

PRIOR KNOWLEDGE

Team ID	PNT2022TMID21349
Project Name	A Novel Method for Handwritten Digit Recognition System

One should have knowledge on the following Concepts:

1. Machine Learning

2. Clustering, Classification, Regression

3. Artificial Intelligence

4. Convolution Neural Networks

Machine Learning:

Machine learning is a branch of **Artificial intelligence (AI)** and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. It is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve performance on some set of tasks.

1. Supervised Machine Learning
2. Unsupervised Machine Learning
3. Semi-Supervised Machine Learning
4. Reinforcement Learning

Semi-Supervised learning is used in Text Classification.

1. Supervised learning:

Supervised machine learning is based on supervision. It means in the supervised learning technique, we train the machines using the "labelled" dataset, and based on the training, the machine predicts the output. Here, the labelled data specifies that some of the inputs are already mapped to the output. More precisely, we can say; first, we train the machine with the input and corresponding output, and then we ask the machine to predict the output using the test dataset.

Categories:

A) Classification

B) Regression

A) Classification:

Classification algorithms are used to solve the classification problems in which the output variable is categorical, such as **“Yes” or No, Male or Female, Red or Blue, etc.** The classification algorithms predict the categories present in the dataset. Some real-world examples of classification algorithms are **Spam Detection, Email filtering, etc.**

B) Regression:

Regression is a technique for investigating the relationship between independent variables or features and a dependent variable or outcome. It's used as a method for predictive modelling in [machine learning](#), in which an algorithm is used to predict continuous outcomes.

2. Unsupervised Learning:

Unsupervised learning is different from the Supervised learning technique; as its name suggests, there is no need for supervision. It means, in unsupervised machine learning, the machine is trained using the unlabeled dataset, and the machine predicts the output without any supervision.

Categories:

A) Clustering

B) Association

2) Clustering, Classification, Regression:

Clustering:

Clustering is an unsupervised technique. With clustering, the algorithm tries to find a pattern in data sets without labels associated with it. This could be a clustering of buying behaviour of customers. Features for this would be the household income, age, ... and clusters of different consumers could then be built.

Classification:

The Classification algorithm is a Supervised Learning technique that is used to identify the category of new observations on the basis of training data. It is used for spam for years now and these algorithms are more or less mature in classifying something as spam or not. With machine data, it could be used to predict a material quality by several known parameters such as humidity, strength, color, etc. The output of the material prediction would then be the quality.

Regression:

Regression is often confused with clustering, but it is still different from it. With a regression, no classified labels (such as good or bad, spam or not spam, ...) are predicted. Instead, regression outputs continuous, often unbound, numbers. This makes it useful for financial prediction.

3. Artificial Neural Networks:

Artificial neural networks, usually simply called neural networks or neural nets, are computing systems inspired by the biological neural networks that constitute animal brains. An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain.

Type of Neural networks:

- Perceptron.
 - Feed Forward Neural Network.
 - Multilayer Perceptron.
 - Convolutional Neural Network.
 - Radial Basis Functional Neural Network.
 - Recurrent Neural Network.
 - LSTM – Long Short-Term Memory.
- Sequence to Sequence Models.

4. Convolution Neural Networks:

A convolutional neural network (CNN) is **a type of artificial neural network used primarily for image recognition and processing, due to its ability to recognize patterns in images.** A CNN is a powerful tool but requires millions of labelled data points for training.

Layers of CNN:

The different layers of a CNN. There are four types of layers for a convolutional neural network: the convolutional layer, the pooling layer, the ReLU correction layer and the fully-connected layer.

Components of CNN:

1)Input layer

2)Output layer

3)One or more hidden layer

5.Flask:

Flask is a lightweight WSGI web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. It began as a simple wrapper around Werkzeug and Jinja and has become one of the most popular Python web application frameworks.

WSGI:

The Web Server Gateway Interface is a simple calling convention for web servers to forward requests to web applications or frameworks written in the Python programming language.

Jinja2:

Jinja2 is a modern day templating language for Python developers. It was made after Django's template. It is used **to create HTML, XML or other markup formats that are returned to the user via an HTTP request**