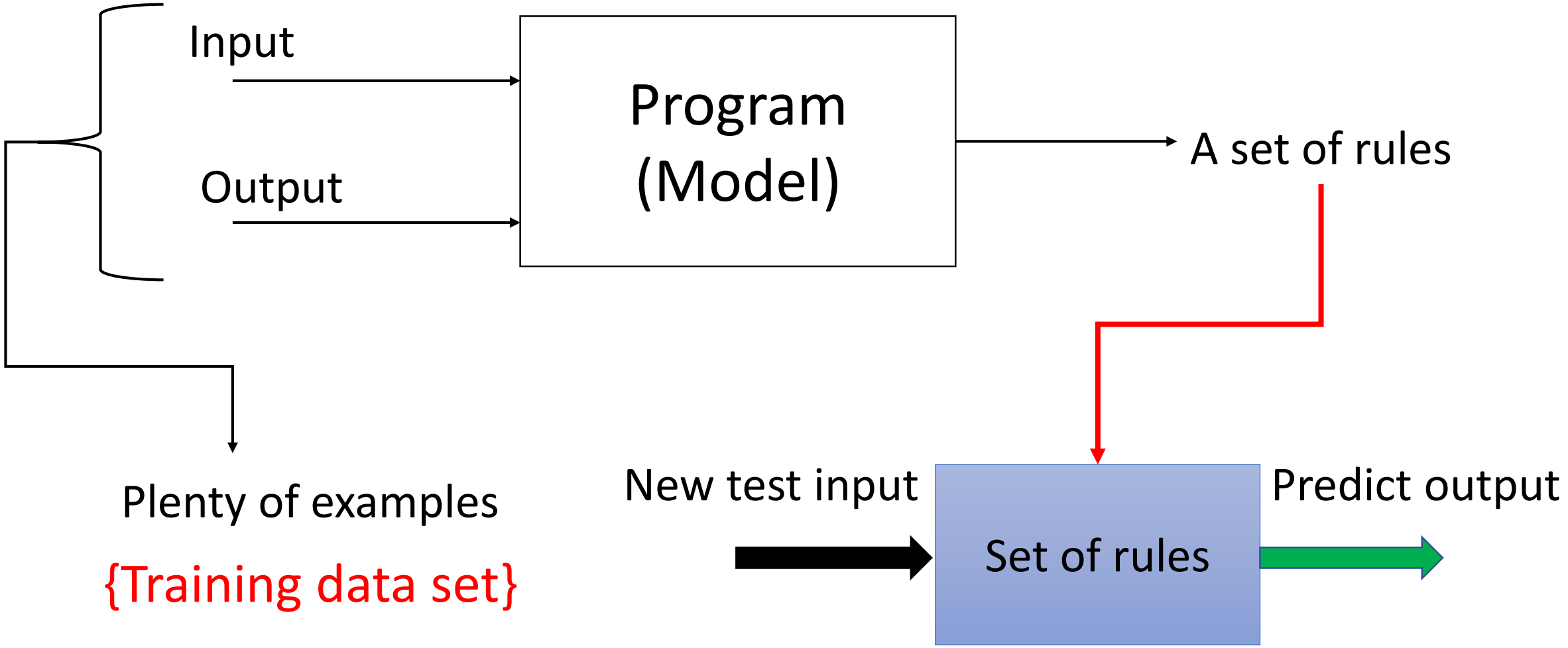
Lecture 2

Supervised Learning

Suman Samui

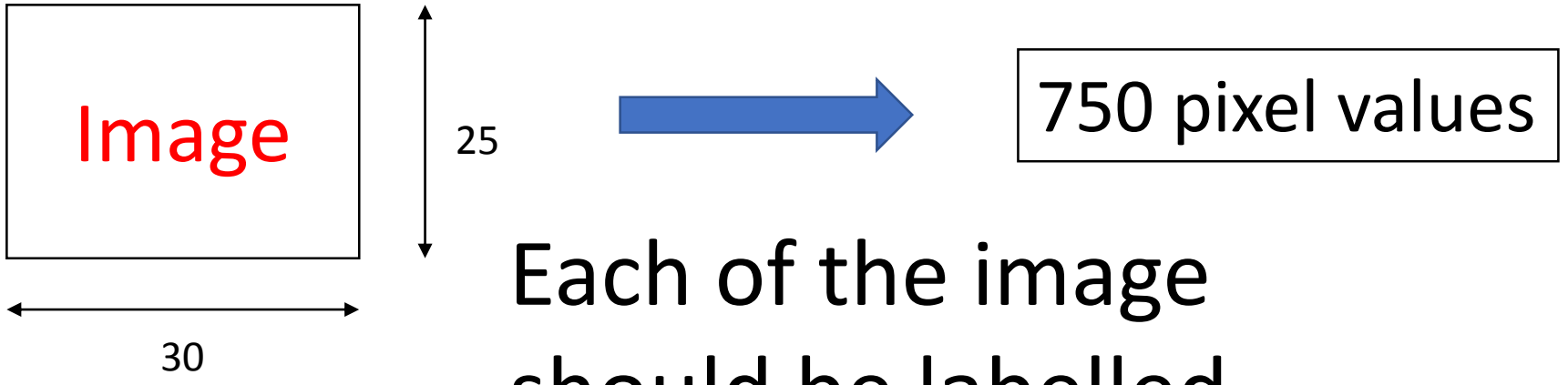
Machine Learning Approach



Classification Task: Computer Vision

Fruits Classification

- Machine learning model will learn from the images of fruits.
- Inputs are basically pixel values of the images.



Each of the image should be labelled

Image	Class
Mango	**0**
Apple	**1**
Banana	**2**

Used Car Price Prediction

Factors
Original Price of the car
Model
How much old?
Condition
Fuel type
Odometer status
Mileage etc. & many more

Car Price Prediction: Regression Problem

- You will have several examples of old cars and their features and resale prices.
- Machine learning model will be trained by the features of the cars
- After training ML model will the predict the resale price of a new car.
- Predicted value → Continuous in nature → Regression type of problems.

Generalization in Machine Learning

- It is desirable that our trained model should perform well on unseen data (which is not used for training).
- If your model performs well on training data but can't perform well on new unseen data then it's more likely that our model is memorizing the data patterns. It is not understanding underlying set of rules that correlating inputs and outputs.
- **Reasons of this type of result → Many folds:**

1. Insufficient data
2. Poor data quality
3. Non-presentative or uninformative features (data)
4. Chosen Model is not perfect, may be too simple → Underfitting
5. Chosen Model may be very complex (too many parameters) →Overfitting

Summary

- Supervised tasks → Classification / Regression

- For any problem

1. You start with a data set (labelled)
2. You select a model → such as neural network (a set of parameters relating input and output)
3. You will train the selected model
4. Finally you apply the trained model to make predictions on new cases (Inference), hoping that your trained model generalize well.