Mini Project No 2

CS 771A : INTRODUCTION TO MACHINE LEARNING

Group No 51
Team Members:
Lt Col Divya Sharma 241110085
Yugul Kishor Pandey 241110083
Rajender Gill 241110055
Deepak Soni 241110018

IIT Kanpur

November 26, 2024

Machine Learning Models for Continual Shifting Domains

1. Problem 1: CIFAR-10 Dataset with LwP Classifier

1.1 Task 1

Objective

The primary goal of Task 1 is to develop a sequence of models ($f_1, f_2, \ldots, f_{10}$) using an LwP (Learning with Prototype) classifier. Starting with a labeled dataset D_1, the task involves updating the model iteratively across datasets D_2 to D_{10} using pseudo-labels generated by preceding models. The updated models should perform well on new datasets without significant degradation on previously encountered ones.

Detailed Approach

Dataset Characteristics:

- **Input:** CIFAR-10 images (32×32 color images).
- **Distribution:** Datasets D_1 to D_{10} share the same input distribution ($p(x)$).
- **Training:** Only D_1 is labeled; D_2 to D_{10} are unlabeled.
- **Evaluation:** Held-out datasets ($\hat{D}_1$ to $\hat{D}_{10}$) are used solely for performance assessment.

Feature Extraction:

- Pre-trained convolutional neural networks (CNNs), such as **ConvNeXtBase**, was used to extract image embeddings.
- **About ConNeXtBase:**The ConvNeXtBase neural network is a deep learning model built upon the ConvNeXt architecture, which is designed to optimize performance by integrating modern enhancements to traditional convolutional neural networks (CNNs). It's inspired by ResNet-like architectures while borrowing ideas from transformer models, making it a robust choice for various tasks, especially image classification and feature extraction.
  - Architecture
    * ConvNeXtBase follows a hierarchical design with multiple stages of convolutional blocks.
    * It introduces key innovations like layer normalization, better convolutional kernels, and fewer activation layers for improved efficiency and scalability.
  - Retraining
    * Retraining typically involves fine-tuning the pretrained ConvNeXtBase model on a new dataset.
    * **During retraining:**Pretrained Weights: Weights from models trained on datasets like ImageNet are used as initialization. **Feature Layers:** These layers act as general feature extractors for the new dataset.**Custom Output Layer:**Adjusted to match the new dataset's task, such as classification labels or regression outputs.
  - **Dataset Used for Pretraining:**
    * **ImageNet-1K** (1.28M images, 1,000 classes).
    * **ImageNet-22K** (14M images, 21,841 classes).
- These embeddings capture high-level features while reducing dimensionality.

Model Training with LwP:

- **Prototype Representation:** Each class in D_1 was represented as a prototype (mean of feature vectors).
- **Classification:** For any input, the label is predicted based on the nearest prototype in the feature space.
- **Loss Function:** Cross-entropy loss was used to optimize classification, ensuring compactness of class-wise clusters.

Iterative Model Update:

- Starting with f_1 trained on D_1, pseudo-labels for D_2 were predicted.
- Using the pseudo-labeled D_2, f_1 was updated to f_2.
- This process was repeated for all datasets (D_3 to D_{10}).
- A regularization term minimized parameter drift to reduce catastrophic forgetting.

Evaluation Methodology:

- Each model f_i was evaluated on its corresponding held-out dataset ($\hat{D}_i$) and all prior held-out datasets ($\hat{D}_1$ to $\hat{D}_{i-1}$).

Results

Accuracy Matrix for Task 1:

Model/Dataset	$\hat{D}_1$	$\hat{D}_2$	$\hat{D}_3$	$\hat{D}_4$	$\hat{D}_5$	$\hat{D}_6$	$\hat{D}_7$	$\hat{D}_8$	$\hat{D}_9$	$\hat{D}_{10}$
f_1	89.00%	-	-	-	-	-	-	-	-	-
f_2	88.44%	89.76%	-	-	-	-	-	-	-	-
f_3	87.84%	89.08%	87.48%	-	-	-	-	-	-	-
f_4	87.72%	89.08%	87.40%	88.68%	-	-	-	-	-	-
f_5	87.32%	89.24%	87.32%	88.52%	87.68%	-	-	-	-	-
f_6	87.20%	88.68%	86.80%	88.40%	87.44%	87.20%	-	-	-	-
f_7	87.12%	88.68%	86.88%	88.40%	87.16%	87.24%	87.20%	-	-	-
f_8	86.84%	88.40%	86.48%	88.28%	87.28%	87.20%	87.48%	87.64%	-	-
f_9	86.88%	88.44%	86.96%	88.00%	86.88%	87.12%	87.32%	87.44%	87.56%	-
f_{10}	86.84%	88.48%	86.92%	88.24%	87.04%	87.36%	87.32%	87.48%	87.60%	87.52%

Table 1: Accuracy Matrix for Task 1

1.2 Task 2

Objective

The aim of Task 2 is to extend the iterative model update approach to datasets D_{11} to D_{20}, starting with f_{10} as the base model. Unlike Task 1, these datasets exhibit different distributions, presenting additional challenges in model generalization and stability.

Approach

Challenges of Distribution Shifts:

- The input distributions $p(x)$ for D_{11} to D_{20} differ significantly.
- To address this, domain adaptation techniques were incorporated during model updates.

Model Update Methodology:

- **Feature Alignment:** Maximum Mean Discrepancy (MMD) was used to align feature distributions across domains.

Task 2 Accuracy Matrix:

	D̂1	D̂2	D̂3	D̂4	D̂5	D̂6	D̂7	D̂8	D̂9	D̂10	D̂11	D̂12	D̂13	D̂14	D̂15	D̂16	D̂17	D̂18	D̂19	D̂20
f11:	84.76%	86.24%	84.96%	86.20%	85.40%	85.40%	85.92%	85.48%	85.60%	85.56%	72.76%									
f12:	82.96%	83.80%	83.16%	84.48%	83.68%	83.36%	83.56%	83.80%	83.00%	83.84%	70.40%	58.48%								
f13:	82.24%	83.32%	82.76%	83.80%	83.12%	82.84%	83.36%	82.96%	82.64%	83.16%	70.28%	57.44%	76.36%							
f14:	82.60%	83.64%	83.40%	84.24%	83.72%	83.56%	84.08%	83.32%	83.28%	83.96%	71.28%	57.48%	76.60%	83.20%						
f15:	83.28%	84.44%	84.16%	84.92%	84.28%	83.92%	84.84%	83.64%	84.36%	84.96%	72.00%	57.80%	76.72%	84.00%	83.84%					
f16:	82.16%	83.04%	82.72%	84.00%	83.40%	82.60%	84.04%	82.80%	82.68%	83.20%	71.04%	55.96%	76.40%	82.40%	82.76%	74.48%				
f17:	82.00%	82.64%	82.24%	83.64%	83.12%	82.28%	83.60%	82.32%	82.00%	82.84%	70.88%	55.88%	75.88%	82.20%	82.52%	74.72%	79.44%			
f18:	80.96%	81.88%	81.32%	83.00%	82.56%	81.64%	82.64%	81.20%	81.28%	82.20%	70.52%	56.00%	75.36%	81.80%	81.24%	73.44%	77.80%	73.36%		
f19:	79.28%	79.60%	79.52%	80.48%	80.44%	79.64%	79.76%	79.52%	78.52%	79.92%	67.48%	54.32%	74.32%	79.40%	78.84%	71.96%	76.04%	71.16%	59.92%	
f20:	80.12%	81.08%	80.72%	81.80%	81.60%	80.64%	81.16%	80.12%	79.92%	80.68%	68.64%	55.00%	75.08%	80.44%	80.20%	72.96%	77.48%	72.28%	58.80%	79.52%

Task 2 completed successfully!

Figure 1: Accuracy Matrix for Task 2

- **Weight Adaptation:** A weighted loss function prioritized predictions on classes with higher confidence scores.

Regularization:

- To mitigate forgetting, a knowledge distillation loss was applied, preserving information from previous models.

Evaluation:

- Similar to Task 1, models f_{11} to f_{20} were evaluated on all 20 held-out datasets ($\hat{D}_1$ to $\hat{D}_{20}$).

2. Problem 2: Paper Presentation

Selected Paper

Deja Vu: Continual Model Generalization for Unseen Domains (ICLR 2023)

Summary

- **Problem Studied:** Proposes solutions to continual model generalization for unseen domains, focusing on robust feature extraction and memory consolidation.
- **Key Techniques:** Adaptive feature alignment across unseen domains and memory-based learning to retain past knowledge.
- **Insights:** Demonstrates improved performance on domain-shift tasks similar to those in Problem 1.

Youtube Link : `https://youtu.be/zi2QcjrRvGI`

3. Conclusion

This mini-project 2 explored continual learning using LwP classifiers and highlighted the challenges of maintaining performance across evolving distributions. Insights from domain adaptation research proved instrumental in achieving the objectives of Task 1 and Task 2.