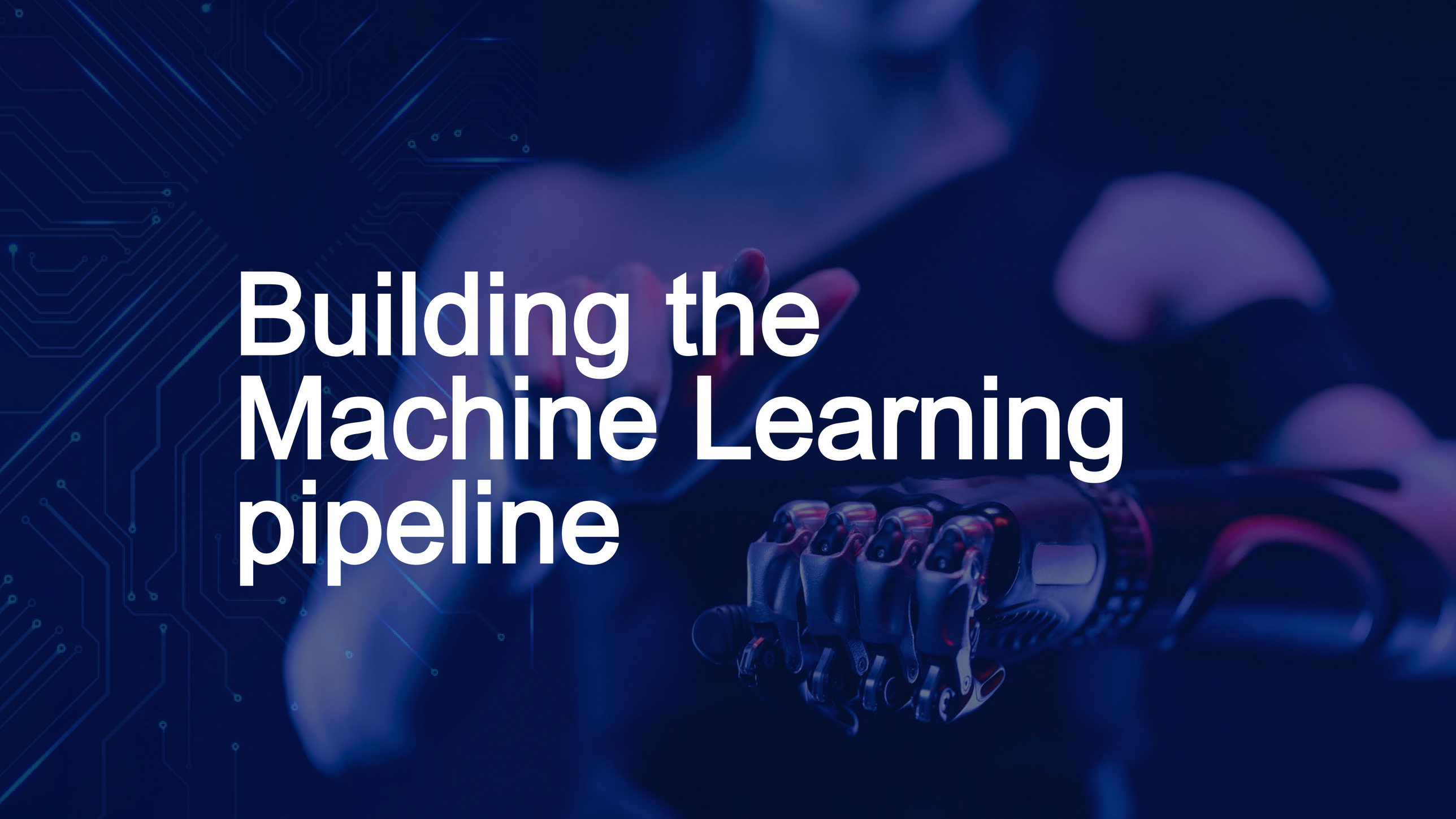
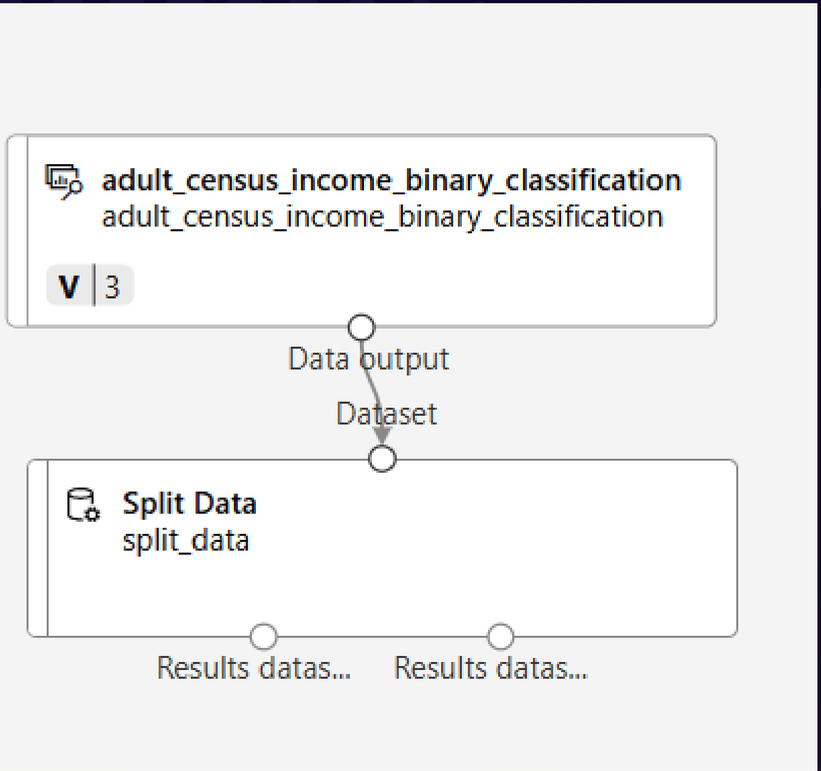
Building the Machine Learning pipeline



Split data

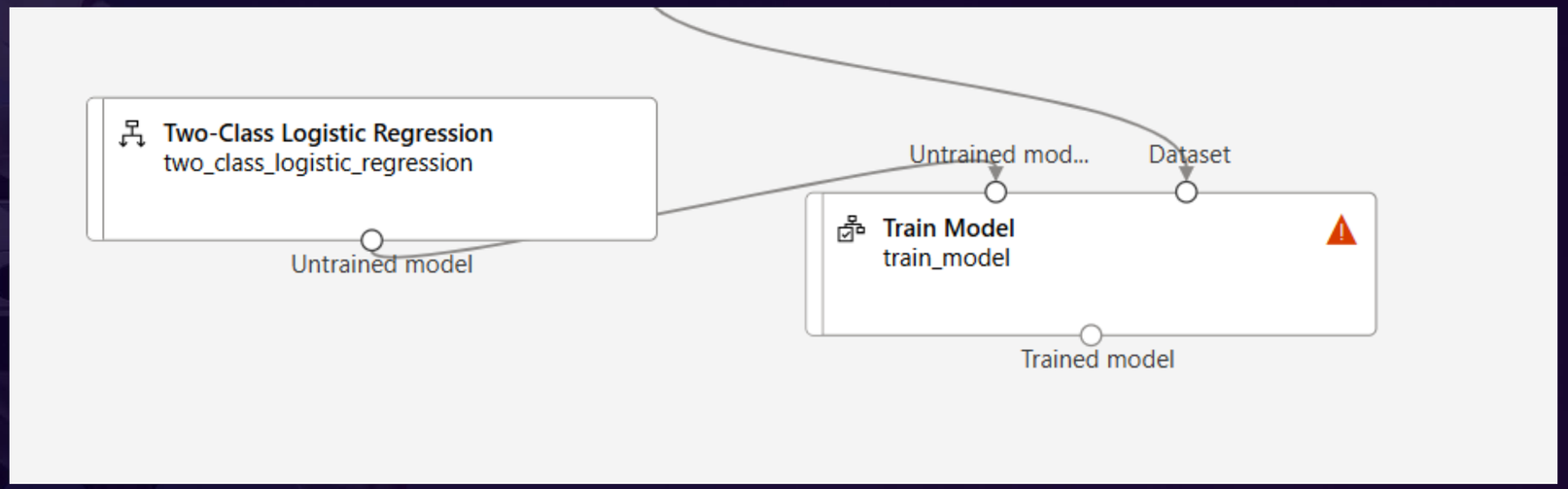


- Our aim is to build a machine learning model. This model will be used to classify the income of individuals into 2 classes - <=50K or > 50K

- The first step was to get the data set and split the set into a training and test set. The training set would have 70% of the data and the test set would have 30% of the data.

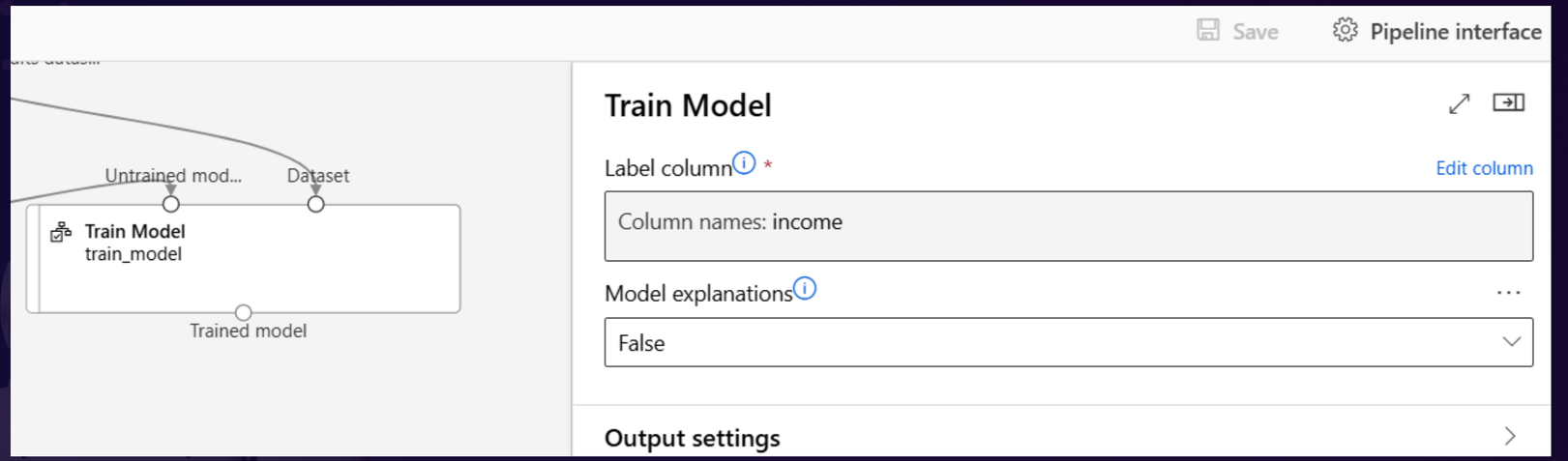
Train Model

- Next we need to add the component to train the model. We will add as the input the 70% of the split data phase.
- We will add the algorithm which will be used to train the machine learning model.



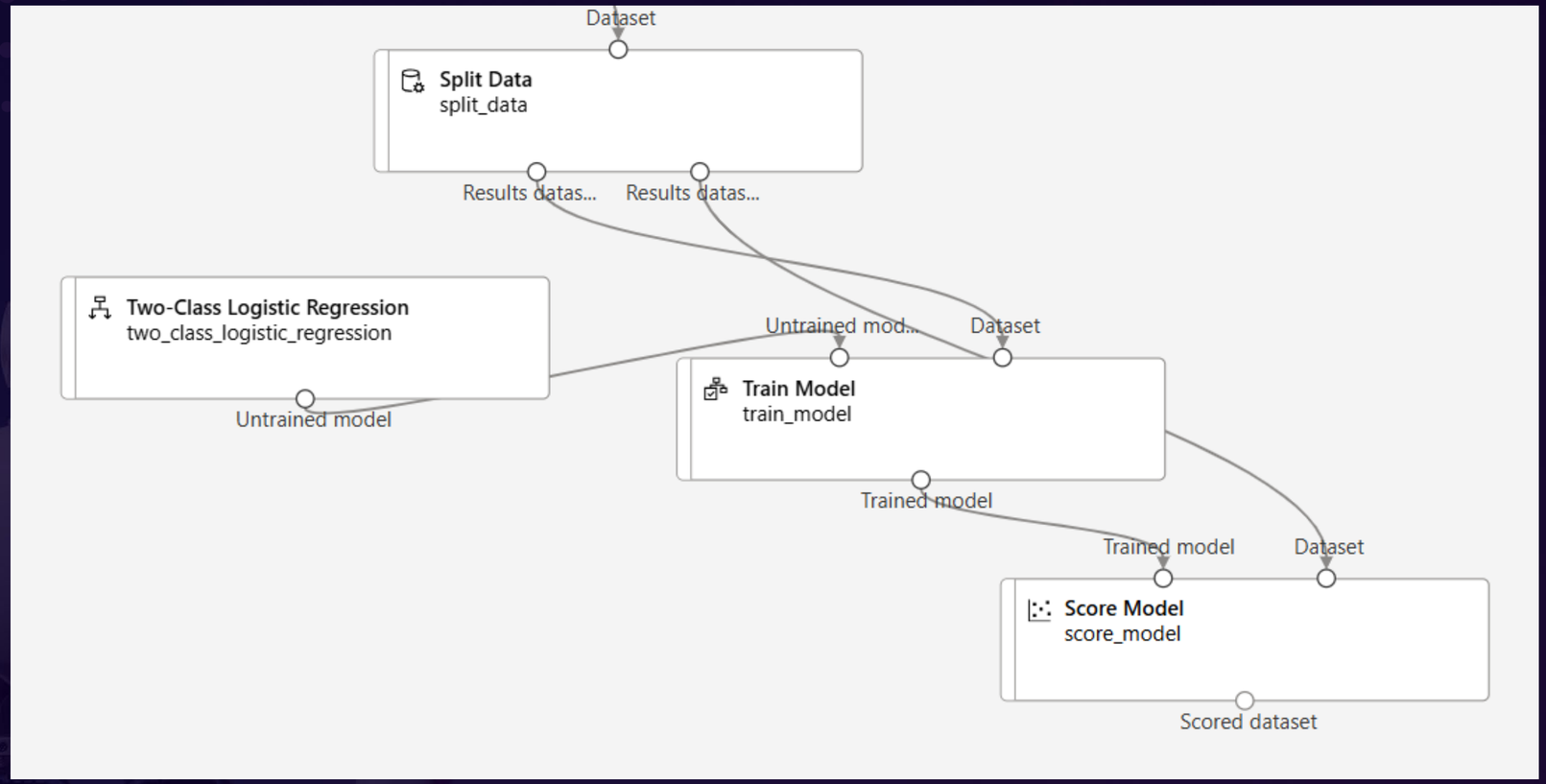
Train Model

- We will also ensure to train the model based on the income column. This is our label.



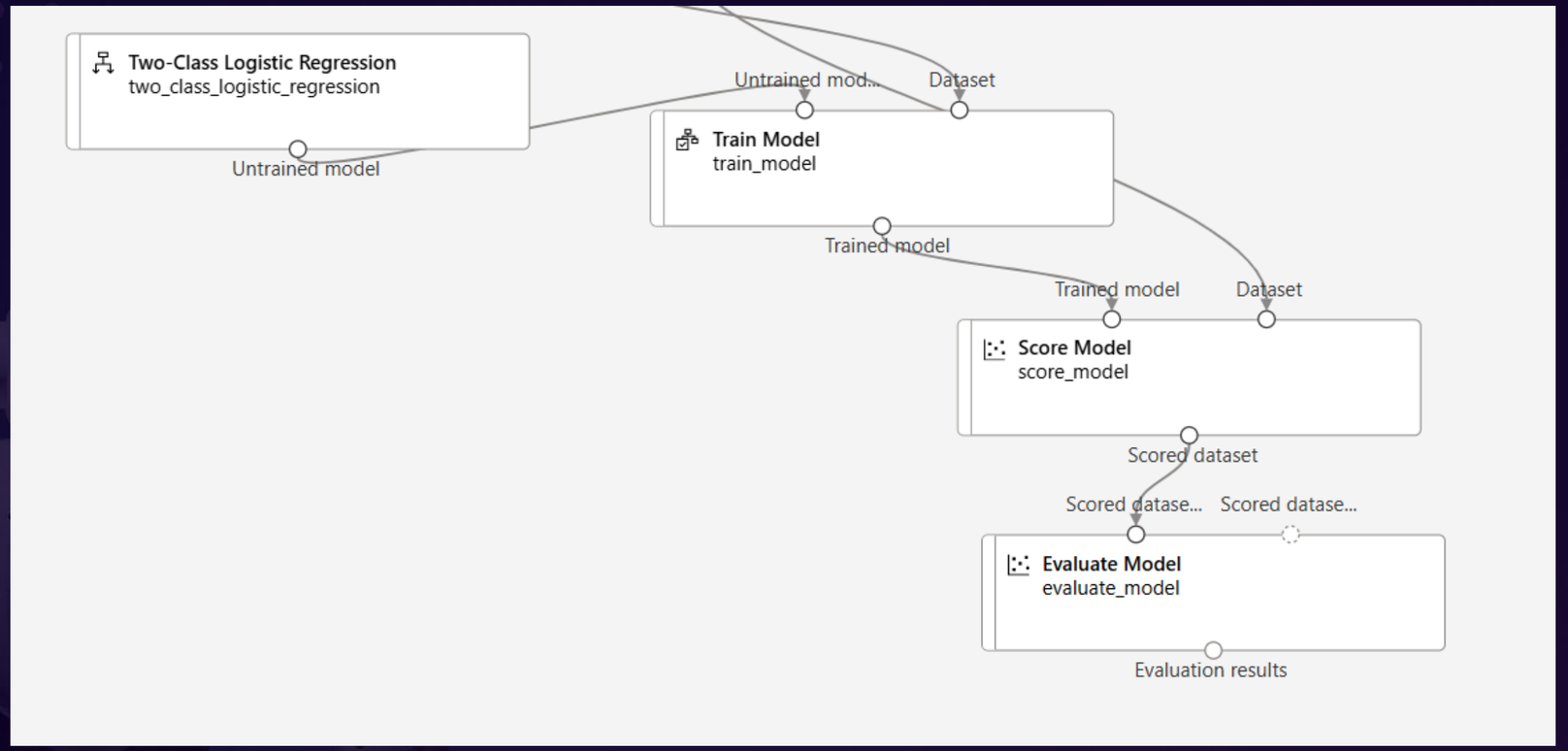
Score Model

- We can now score the model based on the 30% test data. This will help us understand how efficient the model is when it comes to predicting the income value.

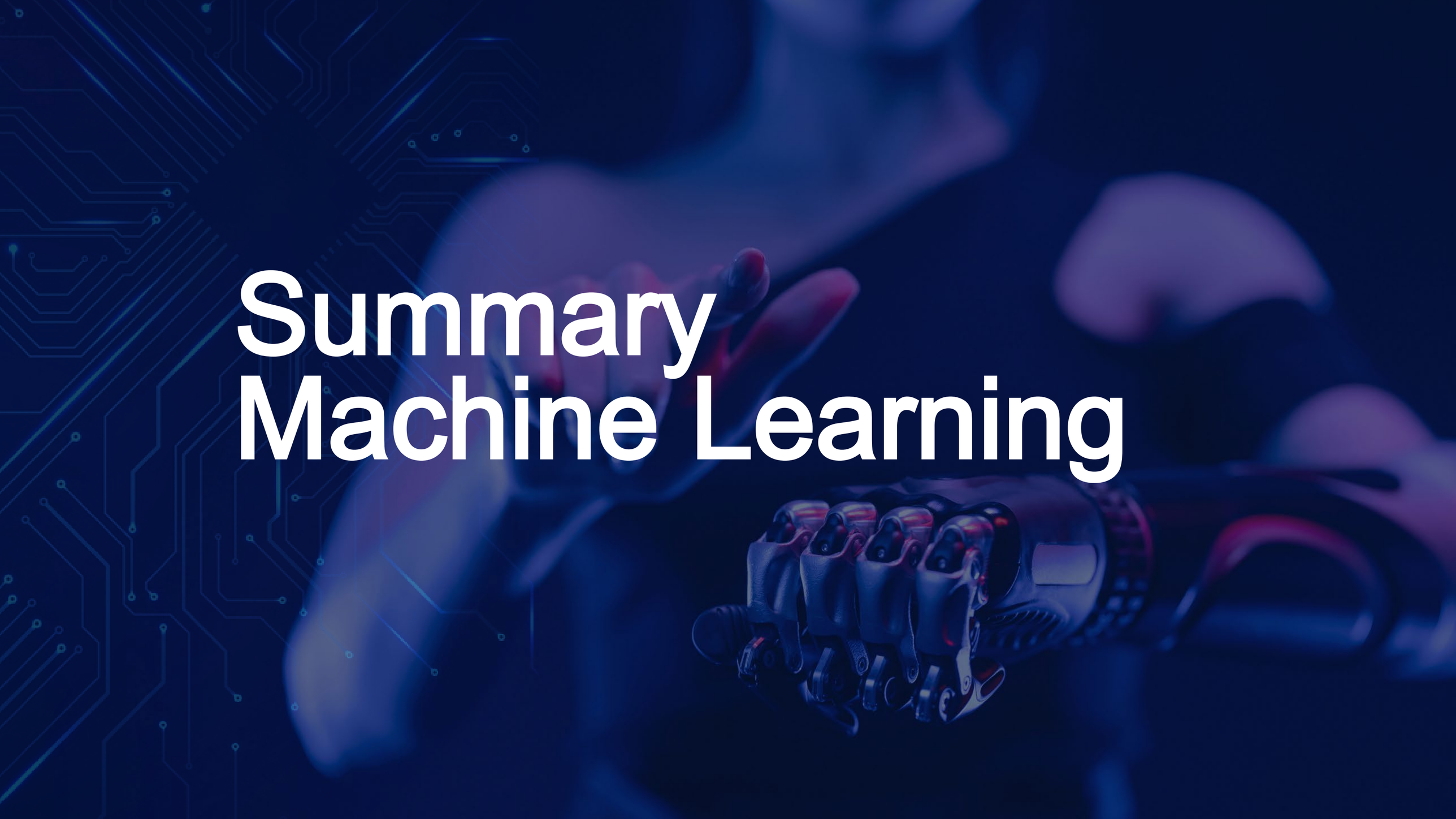


Evaluate Model

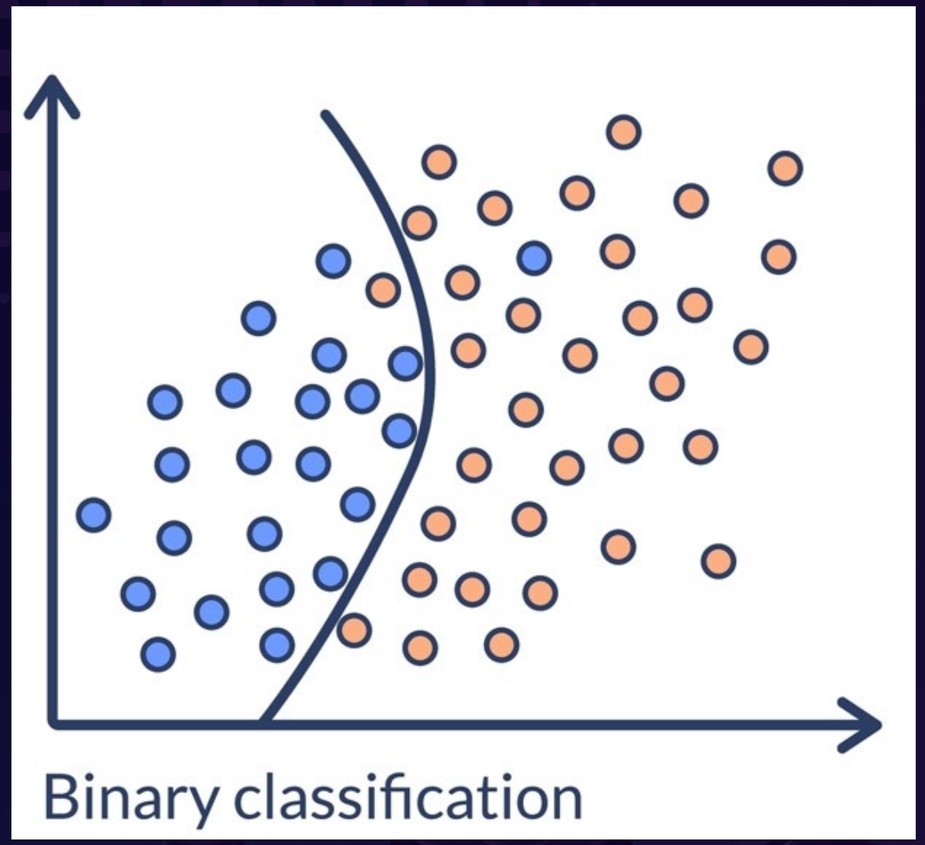
- Based on the scoring and the test data, we can finally evaluate the model.



Summary
Machine Learning



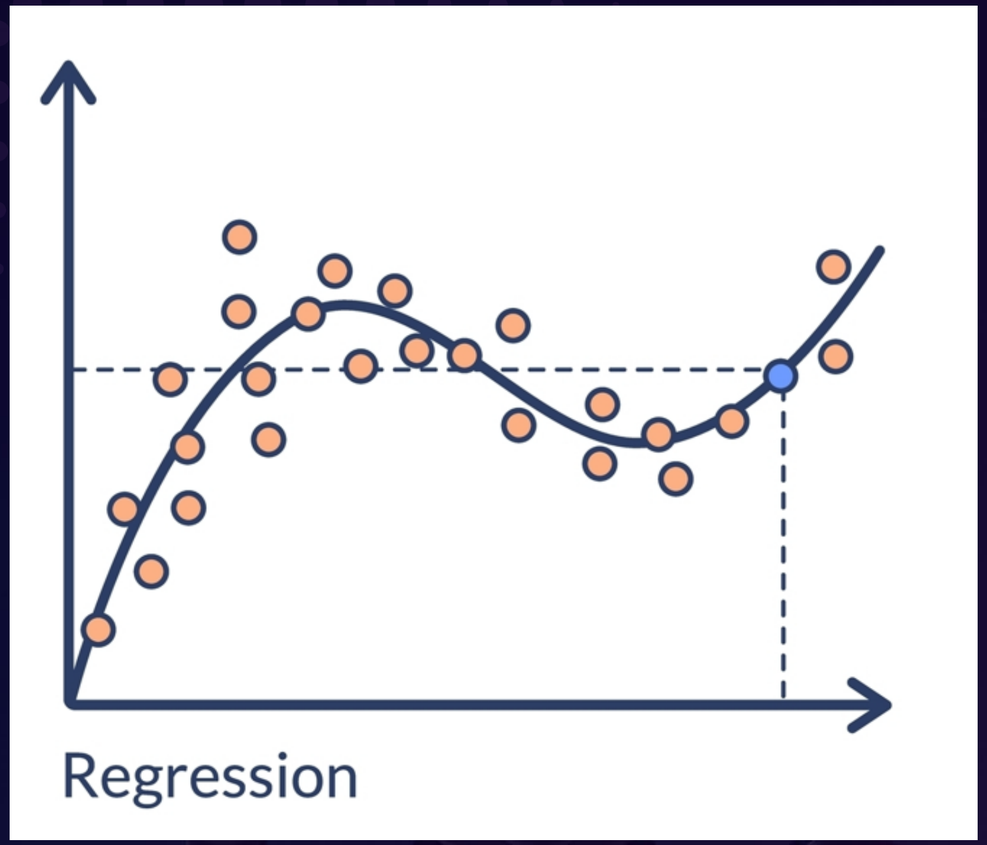
Machine Learning Algorithms



Classification Algorithms - Here we can classify our data into categories. We can have a two-class or binary classification. Example - Is the user going to buy the product if it is launched on our site?

Or we can make user of multi-class classification algorihtms.

Machine Learning Algorithms



Regression Algorithms - Here we want to predict a set of values based on historical data. Example - What would the sales for the various products for the upcoming month.

Machine Learning Algorithms

Clustering Algorithms - Here we want to segregate our data values into different clusters. For example if we want to segregate customers into different tiers for a loyalty program.

Features

- In training your machine learning model ,we first need to feed in training data.
- In supervised learning , the data will contain features and labels.
- The labels are the values that you want to predict.
- Feature selection - Here you choose relevant features when training a model.
- Feature engineering - Here you create new features based on raw data.

Classification Model Evaluation

01 Accuracy

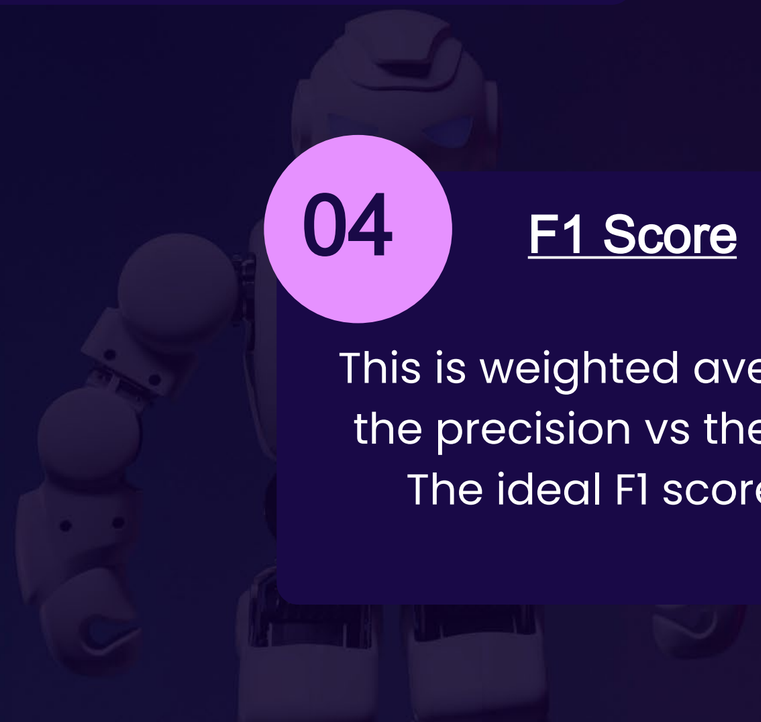
This measures the goodness of a classification model.

02 Precision

This measures the proportion of the true results over the positive results.

03 Recall

This is the fraction of the total amount of relevant cases that were actually retrieved.



04 F1 Score

This is weighted average of the precision vs the recall. The ideal F1 score is 1.

05 AUC

This is the area under the curve that shows the true positives and the false positives.

Regression Model Evaluation

01

Mean absolute error

This is used to measure how close the predictions are to the outcomes. Its ideal to have a lower score.

02

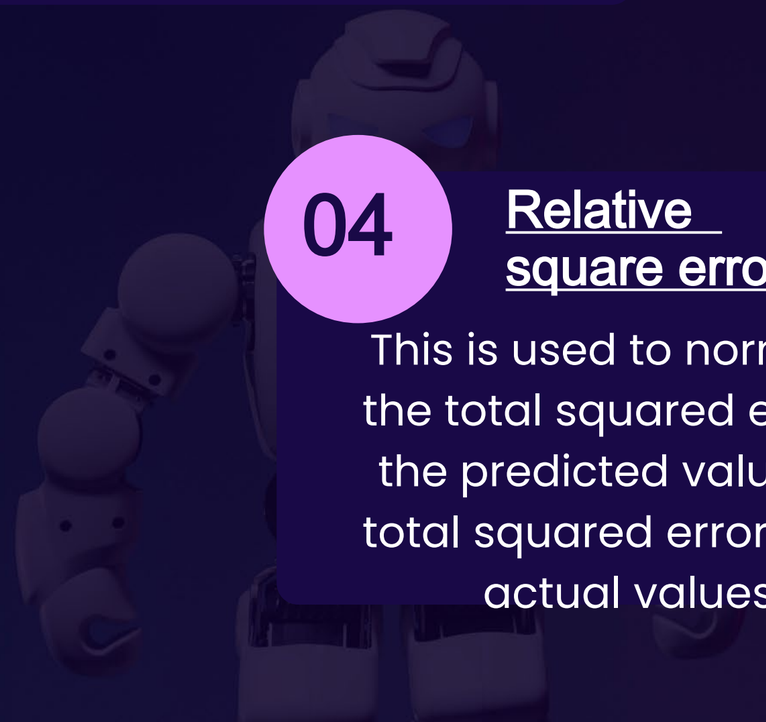
Root mean squared error

This just helps to summarize the error in the model.

03

Relative absolute error

This is the absolute difference between the expected and the actual values.



04

Relative square error

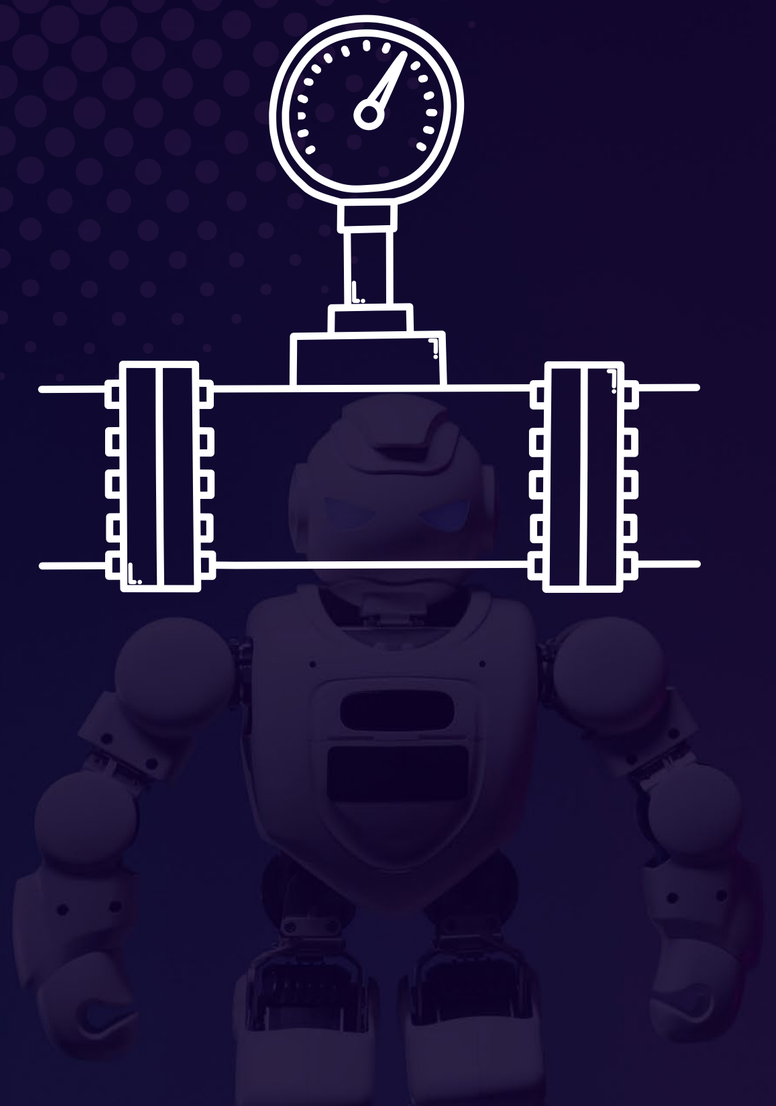
This is used to normalize the total squared error of the predicted values by total squared error of the actual values.

05

Coefficient of determination

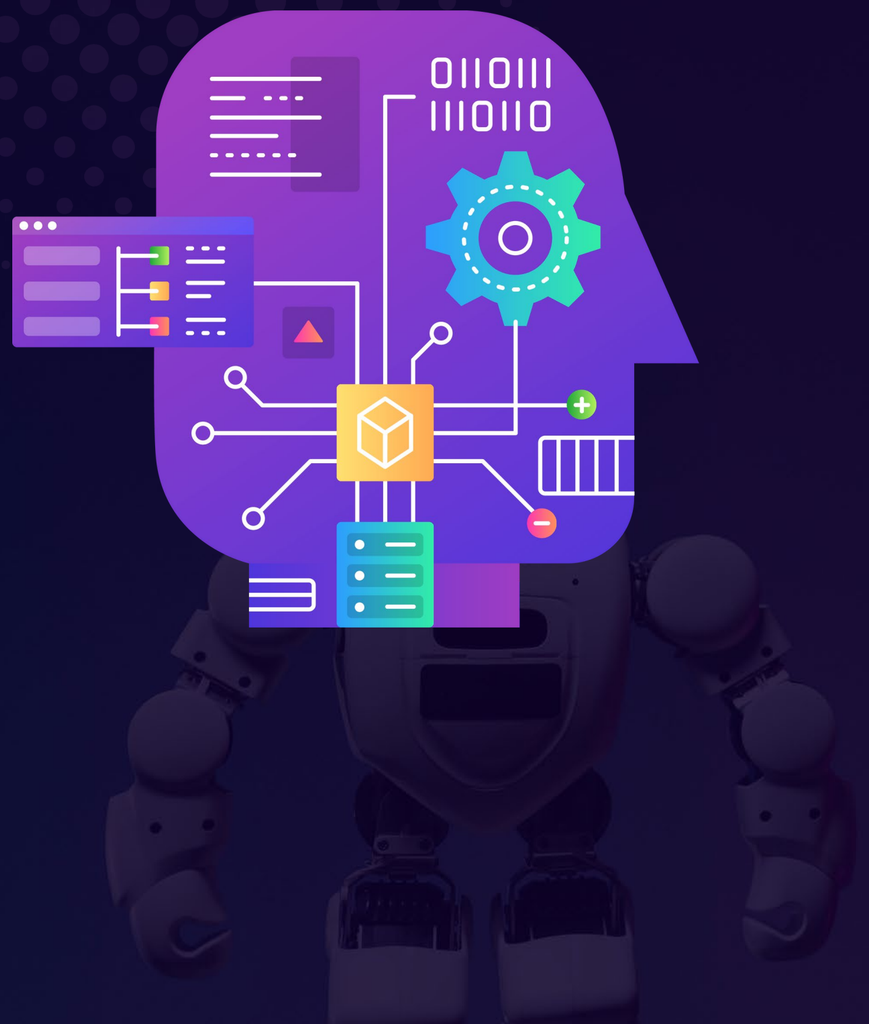
This is just used to represent the predictive power of the model between 0 and 1.

Pipeline



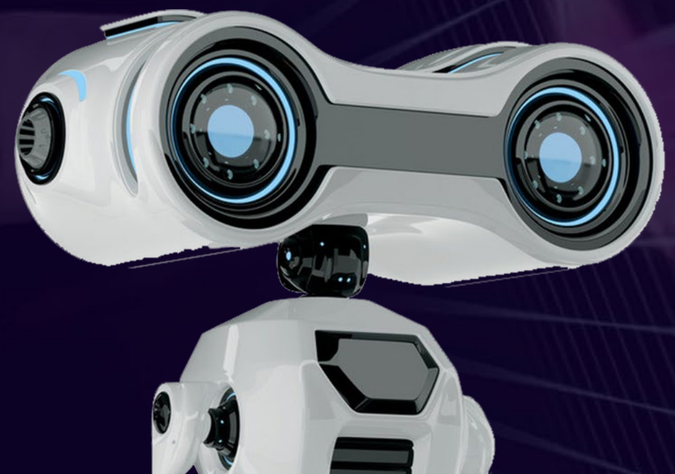
- When you want to consume the pipeline, you first need to deploy the pipeline.
- You can create an inference or a batch pipeline.
- Once the pipeline is ready, you can have an endpoint in place by deploying it to a compute solution such as Azure Kubernetes.
- You can then invoke the pipeline via a REST endpoint and an authentication key.

Automated ML



- Automated Machine Learning will automatically train the model by using different algorithms.
- It will create different pipelines for you for using the different algorithms.
- Here you will provide the training data that has been labelled.
- You don’t need any prior programming experience to use Automated Machine Learning.

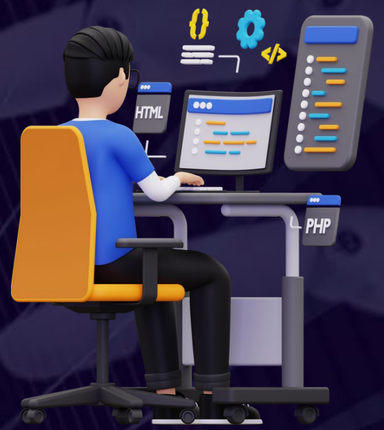
Using AI services

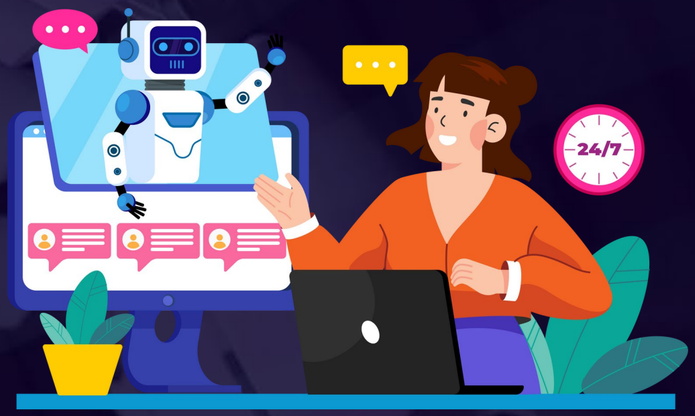


AI services

Azure has a host of in-built AI services. You can use this services to incorporate AI features within your application.

A user can upload a set of images to an application. The application should then generate different tags or divide them into groups based on the objects it finds within the images.





Now instead of developing the functionality of being able to detect objects within an image from scratch , we can leverage existing solutions on Azure to do this for us.

Azure AI Vision

Generate captions

The service can be used to generate captions and descriptions from images.

Detect Objects

We can also detect individual objects from images.

Optical character recognition

We can also extract text from images.

AI services

Azure AI Vision

This is a specific resource for Azure AI Vision. You can use this service if you don't want to use other AI services. And if you want to track just the cost of using this service.

Azure AI services

If you want to use the Azure AI Vision service along with other services - Azure AI Language, Azure AI Custom Vision.

Custom Vision

Custom Vision

- This is a service that allows you to build and deploy your own image identification models.
- You can use your own labels to train the model.

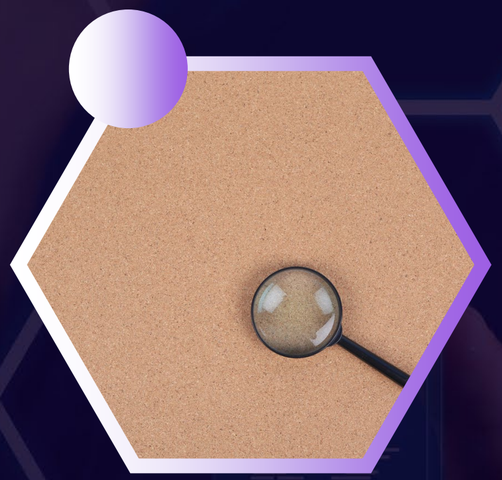
Custom Vision

- You can use Custom Vision via the Custom Vision portal. Or we can use it from a programming language by using the available SDK or REST API.
- The Custom Vision service will allow you to first train a model with your own set of images. You can label the images to train the model.

Custom Vision

Image classification

Here you can apply one or more labels to an image.



Object Detection

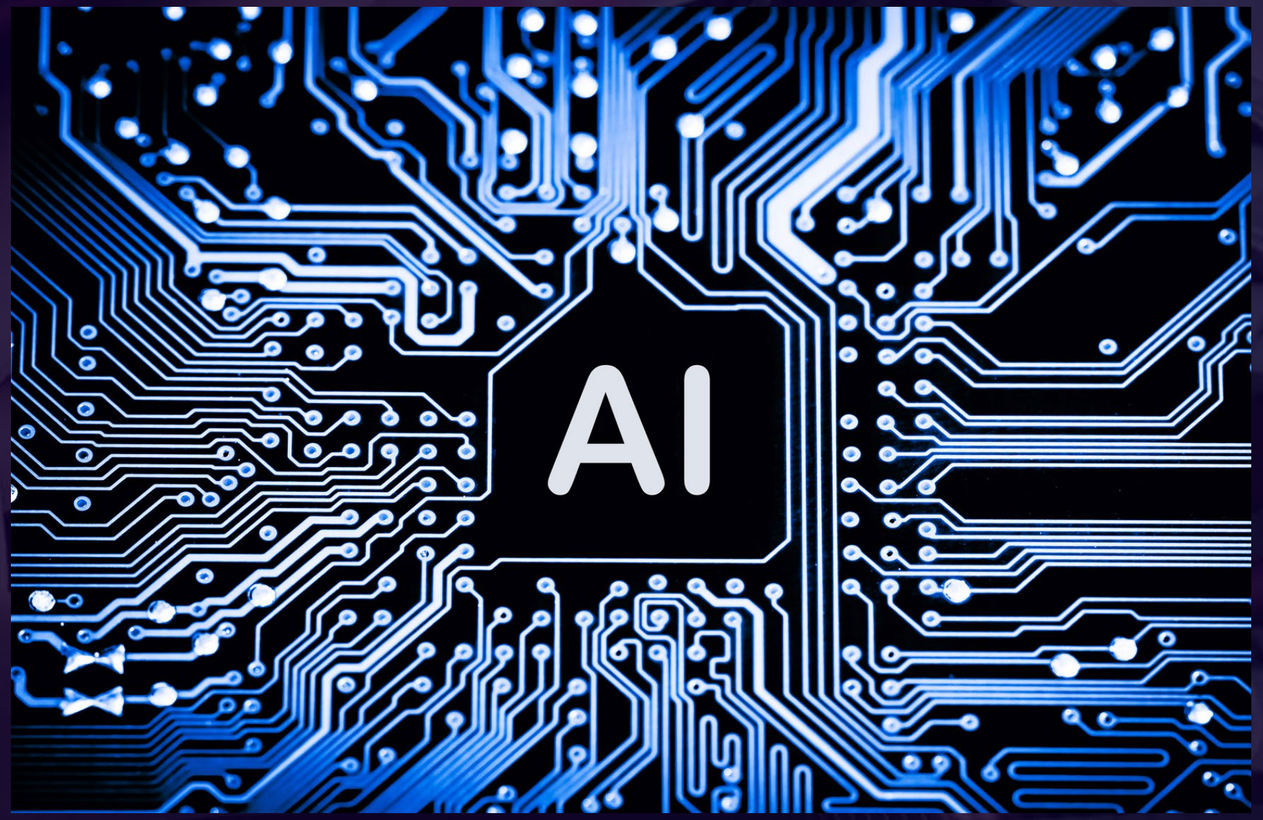
This is used to return the coordinates of the label within the image.

Custom Vision

Pricing and Limits - https://learn.microsoft.com/en-us/azure/ai-services/custom-vision-service/limits-and-quotas

Factor	F0 (free)	S0 (standard)
Projects	2	100
Training images per project	5,000	100,000
Predictions / month	10,000	Unlimited
Tags / project	50	500
Iterations	20	20
Min labeled images per Tag, Classification (50+ recommended)	5	5
Min labeled images per Tag, Object Detection (50+ recommended)	15	15
How long prediction images stored	30 days	30 days

Summary



Azure AI Vision

Image Analysis - We can extract visual features from images. We can get tags from images, generate context and even detect objects from images.

Optical Character Recognition - This can be used to extract text from images.

Face API - Here we can detect , recognize and analyze human faces within images.

Video Analysis - Features such as Spatial Analysis and Video retrieval are also available.

Custom Vision

- This service allows you to build, deploy your own image-based models.
- You can use Custom Vision via the Custom Vision portal. Or we can use it from a programming language by using the available SDK or REST API.
- You create a particular project type in the Custom Vision Service - Multilabel(Zero or more tags to your images) or Multiclass (images are sorted based on the most likely tag)
- You can choose a domain for the project.
- Upload the images, tag them and then train your model.
- Once the model is trained, publish the model and use the endpoint accordingly.

Face Service

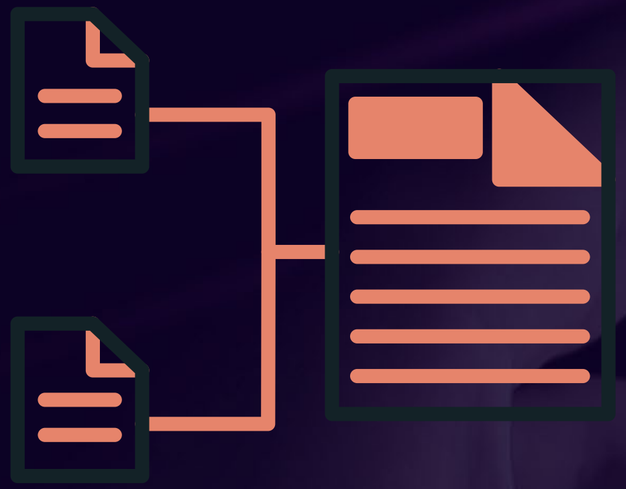
- This is a service that allows you to work with human faces within images.
- We can use the Detect API to detect faces within an image. Each face gets a bounded rectangle and can also be bound to a face ID.
- There are also some certain attributes that you can get about the face - Accessories - whether the human is wearing glasses , mask.
- Face Recognition - Here you can create a PersonGroup object that holds human faces. You can then identify a particular face from the PersonGroup object.
- You can also use the Verify API to verify if a face and a Person object is the same.

Document Intelligence

- This service can be used to process the data within forms and documents.
- There are a number of pre-built models in place. For example we have an invoice model that can be used to extract data from invoices.
- But we can also build our own custom models as well.

Azure AI Language Service

Azure Language service



- Named Entity Recognition
- This feature can be used to identify and categorize several entities within text.
- There is a pre-built set of categories - Person, Location, Event , Organization etc.

Azure Language service

- Language Detection
- This service can be used to detect more than 100 languages from text.

Azure Language service

- Sentiment analysis
- Here you can understand whether there is a positive or negative sentiment based on the written text.
- This feature also returns a confidence rating for the positive, negative or neutral sentences it detects within the text.

Azure Language service

- Key phrase extraction
- This can be used to detect the main topics within the text.

Azure Language service

- Custom question answering
- This service can be used to build a natural conversational client over your data.
- You can build a knowledge based on common asked questions and use it within your application.

Azure AI Language - Q and A

Azure AI Language - Q and A

Here you can add a natural conversional layer over your data.

You can then use this layer to create chat bots.

You can create a project in the Q and A system. Then add common question and answer pairs that are asked by users.

This is an ideal service to use when you have static information in place.