Introduction:

Nowadays, automobiles have become the most convenient mode of transport for every household in India. India has many distinct types of traffic sign each with distinct color, shape, and sizes this makes road traffic environment more complicated, and people expect to have traffic sign assist system for providing timely instructions and warnings to drivers regarding traffic restrictions and information. that provides drivers with traffic sign information, regulates driver operations, or assists in vehicle control to ensure road safety.

Introduction to the dataset:

The given Dataset consists of a train and test file. The train file has 932 image data entries.

Column(s)	Name	Meaning
0	ID	An integer identifier which is unique within a file for each image.
1	Class	An integer identifier the traffic sign type.
2 - 2501	C1 - C2500	Pixels of the 50 by 50 image stored in greyscale, with each pixel an integer value in range [0, 255].

The test file contains 400 data entries as train but in Class column is filled with zeros as we need to predict Class using machine learning models or deep learning models.

We started by exploring our dataset. We found that all classes have at least one image data. Classes like class 13, class 8 and class 26 have the lion's share in data. whereas class 1, class 20 and class 25 have a smaller data share. We noted that classes with a lion share might overfit and classes with a smaller share might underfit the model.

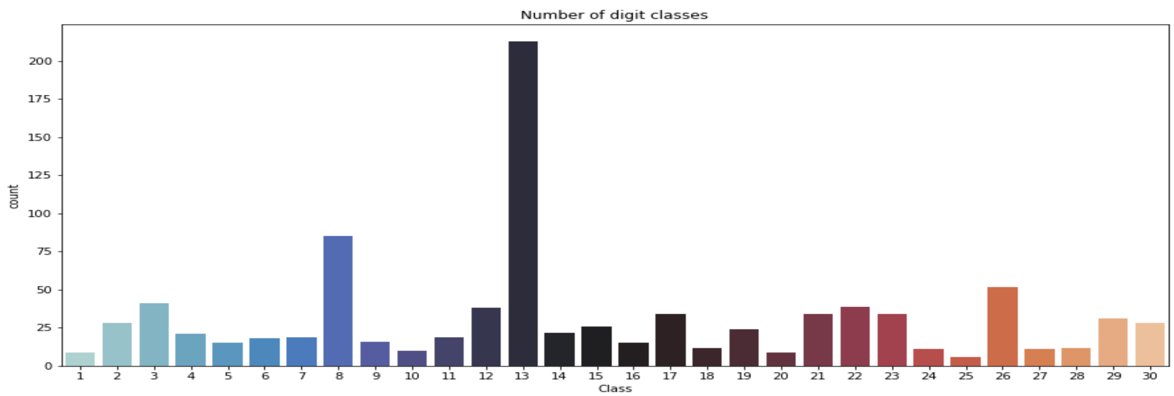


Figure 1

Methodology:

Some of the challenges from a computer vision perspective which are trivial for a human to perform.

1. **Illumination conditions**: Substantial effect of the light on pixel level of image.
2. **Intra-class variation**: Similar types of class images.
3. **Scale variation/Deformation**: Images of same class have different percentage of actual sign occupancy.
4. **Background clutter**: The Images of interest to be mixed with its environment/background.

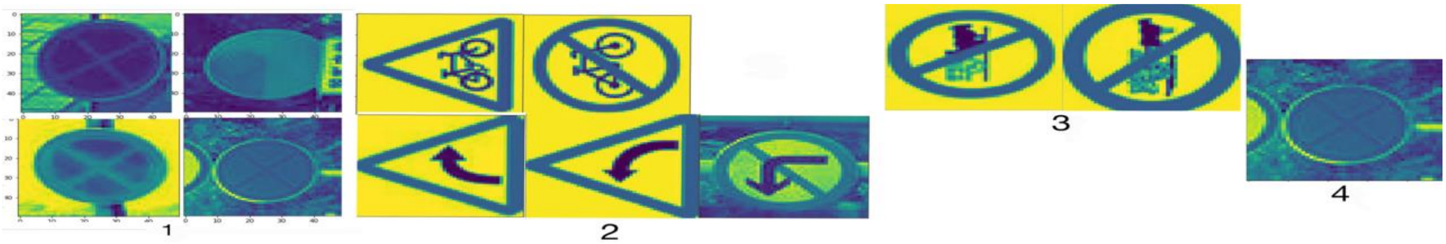


Image preprocessing, preparation and Modeling:

1. **Data normalization/Scaling:** scaled the features by dividing them by 255 to get values ranging from 0 to 1.
2. **Data augmentation:** Data augmentation is used to increase the data for classes which have lesser data by generating new data points from existing data.
3. **Resampling using SMOTETomek:** As we can understand from Figure 1 above, in order to account for the imbalance in the dataset, we have used combined sampling module SMOTETomek from imblearn.combine package to reduce the imbalance between the classes in the dataset.

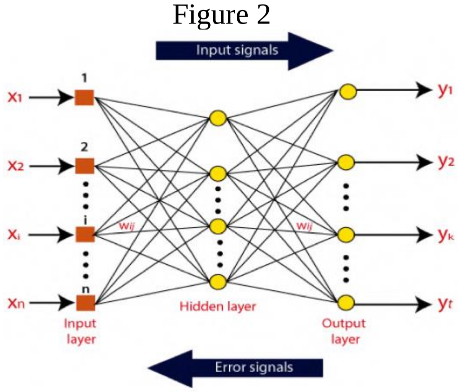
In order to predict the class of test file, we used the following approaches

a) Artificial Neural Network (ANN)
b) Convolutional Neural Network (CNN)
c) Convolutional Neural Network (CNN with dropout)
d) VGG16

(a) Artificial Neural Network (ANN):

Artificial Neural Network can be best represented as a weighted directed graph, where the artificial neurons form the nodes. The Artificial Neural Network receives the input signal from the external source in the form of a pattern and image in the form of a vector. These inputs are then mathematically assigned by the notations x(n) for every n number of inputs.

Figure 2



Each of the input is multiplied by its corresponding weights. and the total of weighted inputs is passed through the activation function. The activation function refers to the set of transfer functions used to achieve the desired output.
In our ANN model, we have used two hidden layers (activation function - ReLu), and a final output layer (activation function - SoftMax), optimizer-Adam, loss-sparse_categorical_crossentropy with 30-40 epochs and a validation split of 20%.
ReLu activation – It determines whether an input node will 'fire' given the input data and will apply a max (0, x) function, thresholding at 0.
Softmax function is used as the activation function in the output layer to predict a multinomial probability distribution.

In ANN, the "layers" are rows of data points that are hosted by neurons that all share the same neural network (fully connected network) and to learn, ANN uses weights because of this it is more prone to overfitting training dataset, and it causes error in prediction on test data.

(b) Convolutional Neural Network (CNN):

In CNN only the last layer is fully connected, has no weights directed graph and CNN uses filtration to analyze image inputs and casts multiple layers on images. Because of this CNN gives better results than ANN.

For this project, we have used three hidden layers (kernel size of (2,2), activation function - ReLu), connecting layer (activation function - ReLu) and final output layer (activation function - SoftMax). Since the number of input data is small, there is a chance of overfitting the model so to overcome this, we use a dropout layer as part of regularization.

(c) Convolutional Neural Network (CNN with dropout):

We applied a Dropout layer with a dropout ratio of 50% every hidden layer as it randomly inactivates some hidden neurons of the layeris are important in training CNN as they act as a regularization technique so to prevent overfitting on the training data. If they aren't present, the first batch of training samples influences the learning in a disproportionately high manner. This, in turn, would prevent the learning of features that appear only in later samples or batches.

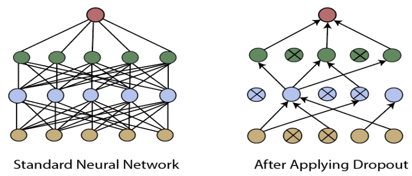


Figure 3

Say we show ten pictures of a circle, in succession, to a CNN during training. The CNN won't learn that straight lines exist; consequently, it'll be confused if we later show it a picture of a square. We can prevent these cases by adding Dropout layers to the network's architecture, in order to prevent overfitting.

Comparison between ANN, CNN and CNN (with dropout) model

We have used the classification_report module from the sklearn package to get an idea about the prediction report of the models. Apart from accuracy, we used **F1-score** value to compare the models, **F1-score** combines the precision and recall of a classifier into a single metric by taking their harmonic mean. **F1-score** ranges from 0 to 1, higher the **F1-score** better the classification of corresponding classes.

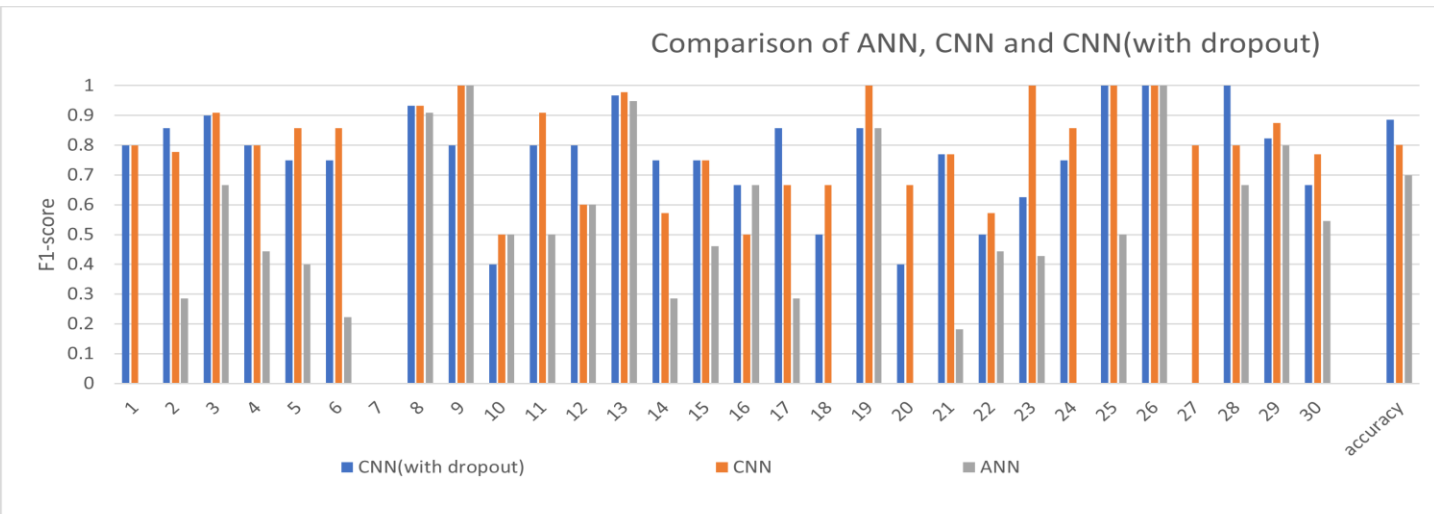


Figure 4

From figure 4 above, for few classes such as 1, 2 ,6 14, 17, 18, 21, 25, **F1-score** improves drastically when we move from ANN to CNN that means classification of this class improved in CNN models.

When we use CNN with dropout layers, **F1-score** stabilized to make the overall classification better than normal CNN model.

(d) Visual Geometry Group (VGG16):

Apart from the above 3 models - ANN, CNN, and CNN (with dropout), due to the small dataset, we tried VGG 16 model which works better with relatively small datasets as part of transfer learning. The VGG 16 model consists of 16 convolution neural network layers trained over 14 million images belonging to 22000 categories. But we could not use the final results in the competition due to some training issues.

References:

1. An illustration of the dropout mechanism within the proposed CNN. (a)... | Download Scientific Diagram (researchgate.net)
2. https://www.kaggle.com/code/pankul/image-classification-w-vgg16-weights