

**INF-477 Redes Neuronales Artificiales**  
**Pauta Control 2**

## 1 Instrucciones

- Este certamen debe ser resuelto individualmente, sin apuntes, en un máximo de 30 minutos. Su nota se obtendrá como la suma de los puntos obtenidos, siendo el puntaje máximo de cada pregunta indicado al inicio de cada enunciado.
- Entregue las respuestas a cada pregunta utilizando un lápiz de tinta indeleble, con letra clara y legible. Recuerde también escribir su nombre y rol en cada hoja.
- Escriba explícitamente cualquier supuesto que crea importante y todos los pasos intermedios que sean necesarios para llegar a un resultado.

## 2 Preguntas

1. (25 pts.) Como usted sabe, la función valor absoluto no es diferenciable en 0. En consecuencia, ¿Es correcto matemáticamente hablando utilizar un método como *back-propagation*, basado en gradientes, para entrenar una red neuronal que se está regularizando con la norma  $\ell_1$ ? Si es así, ¿Cuál es el costo de esta elección?

R: Matemáticamente no es válido, por lo que usan subgradiientes para utilizar *back-propagation*, esto tiene un costo bajo ya que equivale a sumar los signos de cada componente de la columna.

2. (25 pts.) ¿Es cierto o falso que una convolución “válida” se puede obtener a partir de una convolución “completa” usando padding y/o recortando parte del resultado obtenido? Explique.

R: Si restringimos la norma de toda la matriz estaríamos queriendo minimizar también las covarianzas entre las columnas.

3. (50 pts.) ¿Es cierto o falso que una capa convolucional de la forma

$$a^{(i)} = \sigma_i(W^{(i)} * a^{(i-1)} + b^{(i)}),$$

puede implementarse usando una capa feedforward clásica de la forma

$$a^{(i)} = \sigma_i(W^{(i)} \cdot a^{(i-1)} + b^{(i)}),$$

mediante un re-dimensionamiento de la entrada y la salida? Si su respuesta es afirmativa, ¿Qué forma tiene la matriz  $W^{(i)}$ ? ¿Cuántos parámetros libres tiene con respecto a una matriz de pesos tradicional? Hint: recuerde la definición de una matriz de *Toeplitz*.

R: Aplicando la conmutatividad de la convolución, puedo transformar el input en matriz Toeplitz, y de esta forma no es necesario modificar la estructura de la matriz de pesos.