
Homework 14

Regularization-based Lifelong Learning

ML TAs

ntu-ml-2020spring-ta@googlegroups.com

Overview

- Part 1
- Part 2
- Example Result
- Homework Spec
- FAQ

HW14 Part 1

- EWC - Elastic Weight Consolidation
- MAS - Memory Aware Synapse
- Sequential Datasets - SVHN - MNIST - USPS

本次作業中 你將走過以上兩個 regularization based lifelong learning 的方法
運用在三個不同的 dataset 上 並且它們訓練的過程是
先訓練 SVHN 再訓練 MNIST 再訓練 USPS
並做出 accuracy 表現近似最後 evaluation 的圖片

Colab 內有詳細說明實做的細節

HW14 Part 2

兩種選擇:

- 自己實做一個 lifelong learning regularization 的演算法
- Sliced-Cramer Preservation (助教有提供部份的 code 跟 pseudo code)

Based on 前面兩個方法, 助教都把 code 細節都寫好了, 你只需要調參數,

這次你需要參考之前的 block, 自己實做一個演算法

作法跟前面兩個雷同, 只是改變 importance 的算法, 可以參考 paper 的 pseudo code 作 implementation.

Part 2: Pseudo Code

Algorithm 1: Sliced Cramer Preservation (SCP)

Input: Data, $\mathcal{X}^A = \{\mathbf{x}_n^A \sim p_X^A\}_{n=1}^N$, and the optimized neural network, $\phi(\cdot; \theta_A^*)$, for Task A .

Parameters: Number of random projections, L .

Output: Synaptic importance matrix Γ

- 1 Calculate the mean response of the network at the targeted layer: $\bar{\phi}_A^* = \frac{1}{N} \sum_{n=1}^N \phi(\mathbf{x}_n^A; \theta_A^*)$.
 - 2 Initialize the synaptic importance matrix, $[\Gamma]_{i,j} = 0$.
 - 3 **for** $l \leftarrow 1$ **to** L **do**
 - 4 Sample ξ_l from \mathbb{S}^{K-1}
 - 5 Slice the mean response: $\rho = \xi_l \cdot \bar{\phi}_A^*$
 - 6 Calculate $\nabla_{\theta} \rho$ using auto-differentiation
 - 7 Update Γ : $\Gamma += \frac{1}{L} (\nabla_{\theta} \rho) (\nabla_{\theta} \rho)^T$
-

\mathbb{S}^{K-1} 是指一個 k-dimension 的 unit-sphere, 這個 k-dimension 會根據你想 regularize 的 layer 取他的 dimension。作業是取最後一層的 vector, 你需要從高維的球殼任意 sample 出一個 長度為 1 的 vector。

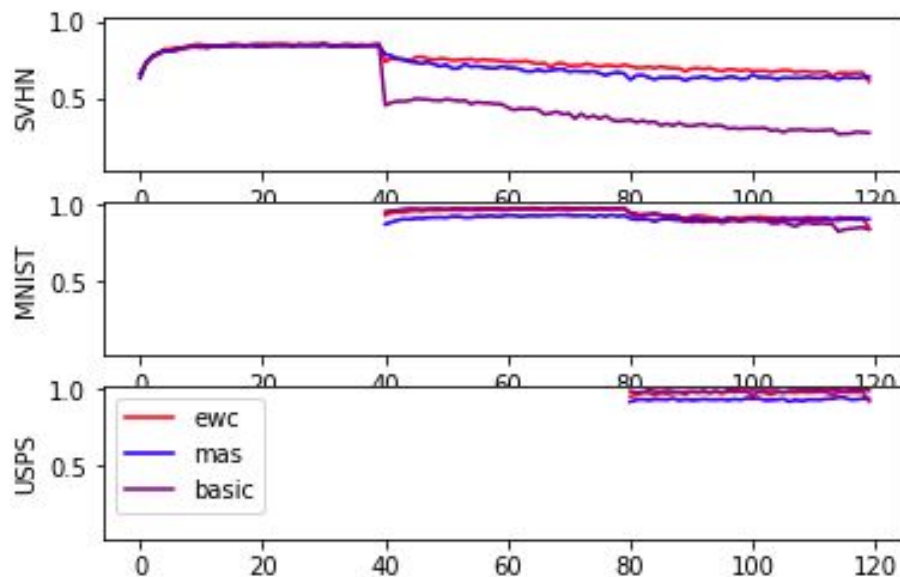
Fixed-model

在這次實驗的例子，採用的是固定的小模型 — 有六層隱藏層的全連接層模型。

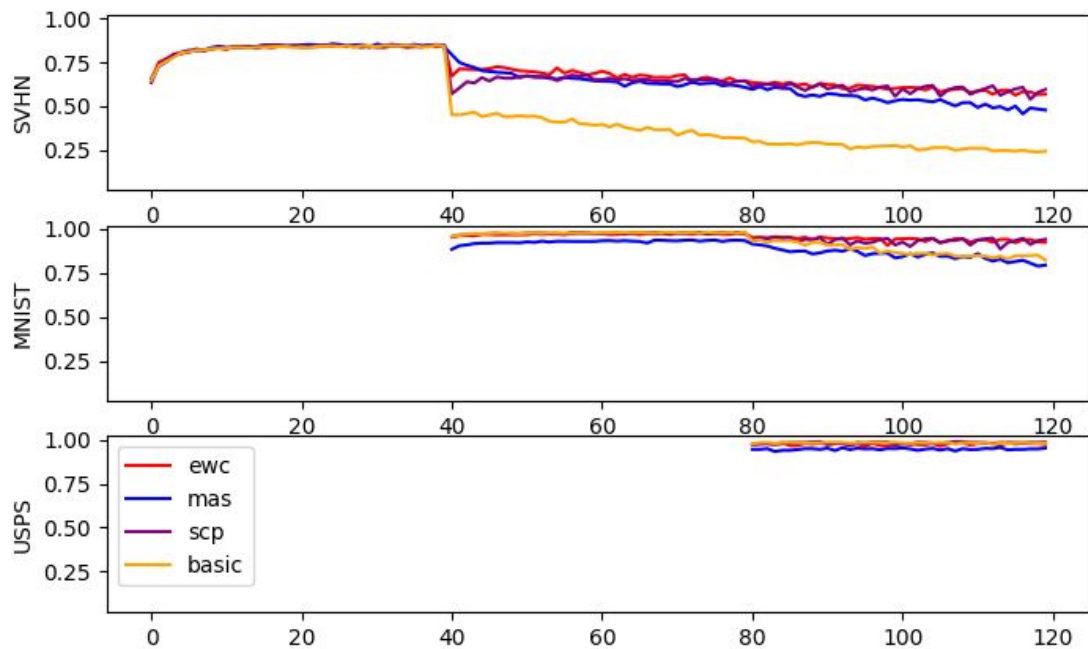
因為本次實驗是注重在 lifelong learning 訓練的方式 重點不在疊模型架構，

所以模型都是固定的。

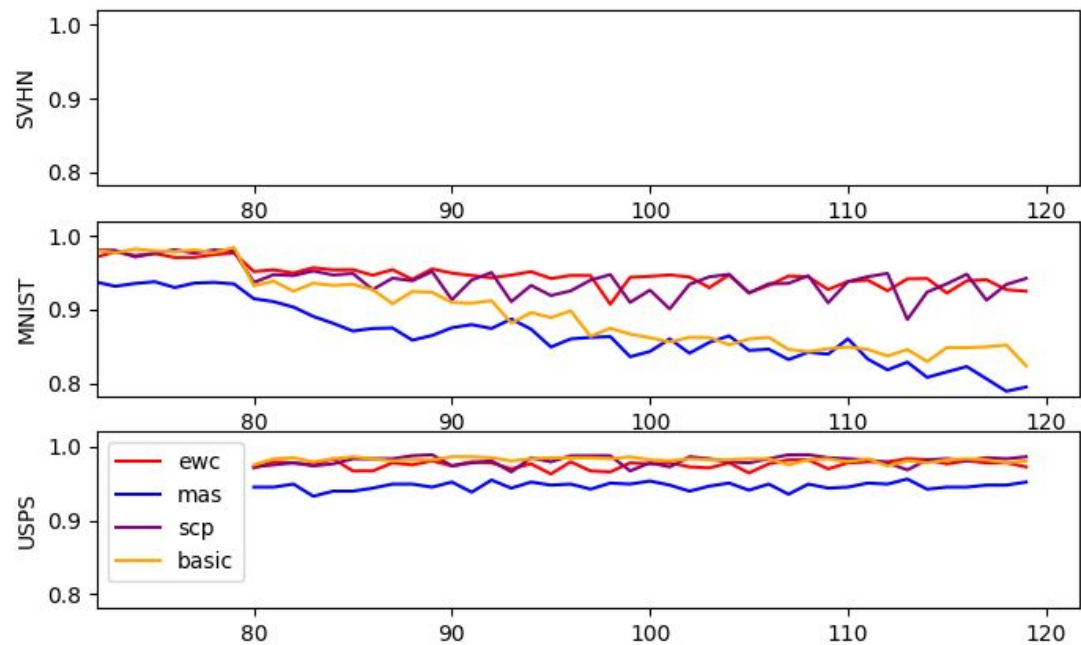
Part 1 Example Result



Part 2 Example Result



Part 2 Example Result



Report

1. (2%)請以中文說明一下 lifelong learning 的中心概念是什麼？
2. (2%)列出 EWC, MAS 的作法是什麼？根據你的理解，說明一下大概的流程該怎麼做 (不要貼 code)。
3. (1%) EWC 和 MAS 方法上所需要的資料最大的差異是什麼？
5. (5%)秀出 part1 及 part2 最後結果的圖，並分析一下結果，以及你跑的實驗中有什麼發現。

Submission Format

GitHub 的 hw14-<account> 必須包含以下檔案：

- report.pdf
- main.py(整理好成 py 並註明哪裡是 part 1, 哪裡是 part 2)

Scoring Policy

- 評分標準 以 report 為準。
- 最後一題的部份，請分別把兩個部份訓練完的圖片放到 report 上，如有額外的訓練數據、實驗，也可以自由補充。

FAQ

- 若有其他問題, 請貼文在 FB 社團裡或寄信至助教信箱, **請勿直接私訊助教**。
- 助教信箱: ntu-ml-2020spring-ta@googlegroups.com