Examen Introducción al Machine Learning Aplicado al Audio: Teoría e Implementación.

Profesor: Rodolfo A. Lobo

Enero 2024

1. Explique cada parte involucrada en una clasificación binaria dentro de la función Binary Cross Entropy Loss (Pérdida de Entropía Cruzada Binaria):

$$\mathcal{L}(y,\hat{y}) = -\sum_{k=1}^{n} y \cdot \log(\sigma(x_i)) + (1-y) \cdot \log(1-\sigma(x_i))$$
(1)

recuerde que en general, bajo nuestra notación $\hat{y} = \sigma(x)$. Utilice las clases $c \in \{0,1\}$ para ejemplificar que ocurre en un problema de clasificación binaria y cómo la expresión anterior nos ayuda a "castigar" el modelo cuando comete errores.

- (1pt) Dibuje cada parte de la función, es decir, la porción que mide el error de la clase 1 y la porción que mide el error de la clase 0. Expliqué por qué y cómo funciona cada parte.
- (1pt) Defina qué es Entropía, ejemplifique utilizando al menos un par de conjuntos (cómo lo visto en clases).
- 2. Explique cada parte del modelo LSTM:
 - (1pt) ¿Cuál es el rol de cada sub modelo al interior de la célula?. Detalle los modelos, funciones de activación, dimensiones de entrada y salida.
 - (1pt) Explique cuál es la principal diferencia entre este modelo y las redes neuronales convencionales. De al menos 3 posibles ejemplos de uso para este modelo.

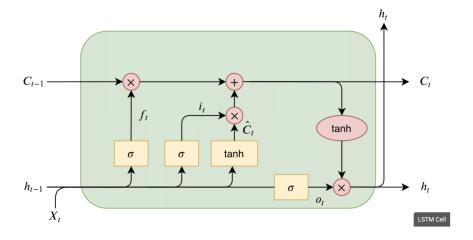


Figura 1: Arquitectura modelo LSTM

3. Dado el modelo lineal:

$$g_{\theta}(x) = \theta_1 x + \theta_0, \tag{2}$$

y la función de pérdida:

$$\mathcal{L}(\theta, x_i, y_i) = \frac{1}{n} \sum_{k=1}^{n} (g_{\theta}(x_i) - y_i)^2$$
(3)

y dados lo siguientes parámetros de inicialización:

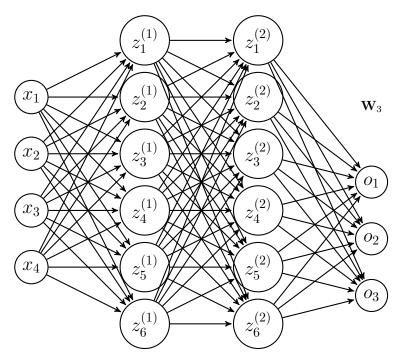
- $\alpha = 0.1$
- $\bullet \ \theta_0 = 0$
- $\bullet \ \theta_1 = 0$
- (1pt) Escriba las ecuaciones de Gradiente Descendiente, calcule las derivadas de forma explícita.
- (1pt) Realice una iteración del modelo de gradiente descendiente para todo el conjunto de entrenamiento (no estocástico):

X	У
1	1
2	0.5
3	2

Tabela 1: Datos de entrenamiento

4. Dado el siguiente diagrama de una red neuronal donde x_j con $j=1,\ldots,4$ siendo las componentes de entrada o información de entrada y además, donde el sesgo o bias es igual a cero para cada capa, es decir $\mathbf{b}_i = 0, \forall i = 1,\ldots,$ numerodecapas:





Responda las siguientes preguntas:

- (a) (2pt) Cuántas capas tiene este modelo.
- (b) (2pt) Cuántos parámetros tiene este modelo.
- (c) (2pt) Sea $z_i^{(j)}$ la salida de la neurona i en la capa j. Cuál es la dimensión de las matrices involucradas si los modelos dentro de cada potencial de activación son modelos lineales de la forma $W_i \mathbf{x} + \mathbf{b}_i$.
- (d) (2pt) Escriba las ecuaciones que representan a este modelo (ya sea de forma anidada o como un sistema de ecuaciones).

5. En relación a las redes convolucionales:

- (a) (2pt) Describa la arquitectura de una red neuronal convolucional (CNN) utilizada para la clasificación de imágenes. Utilice un dibujo de ser necesario.
- (b) (2pt) Describa brevemente los componentes clave como capas convolucionales, capas de pooling, capas completamente conectadas y el propósito de cada una. Además, explique cómo el tamaño de la imagen de entrada afecta a la arquitectura.
- (c) (1pt) Explique con un ejemplo el concepto fundamental de convolución en una CNN.
- (d) (1pt) ¿Por qué sirven para trabajar con audio?