

# Redes LSTM para el reconocimiento de voz aplicado a un conjunto de dígitos

Victor Jesus Sotelo Chico<sup>1</sup>

<sup>1</sup>Universidad Nacional de Ingeniería

Seminario de Tesis II



# Contenido

- 1 Introducción
- 2 Objetivos
- 3 Marco Teórico
  - Redes Neuronales
  - Redes LSTM
- 4 Métodos de Optimización
- 5 Conclusiones y Trabajos Futuros



# Introducción

Las señales de voz proveen una gran cantidad de información a través del tiempo, el estudio de estas ha permitido el desarrollo de sistemas de reconocimiento de voz.



# Objetivos

- Conocer el proceso involucrado en el habla humana.
- Estudiar procesamiento de las señales de voz.
- Diseñar una red neuronal capaz de reconocer un conjunto de audios de números.
- Mostrar los resultados obtenidos y explicarlos basándonos en la teoría estudiada.



# Redes Neuronales Artificiales

Estas redes toman como inspiración la arquitectura del cerebro para la construcción de sistemas inteligente. Actualmente son la base para el desarrollo de la inteligencia artificial.



# Comparación neuronas biológicas y artificiales

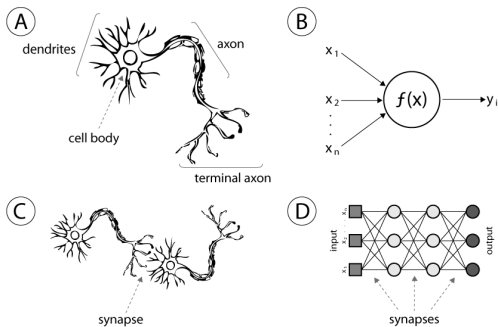


Figura: Redes neuronales biológicas y artificiales

# Redes neuronales Prealimentadas

Es un tipo de red neuronal más simple que existe. Esta red puede clasificarse en:

- Perceptron simple
- Perceptron Multicapas
- Redes neuronales convolucionales



# Esquema Redes neuronales Prealimentadas

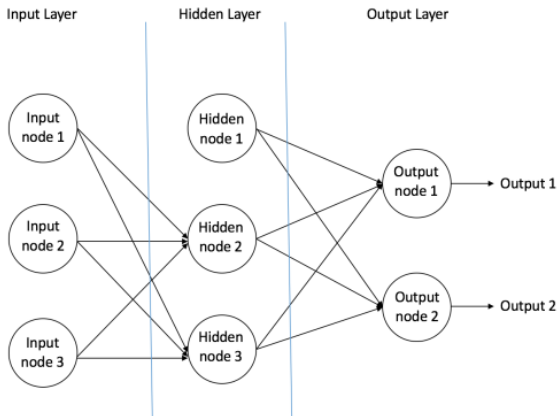


Figura: Esquema de Redes Neuronales Prealimentadas



# Back Propagation

Backpropagation  
+  
Weights Adjusted

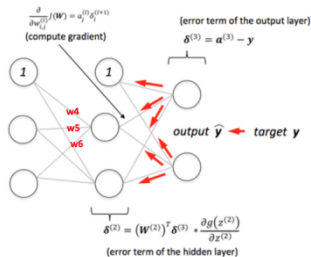


Figura: Propagación hacia atrás

# Redes Neuronales Convolucionales(CNN)

Las CNN son un tipo de redes neuronales especiales para procesar datos como imágenes. La primera CNN fue creada por Yann LeCun.



# Capas de una red neuronal convolucional

- Input Layer
- Convolutional Layer
- Pooling Layer
- Fully Connected Layer
- Output Layer



# Las Redes LSTM o Long Short Term Memory son un tipo de



## redes neuronales recurrentes

# Gradiente de Descenso

La gradiente de descenso es una forma de minimizar la función de costo  $J(\theta)$  parametrizada por los parámetros  $\theta \in \mathbb{R}^d$ .

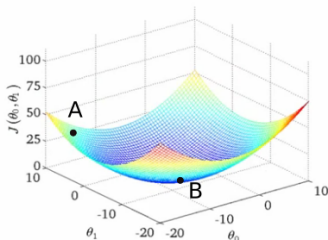


Figura: Gradiente de descenso

# Variantes de la Gradiente de Descenso

Existen 3 variantes de la gradiente de descenso:

- Batch gradient descent
- Stochastic gradient descent
- Mini-batch gradient descent



# Conclusiones

- Los métodos de optimización Adam y RMSprop obtuvieron los mejores resultados de precisión en ambas pruebas.
- A pesar de que el método de optimización Adam fue propuesto a partir del RMSprop. Adam fue superado en algunas de pruebas realizadas.
- Adam es el método que tiene un decaimiento más acelerado al calcular el error en la función de costo cross-entropy.



# Conclusiones

- Entre los métodos adaptativos Adam, RMSprop y Adagrad . Solo este último obtuvo los peores resultados, esto se debió a su dificultad de trabajar con la suma de las gradientes al cuadrado lo cual poco a poco redujo su tasa de aprendizaje.
- El RMSprop como una mejora del Adagrad, obtuvo mejores resultados que este último. Esto debido a que RMSprop trabaja con el promedio de la raíz de la gradiente anterior y tasas de decaimiento para controlar el problema de la disminución de la tasa de aprendizaje del método Adagrad.





# Trabajos Futuro

- Correcto diseño de una red neuronal convolucional.
- Obtener resultados con distintos hardwares.
- Realizar una implementación más interactiva.

