**Paper-1**

**Title**:-**Classification using deep learning neural networks for brain tumors**

**Methodology**:-

The methodology for classification of brain tumor using DNN includes following four main steps.

Step 1: Brain MRIs Dataset acquisition

Step 2: Image segmentation using Fuzzy C-means

Step 3: Feature extraction using discrete wavelet transform

(DWT) and reduction using Principle component

analysis (PCA) technique

Step 4: Classification using DNN

**step-1**

According to the World Health Organization (WHO) classification system to identify brain tumors, there are more than 120 types of brain tumors which differ in origin, location, size, characteristics of the tumor tissues. In this paper three types of malignant brain tumour types are considered :

1. Glioblastoma

2. Sarcoma

3. Metastatic bronchogenic carcinoma

**step-2**

Image segmatation is used to separate different normal brain tissue from brain tumor tissue. Fuzzy C-means is used to segment the brain MRI into 5 sections.

**step-3**

After segmentation features of the segmented tumor is extracted using discrete wavelet transform (DWT). Methodology utilizes a 3-levels decomposition of Haar wavelet to extract 32\*32 i.e 1024 features for each brain MRI. Although this number is not so big compared to the number of feature maps resulted by the convolution filters of CNNs.Thus the principal components analysis (PCA) is used to approximate the original extracted features with lower dimensional feature

vectors.

**Step-4**

After the features are extracted and selected, the classification step using DNN is performed on the resulted feature vector. Classification is performed by using 7-fold cross validation technique for building and training the DNN of 7 hidden layers structure.

**Disadvantages**:-

1.Require more hardware specification and take more time for processing for large size images like(256\*256)

2.Require separate method for segmenation and feature extraction.

**Paper-2**

**Title:-Methods for interpreting and understanding deep neural networks**

In this paper we have studied the problem of interpreting a deep neural network model and explaining its predictions.

Machine learning techniques such as deep neural networkshave become an indispensable tool for a wide range of applica-tions such as image classification, speech recognition, or naturallanguage processing.Techniques for interpreting and understandingwhat the model has learned have therefore become a key ingredient of a robust validation procedure Interpretability is especially important in applications such as medicine or self-drivingcars, where the reliance of the model on the correct features mustbe guaranteed

**1.Interpreting DNN model:-**

This section focuses on the problem of interpreting a conceptlearned by a deep neural network (DNN). A DNN is a collection ofneurons organized in a sequence of multiple layers, where neuronsreceive as input the neuron activations from the previous layer, andperform a simple computation (e.g. a weighted sum of the inputfollowed by a nonlinear activation). The neurons of the networkjointly implement a complex nonlinear mapping from the input tothe output. This mapping is learned from the data by adapting theweights of each neuron using a technique called error backpropagation

**2. Explaining DNN decisions**

In this section, we ask for a given data point x, what makes it representative of a certain concept ω c encoded at the output of the deep neural network (DNN). The output neuron that encodes this concept can be described as a function f ( x ) of the input. A common approach to explanation is to view the data point x as a collection of features ( x i ) di = 1 , and to assign to each of these, a score R i determining how relevant the feature x i is for explaining f ( x ) .

**3.Layer-wise relevance propagation (LRP)**

LRP is a backward propagation technique, specifically designed for explanation. LRP was found to be broadly applicable and to have excellent benchmark performance. The LRP technique is rooted in a conservation principle, where each neuron receives a share of the network output, and redistributes it to its predecessors in equal amount, until the input variables are reached . LRP is furthermore embeddable in the theoretical

framework of deep Taylor decomposition .

**4.Recommendations and tricks**

Machine learning methods are often described in papers at an abstract level, for maximum generality. However, a good choice of hyperparameters is usually necessary to make them work well on real-world problems, and tricks are often used to make most efficient use of these methods and extend their capabilities. Likewise, techniques of interpretation often come with their own set of recommendations and tricks. While this section is mainly focaused on LRP, part of the discussion also applies to interpretation techniques in general.

**5.Evaluating explanation quality**

For general tasks, e.g. in the sciences, it can be difficult to determine objectively whether an explanation technique is good or not, as the concept predicted by the DNN may only be interpretable by an expert. Here, we present some strategies to systematically and objectively assess the quality of explanations.

**6.Applications**

Domains as extraction of domain knowledge, computer-assisted decisions, data filtering, or compliance

1)Model validation

2)Analysis of scientific data

**Paper-3**

**Title:-Visualizing Higher-Layer Feature of a Deep Network**

Beyond the model definitions and the quantitative analyses, there is a need for qualitative comparisons of the solutions

learned by various deep architectures. The goal of this paper is to find good qualitative interpretations of high level features represented by such models.

**1.Models**

The first model is a Deep Belief Net (DBN), obtained by training and stacking three layers as Restricted Boltzmann Machines (RBM) in a greedy manner. The second model, by Vincent et al. (2008), is the so-called Stacked Denoising Auto-Encoder (SDAE).

**2.Maximizing the activation**

We look for input patterns of bounded norm which maximize the activation of a given hidden unit 1 ; since the activation function of a unit in the first layer is a linear function of the input, in the case of the first layer, this input pattern is

proportional to the filter itself. The reasoning behind this idea is that a pattern to which the unit is responding

maximally could be a good first-order representation of what a unit is doing.

**3.Sampling from a unit of a Deep Belief Network**

**4.Linear combination of previous layers’ filters**

**Limitations**

one cannot find a simple representation of a higher layer unit

as we scale the datasets to larger and larger images i.e DNN cannot process larger size iamges efficiently.