1st Post

Convolutional Neural Network (CNNs):

Inspired by the human visual cortex, Convolutional Neural Networks (CNNs) are deep learning algorithms used for tasks like image recognition, segmentation, and classification, common in the field of Computer Vision. They're also versatile enough to be applied in areas such as NLP, time series analysis, and speech recognition.

How does CNN work under the hood?

CNNs process an input image by applying filters (small grids like 3x3 or 5x5) that "slide" over the image to extract features, identifying patterns like edges and textures. These features are combined to generate a feature map, which is passed through multiple layers to refine the output.

Layers in CNNs

I will be describing CNN layers in simple and shot way

1. Convolution layer: This layer applies filters (small grids of numbers like 3x3 or 5x5) over the image. These filters "slide" over the input image, identifying features like edges, textures, and patterns. The result is a feature map that highlights areas where these features are present in the image.

Example: A filter might detect edges, meaning it will emphasize areas of contrast, such as the boundary of an object in the image.

2. Activation Function: After the convolution operation, an activation function like ReLU (Rectified Linear Unit) is applied to introduce non-linearity. It helps the model learn complex patterns by turning negative values to zero and keeping positive values unchanged.
3. Pooling layer: This layer reduces the size of the feature map by summarizing the information in smaller regions. It pulls significant features by applying aggregation operations hence reducing the memory used while training the network.
4. Fully Connected layer: After several convolutional and pooling layers, the output is flattened into a single one dimensional matrix. This flattened data is passed through one or more fully connected (dense) layers, where every input is connected to every output. This step helps in making the final decision or prediction.

Example: If you're classifying handwritten digits, this layer would ultimately output probabilities for each digit (0 to 9), determining which one is most likely.

5. Softmax Layer: This is the final layer in classification tasks. It takes the output from the fully connected layer and converts it into probabilities for each class. The class with the highest probability is selected as the model's prediction

In a Nutshell

The CNN takes the input image, extracts features like edges or textures in the convolutional layers, and reduces data dimensions while retaining important features in the pooling layers, and finally uses the fully connected and softmax layers to make a prediction about the image (e.g., recognizing a digit or identifying an object in the image).