INTRODUCTION TO MACHINE LEARNING

- The field of study known as machine learning is concerned with the question of how to construct computer programs that automatically improve with experience.

- Machine learning is a sub-field of artificial intelligence that involves the development of algorithms and statistical models that enable computers to improve their performance in tasks through experience. These algorithms and models are designed to learn from data and make predictions or decisions without explicit instructions.

TECHNOLOGY STACK

- Anaconda

- Integrated Development Environment (IDE): Pycharm

- Python

- Jupyter Notebook

SUPERVISED MACHINE LEARNING

- The main goal in supervised learning is to learn a model from labeled training data that allows us to make predictions about unseen or future data. Here, the term supervised refers to a set of samples where the desired output signals (labels) are already known.

SUPERVISED MACHINE LEARNING

Labeled Data

Prediction

Square

Triangle

Model Training

Lables

Hexagon   Square

Triangle
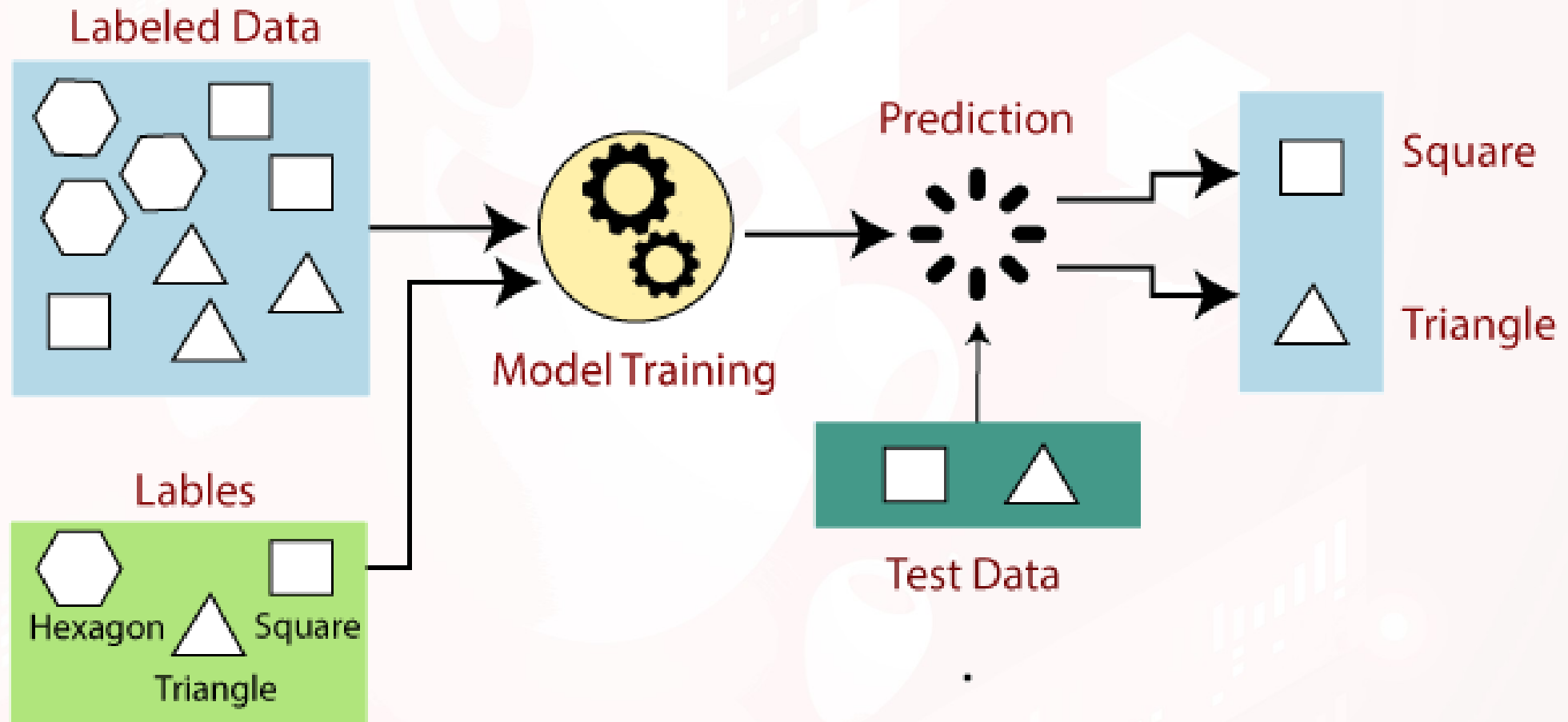
Test Data

SUPERVISED ML STEPS

- Supervised machine learning model involves two steps namely:

- **Learning (Training):** Learn a model using the training data.

- **Testing:** Test the model using unseen test data to assess model accuracy.

  } Accuracy = No of correct classifications/Total no of test cases.

SUPERVISED ML PROBLEMS

- Supervised machine learning problems can be grouped into two major categories:

- **Classification Problems:** This is when the output variable is a category such as male/female, red/blue or rose/lilly.

- **Regression Problems:** A regression problem is when the output variable is a real value such as weight or house price.

COMMON SUPERVISED ML ALGORITHMS

- Some of the common supervised machine learning algorithms include:

- Decision tree.

- K-Nearest Neighbors.

- Support Vector classifier (SVC).

- Logistic Regression.

- Linear Regression.

ADVANTAGES of SUPERVISED ML

- Some of the advantages of supervised ML include:

- It allows you to be very specific with the definition of lables.

- You are able to determine the number of classes you want to have.

- The input data is very well known and is labeled.

DISADVANTAGES of SUPERVISED ML

- Some of the disadvantages of supervised ML include::

- Growing data requires continuous training.

- Cannot discover new classes apart from the already labled ones during training.

UNSUPERVISED MACHINE LEARNING

- Here we only have input data X with no corresponding output variables. The goal is to learn more about the data as a result there is no correct answer and there is no teacher. The algorithm is left on its own to discover and present interesting structure in the data.
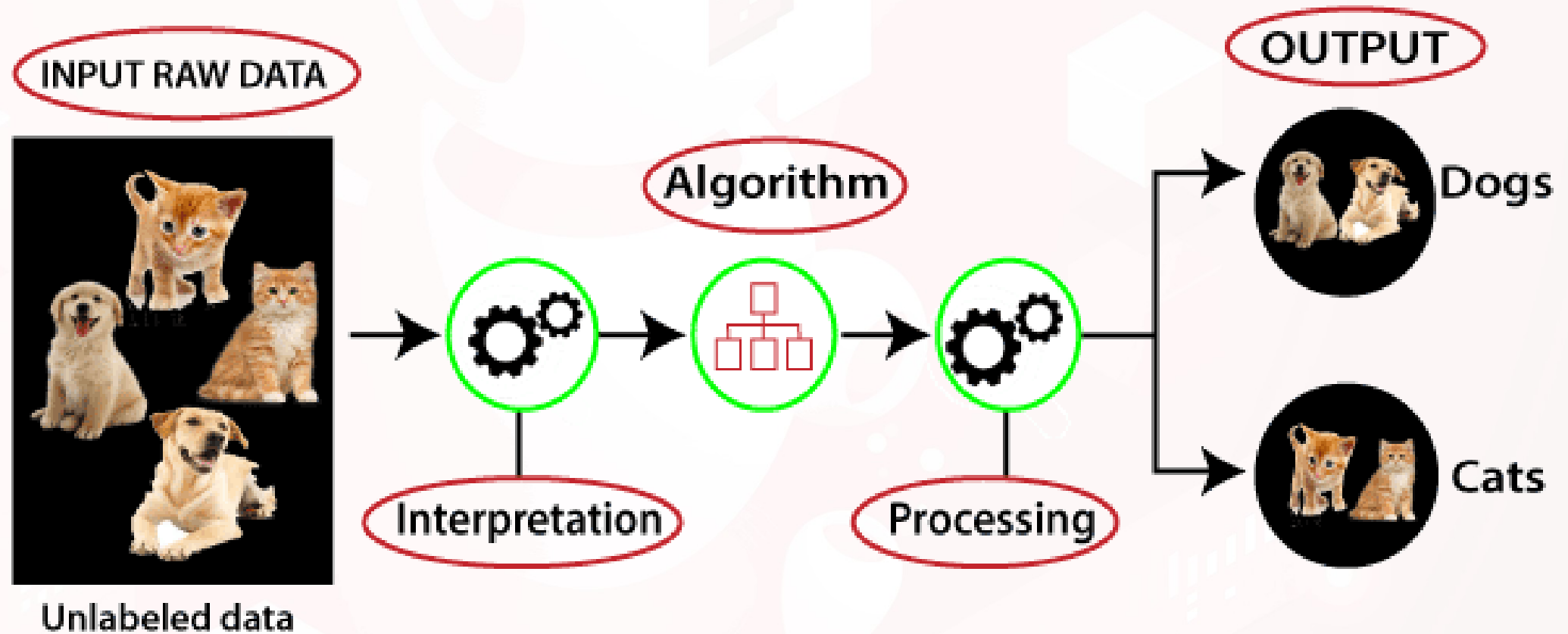
UNSUPERVISED ML GROUPS

- Supervised machine learning problems can be grouped into clustering and association problems:

- **Clustering:** A clustering problem is where you want to discover the inherent groupings in the data for example grouping customers by their purchasing power.

- **Association:** This is where you want to discover the rules that describe large portions of your data, such as people that buy X and also tend to by Y.

UNSUPERVISED ML ALGORITHMS

- Some of the common ML algorithms include:

- K-Means clustering.

- K-Nearest Neighbor (KNN).

- Dimensionality Reduction.

- Hierarchical clustering.

ADVANTAGES of UNSUPERVISED ML

- It has less complexity when compared with supervised machine learning algorithm.

-

DISADVANTAGES of UNSUPERVISED ML

- You cannot be specific about the definition of the sorting classes and output.

- Its difficult to prove the accuracy of results.

- Lack of specificity with the output classes limits its industrial applicability.

Questions

Intentionally Left Blank

RECOMMENDATION ENGINES

Recommender systems are a subclass of information filtering system that seek to predict the rating or preference that a user would give to an item.

Systems for recommending items (e.g. books, movies, CD's, web pages, newsgroup messages) to users based on examples of their preferences.

In this topic our focus will be on: Content based collaborative filtering.

TECHNIQUES OF RECOMMENDATION

1. Collaborative Filtering: This is mostly used on social media engines such as Facebook, Netflix and Twitter. The main logic behind this method is to find users who have similar taste and preferences to the target user then use this subset to provide recommendations. This method works best when you have a large number of user preferences. There are two main approaches when implementing collaborative filtering:

I. User based: Which uses user attributes such as location and gender.

II. Item based: Which focuses on attributes such as items purchased.

TECHNIQUES OF RECOMMENDATION

2. Content based Systems: This recommends items similar to those a user has liked (browsed/purchased) in the past. The major focus is on items and not other user's opinions.

This method has the advantage of not requiring the data on other users to make a recommendation.

DEMO AND QA

DEEP LEARNING

- This is a machine learning technique that teaches computers to do what naturally comes to humans.

- Deep learning is highly used in driver-less cars enabling them to recognize traffic signs, lampposts and pedestrians.

- With deep learning we aim to simulate the brain, helping systems to learn and perform complex tasks with increasing accuracy without human intervention.

Convolutional Neural Networks

- This is a machine learning technique that teaches computers to do what naturally comes to humans.

- Deep learning is highly used in driver-less cars enabling them to recognize traffic signs, lampposts and pedestrians.

- With deep learning we aim to simulate the brain, helping systems to learn and perform complex tasks with increasing accuracy without human intervention.