

Tarea 2

Octubre 2021

Entrega: 12 de noviembre 2021

Profesor: Ángel Jiménez M.

Auxiliar: Rafael De la Sotta V.

Ayudantes: Camila Jorquera, Cristian Correa, Gabriel Iturra, Martín Ríos.

P1.- (3 ptos.) **Backpropagation del error cuadrático medio**¹. Considere una red neuronal feedforward de una capa oculta y una capa de salida, con datos de entrenamiento $T = \{(x_1, y_1), (x_2, y_2)\} = \{(0.05, 0.01), (0.1, 0.99)\}$, dos neuronas en la capa oculta y variable de salida $\hat{Y} = (\hat{y}_1, \hat{y}_2)$. Sea $W^{(1)}$ la matriz de parámetros de la capa oculta y $w_{j,i}^{(1)}$ el parámetro de la conexión entre la entrada x_i y la neurona j de la capa oculta, tal que $w_{1,1}^{(1)} = 0.15$, $w_{1,2}^{(1)} = 0.2$, $w_{2,1}^{(1)} = 0.25$ y $w_{2,2}^{(1)} = 0.3$, mientras que $w_{j,0}^{(1)}$ corresponde al parámetro de sesgo de la neurona j , tal que $w_{1,0}^{(1)} = w_{2,0}^{(1)} = 0.35$. Por otro lado sea U la matriz de parámetros de la capa de salida y $u_{k,j}$ el parámetro de la conexión entre la neurona j de la capa oculta y la componente k de la variable de salida, tal que $u_{1,1} = 0.4$, $u_{1,2} = 0.45$, $u_{2,1} = 0.5$ y $u_{2,2} = 0.55$, mientras que $w_{k,s}$ corresponde al parámetro de sesgo de la componente k de la variable de salida, tal que $w_{1,s} = w_{2,s} = 0.6$. Tanto la capa oculta como la capa de salida se activan con una función sigmoid $\sigma(a) = \frac{1}{1+e^{-a}}$. Considere como función de pérdida del entrenamiento al error cuadrático medio dado por:

$$L = \frac{1}{2} \sum_{k=1}^2 (y_k - \hat{y}_k)^2 \quad (1)$$

- (0.5 ptos.) Dibuje un grafo de computación para la red neuronal feedforward. Utilice el algoritmo forward propagation para computar manualmente los valores \hat{y}_1 , \hat{y}_2 y de la pérdida L .
- (2 ptos.) Utilice el algoritmo de backpropagation para computar manualmente el gradiente de L con respecto a cada uno de los parámetros de la red, es decir, $\frac{\partial L}{\partial w_{1,1}^{(1)}}$, $\frac{\partial L}{\partial w_{1,2}^{(1)}}$, $\frac{\partial L}{\partial w_{2,1}^{(1)}}$, $\frac{\partial L}{\partial w_{2,2}^{(1)}}$, $\frac{\partial L}{\partial u_{1,1}}$, $\frac{\partial L}{\partial u_{1,2}}$, $\frac{\partial L}{\partial u_{2,1}}$ y $\frac{\partial L}{\partial u_{2,2}}$.
- (0.5 ptos.) Considere una tasa de aprendizaje $\eta = 0.5$ y actualice todos los parámetros de la red de acuerdo al algoritmo del descenso del gradiente, luego compute nuevamente los valores \hat{y}_1 , \hat{y}_2 y de la función de pérdida L . Concluya.

P2.- (3 ptos.) **Gestión de riesgo financiero.** En este problema abordaremos una de las tareas básicas de gestión de riesgo financiero, que consiste en identificar el riesgo de cesación

¹ Esta pregunta es larga y mecánica, pero es algo que hay que hacer por lo menos una vez en la vida; se recomienda partir con anticipación

de pago de las tarjetas de crédito por parte de los/las clientes/as de un banco. Una manera de enfrentar este problema es utilizar datos transaccionales para obtener el comportamiento histórico de pago mensual del crédito otorgado a través de las tarjetas, así como datos demográficos de los/las clientes.

Considere el archivo *cesacion.csv* compuesto por las siguientes variables:

- X_1 : Monto en pesos del crédito otorgado.
- X_2 : Género (en este caso se reportaron sólo dos géneros, 1 para masculino y 2 para femenino)
- X_3 : Educación (1 para posgrado, 2 para pregrado, 3 para enseñanza media y 4 para otros tipos de educación)
- X_4 : Estado civil (1 para casado/a, 2 para soltero/a, 3 para otros estados)
- X_5 : Edad en años.
- X_6 a X_{11} : Comportamiento histórico de pago mensual del crédito entre abril y septiembre del mismo año, en que X_6 representa el **estado de pago** en septiembre, X_7 en agosto, X_8 en julio, X_9 en junio, X_{10} en mayo y X_{11} en abril. El estado de pago se define de la siguiente manera: -2 significa saldo pagado en su totalidad y sin transacciones este período (podemos referirnos a esta cuenta de tarjeta de crédito como 'inactiva' este período); -1 saldo pagado en su totalidad, pero la cuenta tiene un saldo positivo al final del período debido a transacciones recientes cuyo pago aún no ha vencido; 0 si el/la cliente pagó el monto mínimo adeudado, pero no el saldo completo, es decir, el cliente pagó lo suficiente para que su cuenta se mantuviera al día, pero giró un saldo; 1 si el pago se postergó por un mes; 2 si se postergó por dos meses, y así sucesivamente hasta 9 en caso que el pago se haya postergado por 9 nueve o más meses.
- X_{12} a X_{17} : Monto del estado de cuenta, en que X_{12} representa el monto en septiembre, X_{13} en agosto, X_{14} en julio, X_{15} en junio, X_{16} en mayo y X_{17} en abril.
- X_{18} a X_{23} : Monto del pago anterior, en que X_{18} representa el monto del pago en septiembre, X_{19} en agosto, X_{20} en julio, X_{21} en junio, X_{22} en mayo y X_{23} en abril.
- Y : Cesación de pago (1 cuando hay cesación, 0 cuando no hay cesación)

Al respecto, desarrolle los siguientes pasos para clasificar si un cliente cae o no en cesación de pago:

1. (1.0 pto.) Preprocese adecuadamente y realice un análisis exploratorio de los datos, discutiendo sus hallazgos.
2. (1.0 pto.) Proponga una red neuronal feedforward que permita clasificar a los clientes que caerán en cesación de pago. ¿Qué funciones de activación y de pérdida propone utilizar y por qué?. Discuta las **métricas de desempeño apropiadas** para este problema y justifique su selección.
3. (1.0 pto.) Implemente la red neuronal con diferentes optimizadores, utilizando fundadamente métodos de mejoramiento y regularización. Entregue un diagrama del modelo final, aplique la red sobre los datos preprocesados y compare los resultados obtenidos con distintas configuraciones.

Evaluación de la tarea

La tarea se realiza en grupos. Para la pregunta 1 se puede entregar la digitalización del manuscrito o su edición en algún procesador. Para la pregunta 2 se debe entregar el código reproducible en un cuaderno de Jupyter o en Google Colab, explicando de manera ordenada los pasos realizados. Debe incluir sus supuestos, justificaciones, procedimientos bien detallados y discusiones en el mismo cuaderno. Cite acórdemente si utiliza alguna fuente de información para realizar su trabajo.

Dado que el problema se puede abordar de diferentes maneras, se evaluará la implementación de modelos, utilidad del análisis estadístico, la coherencia del informe y las respuestas entregadas según los resultados obtenidos.

Recuerden que el foro siempre está abierto para preguntas.