

# *Evaluación Crítica*

Entrenamiento de un reconocedor fonético de dígitos para el español de México usando CSLU TOOLKIT

Por: N.Munive, A vargas, B. Serridge, O. Cervantes

Computación y sistemas Vol. 3 No 2

Articulo recibido el 10 de junio , 1999; aceptado el 27 de agosto de 1999

DEIBY ALEXANDER FANDIÑO RODRÍGUEZ  
COD:256349

UNIVERSIDAD NACIONAL DE COLOMBIA  
FACULTAD DE INGENIERÍA  
MAESTRÍA EN INGENIERÍA DE SISTEMAS  
SEMINARIO DE INVESTIGACIÓN I



# 1. RESUMEN (1)

- El artículo resume la investigación realizada, sobre la implementación de un reconocedor fonético para dígitos en español, este proyecto fue implementado usando el Toolkit, el cual es un paquete de aplicaciones desarrollado específicamente para sistemas de reconocimiento de voz.

la investigación se centra en 3 partes:

1. Creación del corpus
2. Desarrollo del clasificador
3. Evaluación de resultados

# RESUMEN (2)

## 1. Creación del corpus

El corpus comprende toda la base de conocimiento del programa, esto incluye la colección de los fonemas y las palabras del grupo de muestra tomado, que en este caso fue de 50 personas, 25 mujeres y 25 hombres, el resultado de este proceso fue la colección de 1933 archivos de voz.

# RESUMEN (3)

## 2. Desarrollo del clasificador

En esta fase se explica el proceso para la clasificación de los fonemas, el cual reduce el problema y permite una mayor probabilidad que el sistema lo interprete correctamente, después de este proceso se sigue con el entrenamiento de la red neuronal de arquitectura *feed forward* de tres niveles que emplea el algoritmo *back propagation*.

# RESUMEN (4)

La red neuronal esta constituida por 130 nodos de entrada , 200 nodos ocultos y un nodo de salida el cual es la categoría fonética que se desea reconocer. El proceso consistió en dividir la señal (fonema) en frames, se calculan los vectores de características a partir del espectro de la señal (voz), se clasifican teniendo en cuenta características como energía y anchos de frecuencia, y después estos vectores se introducen a la red neuronal.

# RESUMEN (5)

Después de entrenar la red se paso a la etapa de desarrollo, en la cual se selecciona la red con mejor desempeño, lo cual se realiza evaluando el nivel de error en el proceso de reconocimiento. Se desarrollaron 3 clasificadores:

1. Clasificador independiente de contexto
2. Clasificador dependiente de contexto
3. Clasificador dependiente del contexto agrupando los fonemas en clases generales

Para esto se implementaron las pruebas con un corpus grabado por teléfono.

# RESUMEN (6)

## 3. Evaluación de resultados

En esta fase se determina cual es la mejor red, esto se realiza teniendo en cuenta el mayor porcentaje de reconocimiento en la fase de desarrollo, con un conjunto de datos desconocidos para la red. Esto dio como resultado que fue mas optima cuando se dividieron los fonemas en partes, y se concluyo que el sistema es sensible al ambiente, producto de la prueba con el corpus telefónico.

# EVALUACIÓN (1)

- METODOLOGÍA

- Se observa un proceso estructurado y claro, lo cual facilita el entendimiento del artículo.

- No presenta un marco teórico, y omite definiciones, las cuales da por hecho que el lector tiene conocimiento.

# EVALUACION(2)

- RESULTADOS

- Los resultados descritos son claros, y concluyentes, sin embargo la forma en que son presentados es muy simple, se podrían haber utilizado recursos como graficas para las comparaciones.
- Se ignoran datos de los tiempos de procesamiento de las redes, así como el numero de pruebas que se realizaron, si se realizaron mas de una.

# EVALUACION(3)

## ● CONCLUSIONES

- Las conclusiones son precisas de acuerdo a los resultados obtenidos.
- Es un artículo el cual maneja una literatura muy técnica, la cual aborda mucho de los conceptos que se manejan en el área.
- Es interesante además presenta un enfoque muy técnico, y para una audiencia que conoce del tema sería muy enriquecedor debido a los resultados que presentan.