Classification in Machine Learning

Classification in Machine Learning

- Classification is a task of Machine Learning which assigns a **label value** to a **specific class** and then can identify a particular type to be of one kind or another.
- Types of predictive models in machine learning are:
  - Binary classification
  - Multi-Label Classification
  - Multi-Class Classification
  - Imbalanced Classification

Binary Classification for Machine Learning:

- The most popular algorithms which are used for binary classification are :
  - K-Nearest Neighbours (Also supports multiple labels)
  - Logistic Regression (Only for two labels)
  - Support Vector Machine (Only for two labels)
  - Decision Trees (Also supports multiple labels)
  - Naive Bayes (Also supports multiple labels)

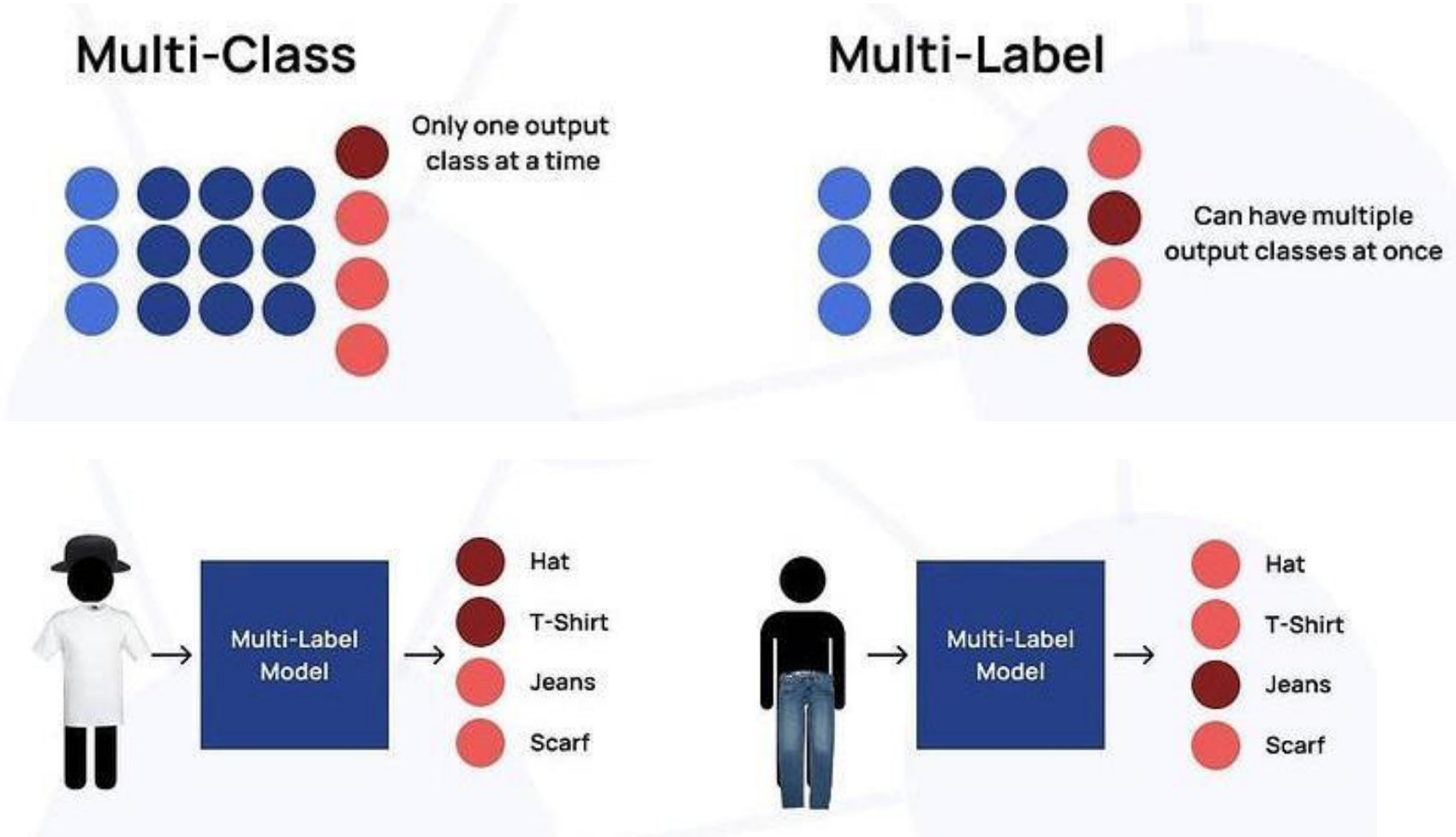
Multi-Class Classification

- These types of classification problems have no fixed two labels but can have any number of labels.
- Some popular examples of multi-class classification are :
  - Plant Species Classification
  - Face Classification
  - Optical Character recognition
- The most common algorithms which are used for Multi-Class Classification are :
  - K-Nearest Neighbours
  - Naive Bayes
  - Decision trees
  - Gradient Boosting
  - Random Forest

Multi-Label Classification for Machine Learning

- Here, we refer to those specific classification tasks where we need to assign two or more specific class labels that could be predicted for each example.
- A basic example can be photo classification where a single photo can have multiple objects in it, like a dog or an apple, etc.
- The main difference with multi-class is the ability to predict multiple labels and not just one

Multi-Class and Multi-Label Classification



Imbalanced Classification for Machine Learning

- An Imbalanced Classification refers to those tasks where the number of examples in each of the classes are unequally distributed.
- Generally, imbalanced classification tasks are binary classification jobs where a major portion of the training dataset is of the normal class type and a minority of them belong to the abnormal class.
- Some examples of these use cases are :
  - Fraud Detection
  - Outlier Detection
  - Medical Diagnosis Test
- Special modelling algorithms can be used to give more attention to the minority class when the model is being fitted on the training dataset which includes cost-sensitive machine learning models. Especially for cases like :
  - Cost-Sensitive Logistic Regression
  - Cost-Sensitive Decision Trees
  - Cost-Sensitive Support Vector Machines