Mini-Project-2

Author 1- Atharv Moghe
Roll No.- 220250

Author 2- Kuldeep Sandip Thakare
Roll No.- 220557

Author 3- Madhav Maheshwari
Roll No.- 220598

Author 4- Prabhat Kumar Yadav
Roll No.- 220774

Author 5- Siddhant Lalpuria
Roll No.- 221050

1 Problem 1

1.1 Task 1

1.1.1 Initial attempt

We initially attempted a simple, standard Learning with Prototypes (LwP) approach without any feature extraction, where the prototypes were chosen as the average of inputs of each class. For this, we used squared Euclidean distance, and the accuracy achieved was 80%. We then considered applying a ResNet50 feature extractor, which resulted in an accuracy of around 90%, so we discarded the first approach.

In the next step, we trained our LwP classifier on Dataset 1 using the true labels provided. Then, we used this model to predict labels for Dataset 2 and updated the prototypes accordingly. We repeated this process sequentially for all datasets from Dataset 1 to Dataset 10. However, this approach did not work effectively as the errors accumulated with each new dataset, degrading the performance over time.

89.59	None	None	None	None	None	None	None	None	None
88.96	88.47	None	None	None	None	None	None	None	None
87.98	87.22	87.00	None	None	None	None	None	None	None
86.47	85.91	86.02	84.88	None	None	None	None	None	None
85.42	85.44	85.14	84.65	85.14	None	None	None	None	None
84.87	84.28	84.26	83.10	83.65	84.88	None	None	None	None
85.28	85.34	85.07	84.58	84.58	85.06	83.60	None	None	None
84.87	84.88	84.56	83.59	83.47	85.21	83.32	84.83	None	None
83.83	84.51	83.84	82.91	83.23	84.51	83.30	84.51	83.87	None
84.18	84.01	84.12	82.34	83.23	84.18	83.14	83.56	83.46	84.90

Table 1: Without using Transfer Learning

1.1.2 Transfer Learning Approach

Instead of using knowledge distillation, we decided to use *transfer learning*, a technique in which a pre-trained model is fine-tuned on a new task. In our case, we leveraged pre-trained feature extractors, such as ResNet50, and fine-tuned them for our specific classification task. This approach involves the following steps:

- **Pre-trained Model:** We used a pre-trained model, such as ResNet50, which was previously trained on large datasets (e.g., ImageNet). This model has learned to extract general features that can be useful for many tasks.
- **Fine-Tuning:** We fine-tuned the pre-trained model on Dataset 1. This involved adjusting the weights of the pre-trained model to better fit the characteristics of our specific dataset.
- **Fine-Tuning on Sequential Datasets:** Once the model was fine-tuned on Dataset 1, we used it to predict labels for Dataset 2, and this process was repeated sequentially for all datasets up to Dataset 10. During each step, the model was fine-tuned to adapt to the new data and improve performance.
- **Transfer of Knowledge:** The pre-trained model, having learned general features from the large dataset, helps improve performance on smaller or specialized datasets by transferring learned knowledge. This transfer allows the model to learn faster and more effectively than starting from scratch.

1.1.3 Conclusion for Dataset 1

In conclusion, transfer learning provided a significant improvement over the initial LwP approach. By fine-tuning a pre-trained model, we achieved higher accuracy, especially when dealing with the complexities of sequential datasets. This approach minimized the error accumulation and allowed the model to generalize better across datasets.

88.68	None	None	None	None	None	None	None	None	None
88.84	89.91	None	None	None	None	None	None	None	None
88.67	89.55	88.74	None	None	None	None	None	None	None
89.02	90.31	89.13	89.24	None	None	None	None	None	None
89.06	90.23	89.56	89.06	87.73	None	None	None	None	None
89.93	89.94	89.94	89.51	88.31	88.91	None	None	None	None
89.96	89.96	89.66	89.24	88.88	89.07	88.53	None	None	None
89.96	90.17	89.88	89.11	89.06	89.85	88.69	89.50	None	None
89.66	90.16	90.16	89.57	89.55	89.85	88.92	90.21	88.89	None
89.96	90.31	89.89	89.05	89.00	89.85	88.62	89.79	89.07	90.31

Table 2: Using Transfer Learning

1.2 Task 2

1.2.1 Initial Attempt

Applying the same approach as in Task-1 leads to a decline in accuracy because the input distributions vary significantly, causing the prototypes to become inconsistent. However, using the features from dataset D10 on all datasets from D11 to D20 without updating yields decent accuracy, as the input distributions are relatively similar to the overall distribution of datasets D1 to D10. Subsequently, we experimented with Top-2 Pseudo Labeling (T2PL), which provided satisfactory results.

1.2.2 Top-2 Pseudo Labeling (T2PL)

This technique involves selecting only those examples where the combined probability of the top two predicted classes exceeds a specified threshold. By filtering out uncertain examples, this method helps enhance accuracy.

2 Task 2

Youtube link - Task 2

88.36	88.40	87.44	87.36	87.17	88.20	86.86	87.86	87.21	87.74	85.13	nan	nan	nan	nan	nan	nan	nan	nan	
85.99	86.32	84.86	84.67	84.94	85.36	84.44	85.98	84.94	85.67	74.92	68.20	nan	nan	nan	nan	nan	nan	nan	nan
87.71	87.97	86.71	86.94	87.48	87.40	86.75	88.32	87.32	87.63	77.64	64.44	79.33	nan	nan	nan	nan	nan	nan	nan
87.67	87.48	86.25	85.90	86.36	86.86	85.90	87.40	86.28	87.24	79.99	59.71	77.38	87.13	nan	nan	nan	nan	nan	nan
87.74	87.63	86.94	86.44	86.36	87.36	86.21	87.63	86.36	87.01	79.33	60.98	76.88	86.32	87.28	nan	nan	nan	nan	nan
87.17	86.55	86.09	85.82	86.55	86.86	85.67	87.48	85.48	86.51	74.04	65.36	78.07	84.40	86.36	78.26	nan	nan	nan	nan
86.55	85.90	84.98	85.36	85.09	85.75	84.86	86.51	85.29	85.63	79.10	60.63	76.00	85.44	85.32	74.57	85.42	nan	nan	nan
84.86	84.75	84.48	83.79	84.25	84.59	84.52	84.29	83.52	84.70	71.69	66.28	77.53	83.37	84.91	74.42	82.10	77.23	nan	nan
84.59	84.63	83.56	82.87	83.90	83.52	83.90	84.02	83.02	84.06	70.19	62.47	74.92	82.29	83.25	72.88	81.06	72.30	74.15	nan
82.21	82.71	81.98	81.25	81.56	82.32	80.60	82.02	81.77	82.02	73.34	53.33	69.12	80.18	82.17	64.97	79.22	68.73	49.53	83.02

Figure 1: 10×20 table

3 Acknowledgment

We thank Prof. Piyush Rai for his guidance and explaining the concepts in a well defined manner and the TAs of the course for their support in execution of the course.