Members:

Pankhuri Bhatnagar pxb220012

Mohit Gullani mxg210090

Garima Rahangdale gxr220005

Hanumath Sai Abhilash Babu Veluvali hxv220011

Dataset Information:

Dataset link:

https://archive.org/download/inteldataset_20230726_0457/Intel_Dataset.zip

Dataset Characteristics:

The image collection contains 1478 photos of mountains and 1591 photos of buildings, all of which have a resolution of 150x150 pixels and are in the JPEG format. These images are split into two sets for training and validation. The training set consists of 1478 mountain photos and 1591 building photos, while the validation set contains 525 mountain photos and 437 building photos.

Class Information:

There are two categories of nature surroundings: Mountains and buildings

The data distribution of the images into different scenes is tabularized below:

Categories	Training Data	Validation Data	Testing Data
Mountains	1478	296	525
Buildings	1591	318	437

Libraries Used:

- PySpark
- Scikit Learn
- TensorFlow
- NumPy
- MatplotLib

Log Table:

Parameters Chosen			Results			
Iterations	**Learning Rate**	**Hidden Layers**	**Train Loss**	**Train Accuracy**	**Test Loss**	**Test Accuracy**
20	0.2	32	0.237422	0.649837	0.482328	0.275362
30	0.2	32	0.237875	0.625000	0.494802	0.268286
20	0.3	32	0.244604	0.612378	0.527027	0.263201
30	0.3	32	0.235981	0.641287	0.512474	0.263843
20	0.2	64	0.247650	0.603420	0.511435	0.274667
30	0.2	64	0.227156	0.650244	0.483368	0.302072
20	0.3	64	0.224887	0.655130	0.484407	0.297065
30	0.3	64	0.226300	0.646580	0.521830	0.279730
20	0.2	128	0.252797	0.597313	0.462578	0.299584
30	0.2	128	0.227806	0.662459	0.465696	0.321494
20	0.3	128	0.221137	0.686889	0.476091	0.299862
30	0.3	128	0.216289	0.697883	0.476091	0.310645

Summary:

Best Result Obtained at following parameters:

Iterations: 30 Rate: 0.3 Hidden Layers: 128 with result

Training Accuracy: 0.697883

Test Accuracy: 0.310645