|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Experiment** | **With** | | | **Without** | | |
|  | **Training Accuracy** | **Validation Accuracy** | **Testing Accuracy** | **Training Accuracy** | **Validation Accuracy** | **Testing Accuracy** |
| SGD/Adam | 98.82 | 92.83 | 91.48 | 99.57 | 96.50 | 96.70 |
| SMOTE |  |  |  |  |  |  |
| ReLU/Leaky-ReLU | 98.82 | 92.83 | 91.48 | 99.03 | 96.18 | 95.14 |
|  |  |  |  |  |  |  |

Experiment 5 Changing Layers in VGG16 network and Adding more Data Augmentation

* Optimizer = Adam
* learning rate = 0.0001
* Data Augmentation Techniques = image zooming, brightness change, horizontal and vertical flip, width and height shift,
* batch size = 20 while training
* Architecture = VGG16
* two hidden layer with 1024 and 512 units as classification layer followed by dropout of 0.25 respectively
* activation function = ReLU
* SMOTE = None

|  |  |  |
| --- | --- | --- |
| **Results** | | |
| **Training Accuracy** | **Validation Accuracy** | **Testing Accuracy** |
| 78.91 | 77.75 | 78.84 |

**Remarks**

Increasing model complexity has reduced generalizability and model tends to be noisy. It didn’t predict class 0 and 1 at all. Solution is to simplify the network

Experiment 8 Train AlexNet from Scratch

* Optimizer = Adam
* learning rate = 0.0001
* Data Augmentation Techniques = image zooming, brightness change, horizontal flip
* batch size = 20 while training
* Architecture = AlexNet
* activation function = reLU
* SMOTE = None

|  |  |  |
| --- | --- | --- |
| **Results** | | |
| **Training Accuracy** | **Validation Accuracy** | **Testing Accuracy** |
| 96.73 | 89.49 | 90.72 |

**Remarks**

Predict very well for class 1 and 3 i.e. mild and very moderate but shows poorer prediction on rest of the cases i.e. it is around 75%.

Experiment 9 Train VGG19

* Optimizer = Adam
* learning rate = 0.0001
* Data Augmentation Techniques = image zooming, brightness change, horizontal flip
* batch size = 20 while training
* Architecture = VGG19
* one hidden layer with 1024 units as classification layer followed by dropout of 0.25
* activation function = reLU
* SMOTE = None
* Epochs = 20

|  |  |  |
| --- | --- | --- |
| **Results** | | |
| **Training Accuracy** | **Validation Accuracy** | **Testing Accuracy** |
| 96.37 | 86.62 | 86.64 |

* **Remarks**

Does not predict class 1 at all. Shows poorer prediction on rest of the cases. Not satisfied results