

REDES NEURONALES 2025

Trabajo Final Integrador

Autoencoder y clasificador convolucional profundo sobre Fashion-MNIST y con pre-entrenamiento

Nota:

- Entreguen el trabajo final integrador **sólo** en formato **.pdf**. Si desean pueden enviar la notebook, pero por separado.
- El **.pdf** no puede tener más de cuatro (5) páginas.

En lo que sigue, trabaje utilizando el conjunto de datos **FashionMNIST**. En todos los casos, recuerde dividir el conjunto completo en un subconjunto de **entrenamiento** y un subconjunto de **validación**, que permitan evaluar el grado de sobreajuste y la capacidad de generalización de las redes neuronales empleadas.

Red autoencoder convolucional profunda

Defina, implemente y entrene un **autoencoder convolucional profundo**, es decir, una red compuesta por un **módulo codificador** y un **módulo decodificador**, ambos convolucionales y con múltiples capas. Elija un conjunto adecuado de hiperparámetros (por ejemplo: tamaño y cantidad de capas intermedias, probabilidad de *dropout*, tamaño y características de los *kernels*, tasa de entrenamiento, optimizador, etc.) y realice distintos experimentos variando dichos hiperparámetros hasta obtener un modelo con un rendimiento razonable. Utilice el **error cuadrático medio (MSE)** como función de pérdida u objetivo a minimizar durante el entrenamiento del autoencoder.

Pre-entrenamiento con una red clasificadora convolucional profunda

Defina, implemente y entrene una **red clasificadora convolucional profunda** empleando como **módulo de entrada** el **módulo codificador** del autoencoder previamente entrenado, seguido de una red convolucional multicapa que actúe como **módulo clasificador**.

Note que la salida del codificador debe utilizarse como entrada del clasificador. **Crucialmente**, durante este entrenamiento mantenga **congelados** los parámetros del módulo codificador, de modo que **solo** se entrenen los parámetros del módulo clasificador. En otras palabras, utilice el entrenamiento del autoencoder como una etapa de **pre-entrenamiento** para parte de la red clasificadora.

Seleccione un conjunto de hiperparámetros adecuados para el módulo clasificador y realice distintos experimentos variándolos hasta alcanzar un rendimiento satisfactorio. Utilice la **entropía cruzada** con **LogSoftMax** como función de pérdida. Calcule adicionalmente la **precisión** (*accuracy*) tanto en el conjunto de entrenamiento como en el de validación, con el fin de identificar la época a partir de la cual la capacidad de generalización deja de mejorar.

Evaluación de la efectividad del pre-entrenamiento

Repita los experimentos anteriores reentrenando la red clasificadora que haya obtenido los mejores resultados, considerando los siguientes esquemas:

- **1. Entrenamiento completo sin pre-entrenamiento:** entrene la red clasificadora completa partiendo de una red completamente sin entrenar, es decir, con codificador y clasificador inicializados aleatoriamente.

- **2. Entrenamiento completo (o fine tuning) con codificador pre-entrenado:** entrene toda la red clasificadora (codificador + clasificador) partiendo de un codificador pre-entrenado en el autoencoder y un clasificador sin entrenar.
- **3. Entrenamiento solo del clasificador sin pre-entrenamiento del codificador:** entrene únicamente el módulo clasificador utilizando un codificador sin entrenar (inicializado aleatoriamente).

Análisis de resultados

Presente y discuta los resultados más relevantes obtenidos en:

- el entrenamiento y selección de hiperparámetros del autoencoder convolucional,
- el entrenamiento y selección de hiperparámetros de la red clasificadora con pre-entrenamiento, y
- la comparación de los **cuatro esquemas** de entrenamiento mencionados:
 - el esquema original con codificador pre-entrenado,
 - y los esquemas **1.**, **2.** y **3.** detallados en la sección anterior.

Incluya las curvas de error y precisión, y describa las observaciones y conclusiones más relevantes.

Indicaciones para la redacción del trabajo final

El manuscrito debe incluir:

- **Título** del trabajo.
- **Autores** y sus afiliaciones académicas o laborales.
- **Resumen** (*abstract*) breve.
- Una sección de **Introducción** con referencias bibliográficas pertinentes.
- Una sección de **Teoría y/o Métodos**, describiendo con claridad las arquitecturas de redes utilizadas.
- Una sección de **Resultados**, con gráficos claros y debidamente rotulados.
 - No olvide incluir leyendas que describan los experimentos de manera tal que puedan reproducirse.
 - Incluya una matriz confusión para reportar el desempeño de la red clasificadora.
 - En el texto, resuma y discuta las observaciones más relevantes.
- Una sección de **Conclusiones**, donde se presenten las ideas más importantes derivadas del trabajo.
- Una sección con la lista de **referencias bibliográficas** citadas en el manuscrito.