

## Assignment. No. 05

**Title:** Binary classification using Deep Neural Networks.  
Example - classify movie reviews into positive & negative reviews, just based on the total text content of the reviews. Use IMDB Dataset.

**Objective:** students should be able to classify movie reviews into positive reviews and negative reviews on IMDB Dataset.

**Theory:**

**classification**

classification is a type of supervised learning in machine learning that involves categorizing data into predefined classes or category based on a set of features or characteristics.

It is used to predict the class of new unseen data based on the patterns learned from the labelled training data.

In classification, a model is trained on labelled dataset, where each data process has a known class label.

The model learns to associate the features with the corresponding class labels and then be used to classify new, unseen data.



E.g. we can use classification to identifying whether an email is spam or not based on its content and metadata, data predict a patient has a disease based on their records and symptoms or to classify images into different categories based on their features.

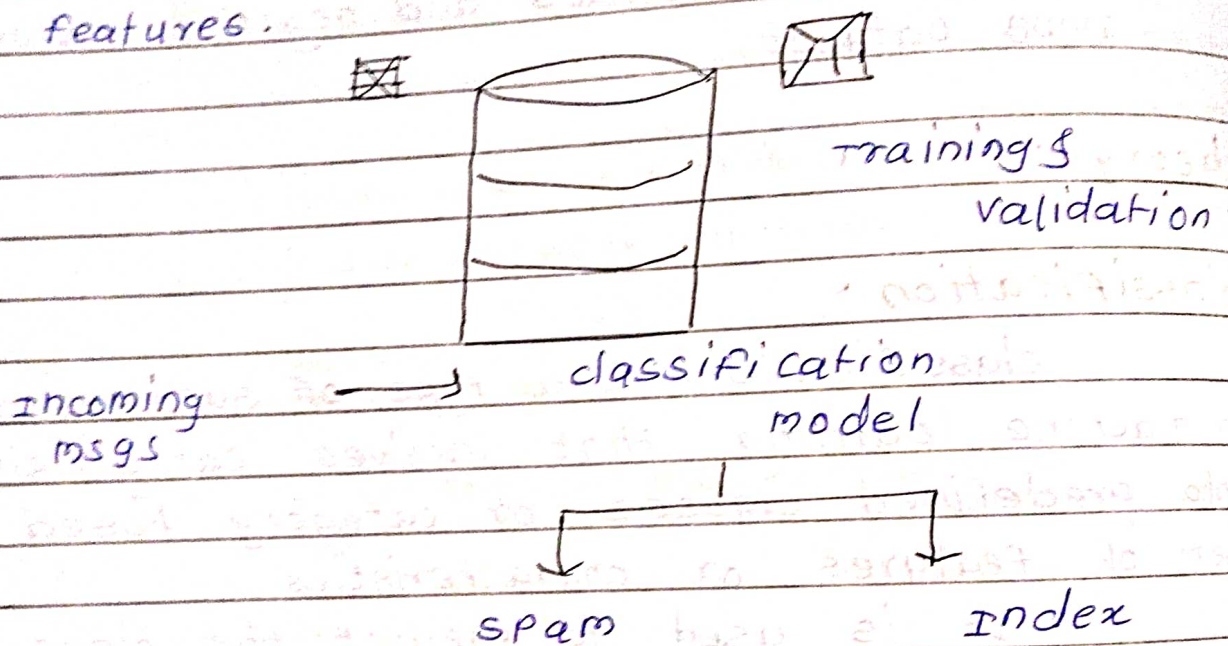


Fig. Email spam classification

The MNIST - dataset contains 60,000 training images and 10,000 testing images of handwritten digits from 0 to 9.

Each image is gray scale  $28 \times 28$  pixel in size and the task is to classify each image into one of the 10 classes corresponding to the digits 0 to 9.

We can use a convolutional neural network to classify the MNIST dataset.



A CNN is a type of deep Neural network that is commonly used for image classification tasks.

How deep Neural network work on classification

Deep Neural Network are commonly used for classification tasks because they can automatically learn to extract relevant feature from raw input data and map them to the correct output class.

The basic architecture of a deep neural network for classification consists of three main parts - an input layer, one or more hidden layers and an output layer.

The input layer receives the raw input data which is usually preprocessed to fixed size and format.

The hidden layer are compared of new data that apply linear transformation and non activations to the input features to extra relevant pattern and representations.

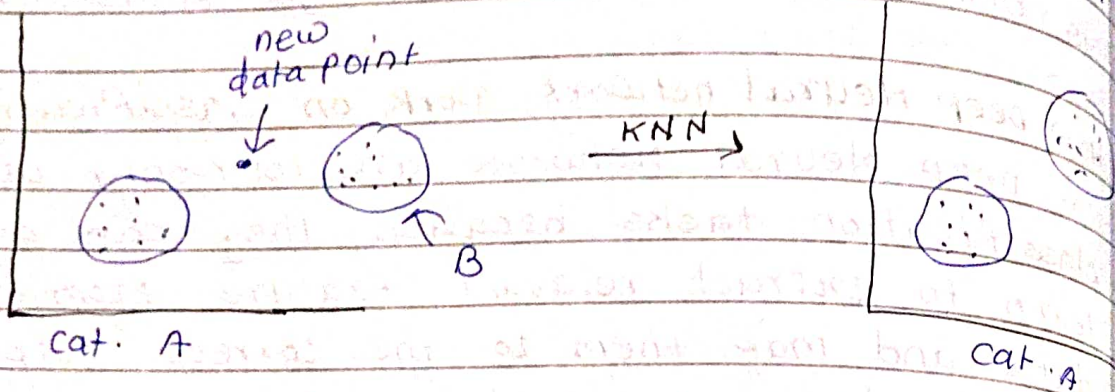
Finally the output layer produces the predict class labels usually as a probability distribution over the possible classes.

When trained properly deep neural networks can achieve state of the art performance on a wide range of classification tasks from image recognition to natural language processing.



Before KNN

After KNN



IMDB dataset :

The IMDB dataset is a large collection of movie reviews collected from the IMDB website, which is a popular source of user-generated ratings and reviews.

The dataset consists of 50,000 movie reviews, split into 25,000 reviews for training & 25,000 for testing.

The IMDB dataset is widely used in research & education for Natural language processing, machine learning, as it provides a review set of labelled text data for training & testing deep learning models.

Conclusion :

In this way, we can classify the reviews by using ONN.