

Tarea 4: Redes Neuronales

Objetivo

Este laboratorio tiene por objetivos: a) la construcción, utilizando bibliotecas de código abierto, de algoritmos basados en redes neuronales para ser aplicadas sobre un problema de clasificación; b) la evaluación de los modelos construidos

Problema

Considere al conjunto de datos «AIDS Clinical Trials Group Study 175»¹, con más de 2000 instancias y 23 atributos, ya utilizado en tareas anteriores:

- Construya conjuntos de entrenamiento y test a partir del conjunto de datos, con una proporción 80/20. A partir del conjunto de entrenamiento, separe un 10% para validación (a partir de ahora llamaremos “conjunto de entrenamiento” al 90% restante, y “conjunto de entrenamiento completo” al original). No incluya en estos conjuntos el campo time del conjunto de datos.
- Modelo 1 (base): utilice la biblioteca scikit-learn para entrenar un modelo de regresión logística a partir del conjunto de entrenamiento. Calcule la accuracy del modelo construido sobre el conjunto de validación.
- Modelo 2: construya, utilizando pytorch, una red neuronal compuesta por una sola neurona lineal (sin función de activación), con dos salidas, una para cada clase objetivo posible del conjunto de datos. Entrene esta red durante 100 épocas, utilizando como función de pérdida la función de entropía cruzada (CrossEntropyLoss). Grafique la pérdida y la accuracy tanto en el conjunto de entrenamiento como en el de validación, en función del número de época. Comente los resultados.
- Modelo 3: construya una red neuronal compuesta por una sola neurona, con función de activación sigmoide, y una sola salida, que estime la probabilidad de cada una de las clases objetivo. Entrene esta red durante 100 épocas, utilizando como función de pérdida la función de entropía cruzada binaria (BCEntropyLoss). Grafique la pérdida y la accuracy tanto en el conjunto de entrenamiento como en el de validación, en función del número de época. Comente los resultados.
- Modelo 4: construya una red neuronal consistente en una red feedforward con una capa oculta de 16 unidades, y que utilice la función sigmoide como activación, que devuelva la probabilidad de cada una de las clases objetivo. Repita el procedimiento de entrenamiento y evaluación del modelo anterior. Comente.
- A partir de los resultados anteriores, proponga tres arquitecturas adicionales que busquen mejorar los resultados, modificando la cantidad de unidades, la cantidad de capas ocultas, y/o diferentes funciones de activación, así como mecanismos de regularización. Para cada una, evalúe su rendimiento sobre el conjunto de validación, con diferentes valores de tasa de aprendizaje. Llamaremos Modelo 5 al modelo con mejores resultados sobre el conjunto de validación.

¹<https://archive.ics.uci.edu/dataset/890/aids+clinical+trials+group+study+175>

- Entrene los cinco modelos anteriores sobre el conjunto completo de entrenamiento (es decir, incluyendo el conjunto de datos de validación), y evalúe su performance en el conjunto de evaluación. Muestre en una tabla los valores de Accuracy, Precision, Recall y medida F1 de cada uno de los modelos. Comente los resultados y compare con los valores obtenidos en los laboratorios 1 y 2.

El trabajo debe realizarse utilizando PyTorch. Puede utilizar cualquier biblioteca adicional de Python que considere necesaria.

Entregables

- Informe con las pruebas realizadas y los resultados obtenidos. El informe a entregar debe ser un Jupyter Notebook, e incluir todos los archivos necesarios para su ejecución en un entorno local.
- Código escrito para resolver el problema.

Fecha límite de entrega

Miércoles 30 de octubre (inclusive)