|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Train-accuracy | Train-Loss | Validation-Accuracy | Validation-Loss |
| Baseline Model | 0.4728 | 1.4732 | 0.4676 | 1.4897 |
| Initialiser as random and zero | 0.4777 | 1.4596 | 0.4816 | 1.4654 |
| Weights as ones | 0.100 | 14.5073 | 0.100 | 14.5073 |
| Constant values | 0.0972 | 2.3027 | 0.100 | 2.3027 |
| Random Normal | 0.4761 | 1.4600 | 0.4773 | 1.4583 |
| Random Uniform | 0.4741 | 1.4653 | 0.4637 | 1.5043 |
| Truncated Normal | 0.4797 | 1.4520 | 0.4838 | 1.4516 |
| Variance Scaling | 0.4791 | 1.4554 | 0.4827 | 1.4596 |
| Orthogonal | 0.4681 | 1.4881 | 0.4449 | 1.5375 |
| Identity | 0.2469 | 1.9481 | 0.2471 | 1.9408 |
| Lecun\_uniform | 0.4816 | 1.4480 | 0.4646 | 1.4871 |
| Glorot\_normal | 0.4761 | 1.4659 | 0.4633 | 1.5064 |
| Glorot\_uniform | 0.4791 | 1.4555 | 0.4653 | 1.4953 |
| He\_uniform | 0.4743 | 1.4707 | 0.4674 | 1.4898 |
| Lecun\_normal | 0.4836 | 1.4428 | 0.4688 | 1.4875 |
| He\_uniform | 0.4737 | 1.4688 | 0.4710 | 1.4752 |

**Observations for prac 2 keras initializers**

* The best method is the one with the lowest train and validation loss and at the same time the highest train and validation accuracy. Thereby on these terms the truncated model out performs all the other models.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Train-accuracy | Train-Loss | Validation-accuracy | Validation-Loss |
| Cifar10 | 0.4796 | 1.4530 | 0.4837 | 1.4416 |
| Mnist | 0.9896 | 0.0326 | 0.9713 | 0.1116 |
| Cifar100 | 0.2224 | 3.2953 | 0.2156 | 3.2838 |
| Fashion\_mnist | 0.8992 | 0.2713 | 0.8759 | 0.3510 |

Therefore the best accuracy comes on Mnist using Truncated normal.