



# Instituto Politécnico Nacional Escuela Superior de Cómputo



## Comunicación Oral y Escrita

“Técnicas de NLP para la detección de errores fonéticos en la pronunciación”

**Autor: González Trejo Omar Daniel**

Correo electrónico: [ogonzalez1500@alumno.ipn.mx](mailto:ogonzalez1500@alumno.ipn.mx)

# Contenido

- Introducción
- Técnica de agrupamiento k-medias
- Redes neuronales profundas
- Método GOP
- Conclusiones

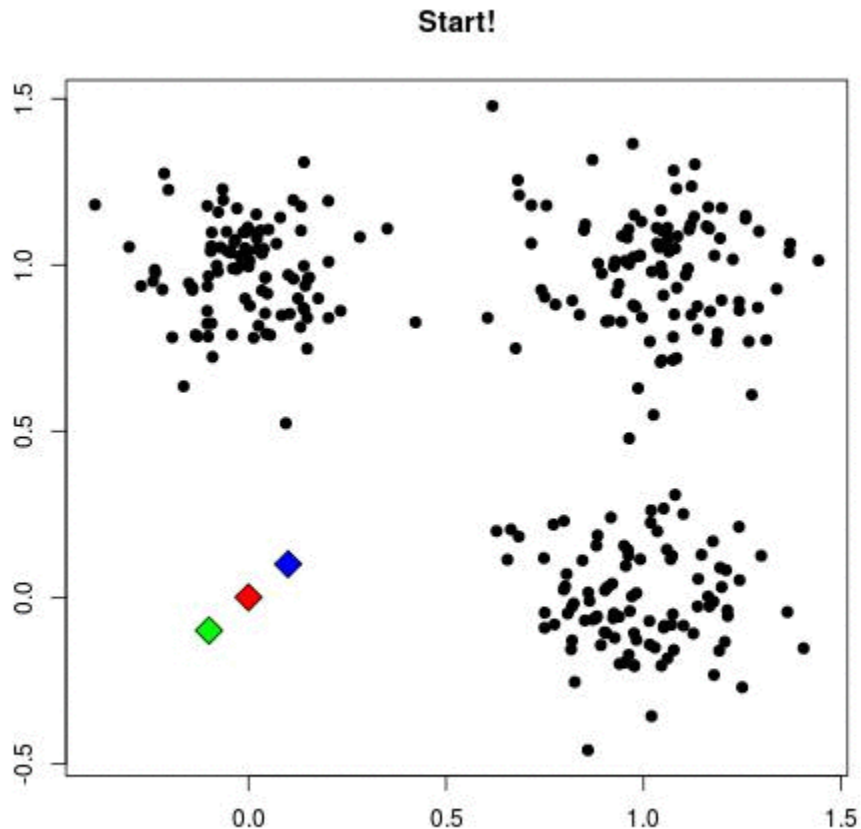
# Introducción



*Figura 1. Emisión de ondas al hablar*

- ❑ Sistemas de aprendizaje de idiomas asistido por computadora (CALL, por sus siglas en inglés)
- ❑ Técnicas para detección del lenguaje
- ❑ Modelo matemático, combinación de técnicas o pruebas

# Técnica de agrupamiento k-medias



- I. Elegir  $K$  centroides de manera aleatoria para definir el estado inicial
- II. Asignar cada punto  $X_i$  al cluster del centroide que más cerca tengan en función de la medida de distancia elegida
- III. Encontrar un nuevo centroide para cada cluster, promediando los puntos que han caído en cada cluster (**recálculo de centroides**)
- IV. Repetir pasos 2-3 hasta que en el paso 3 los centroides no varíen

Si el número de distribuciones es muy alto, el coste computacional puede ser muy grande.

Figura 2. Representación del método k-medias

# Redes neuronales profundas

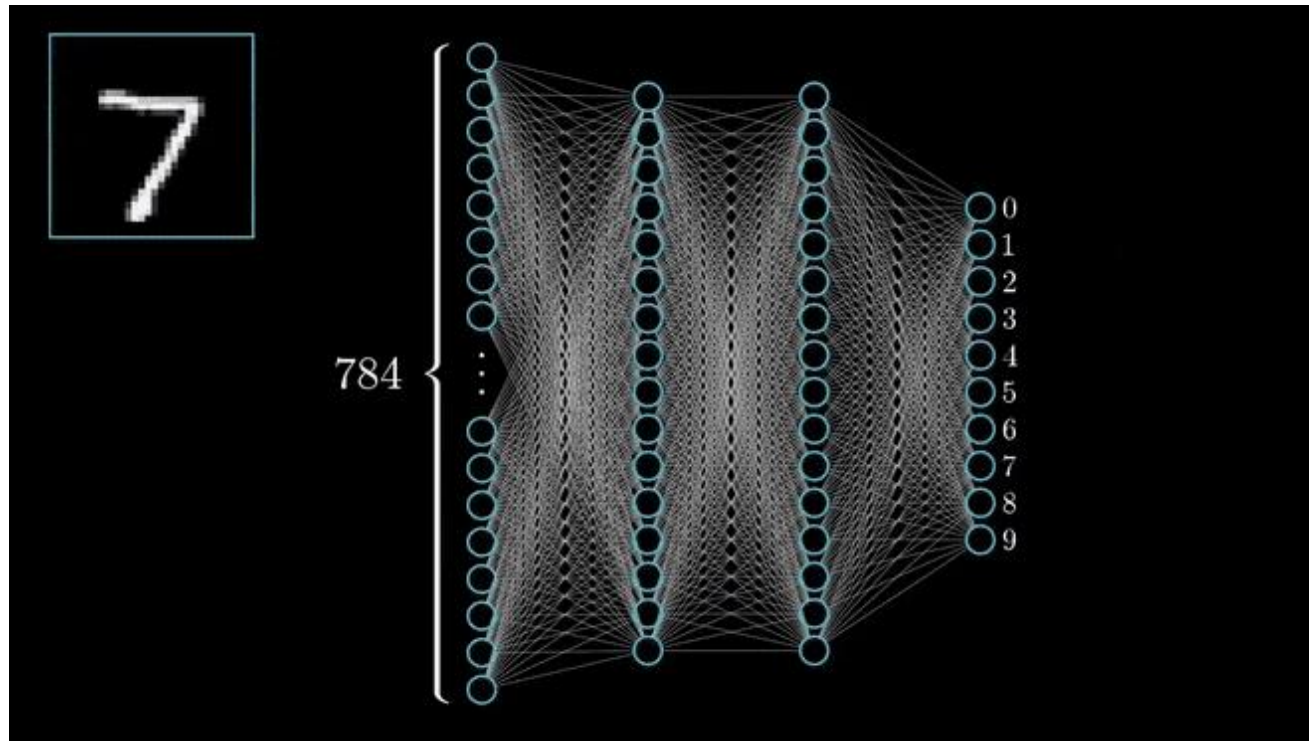


Figura 3. Red neuronal para el reconocimiento de un número escrito

- También conocido como sistemas conexionistas
- Cada enlace asigna un “peso” a la neurona
- Función limitadora o umbral (función de activación)
- La función de pérdida
- Propagación hacia atrás

# Método GOP



Figura 4. Fonema /m/

- ❑ Reconocimiento de voz forzada
- ❑ Reconocimiento de voz libre
- ❑ Se calcula una puntuación GOP de una realización específica del teléfono tomando la diferencia absoluta entre el logaritmo de la probabilidad forzada y logaritmo de la fase de reconocimiento libre
- ❑ Los teléfonos con puntuaciones GOP por encima de un umbral predefinido son probablemente mal pronunciado.

Calcula la relación de probabilidad que una realización del teléfono corresponde al fonema que debería haber sido hablado

# Conclusión

- ❑ La efectividad de cada uno de los métodos mostrados varía con respecto a la aplicación que se le de.
- ❑ Su desempeño depende del modelo matemático que se implemente.
- ❑ Es importante conocer el problema y las posibles variables que puedan afectar nuestra implementación.

# Referencias

- ❑ Apple (s.f). SmallTalk phonemes. [Figura 4]. Recuperado de <https://itunes.apple.com/us/app/smalltalk-phonemes/id384170714?mt=8>
- ❑ Ruiz, P. (2018). Clustering. [Figura 2]. Recuperado de <http://www.pabloruizruiz10.com/resources/Curso-Machine-Learning-Esp/5---Aprendizaje-No-Supervisado/Intro-Clustering.html>
- ❑ Varone, M., Mayer, D & Melegari, A. (2016). What is natural language processing? [Figura 1]. Recuperado de <https://www.expertsystem.com/natural-language-processing/>
- ❑ 3Blue1Brown (2018). Neural Network Gif. [Figura 3]. Recuperado de <https://gfycat.com/gifs/search/neural+networks>



Gracias por su atención