## 1. Number of samples in each folder

**Answer:**

We have three folders. The number of samples in each folder are as follows:

- train samples = 4176
- validation samples = 1392
- test samples = 1392

And each sample has 10 different classes.

## 2. Defining models and hyper parameter tuning

**Answer:**

For each model we updated the following things:

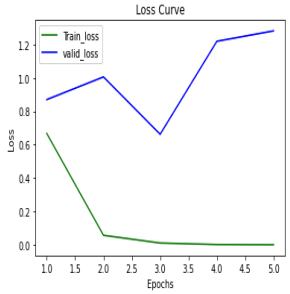
- Number of layers (2 or 3)
- Kernel Size - Tried values (2,2), (3,3) and (4,4).
- Number of Neurons in each layer - We tried different model having 16,32 and 64 neurons in different layers.
- Activation Functions - We tried different activation functions such as RELU, LEAKY RELU.
- Optimizers - We tried different optimizers such as Adam, RMSProp.

In the end, we chose the model which gave out best accuracy on our test set.

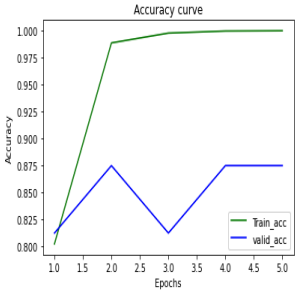
## 3. Training process without regularization

**Answer:**

Loss trend for training process




Accuracy trend for training process




## 4. Accuracy on test data without regularization

**Answer:**

Percentage accuracy for our test set is 81.25%

## 5. Network defined with regularization methods (L1 and L2)

**Answer:**

Model complexity is the primary cause of overfitting. As a result, we can prevent overfitting by regulating the complexity, which is exactly what regularization achieves. Regularization reduces the complexity of the model by punishing higher terms. When regularization terms are included, the model attempts to minimize both loss and model complexity.

There are two types of regularization: L1 Regularization: It's also known as sparsity regularization. It is used to manage sparse vectors that are primarily zeroes, as the name implies. Sparse vectors produce a high-dimensional feature vector space in most cases. As a result, the model becomes extremely difficult to work with.

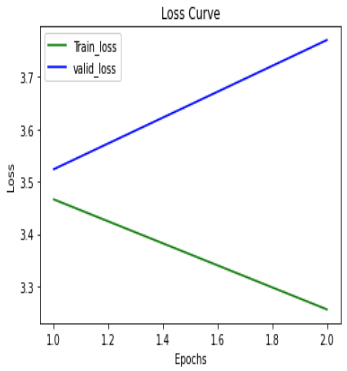
By subtracting a little amount from the weight at each iteration and finally making the weight zero, L1 regularization drives the weights of uninformative features to be zero.

L2 Regularization: For simplicity, it's also known as regularization. When we look at model complexity as a function of weights, we can see that the complexity of a feature is proportional to its absolute weight value.

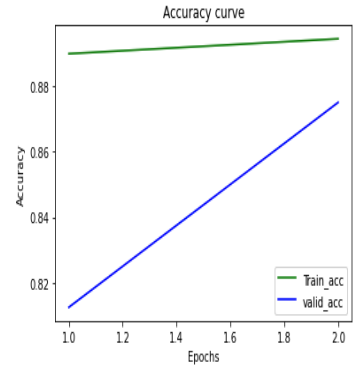
The regularization rate is an additional parameter that can be used to tweak the L2 regularization term (lambda). The L2 regularization term is multiplied by the regularization rate, which is a scalar.

6. Training process with L1 regularization
**Answer:**

Loss trend for training process



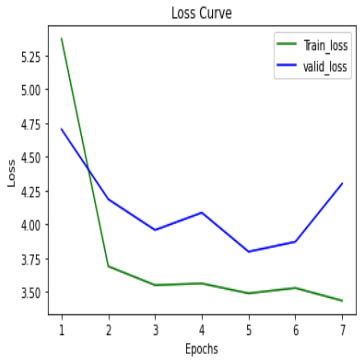

Accuracy trend for training process



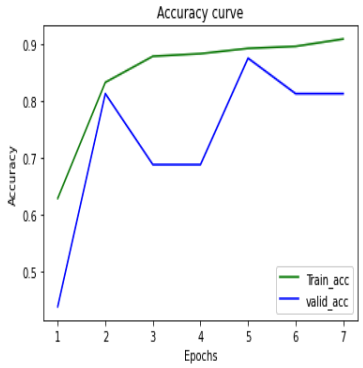

7. Accuracy on test data with L1 regularization
**Answer:**
Percentage accuracy for our test set is 65.01%

8. Training process with L2 regularization
**Answer:**

Loss trend for training process




Accuracy trend for training process




9. Accuracy on test data with L2 regularization

**Answer:**

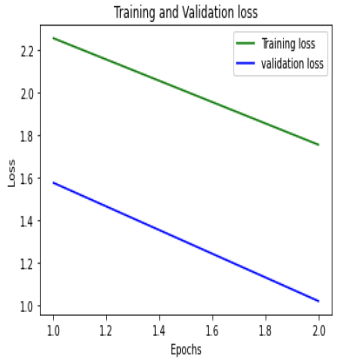
Percentage accuracy for our test set is 66.16%

10. Training process with Resnet

**Answer:**

Used Dropout method for regularization

Loss trend for training process



Accuracy trend for training process

```
[0.23084291815757751, 0.3773946464061737]
```


**Contribution**

| Team Member | Contribution (%) |
|---|---|
| Aryan Saini | 33.33 |
| Divyansh Chopra | 33.33 |
| Pragati Nagar | 33.33 |