DOCUMENTACIÓN TÉCNICA

Universidad de Bogotá Jorge Tadeo Lozano

Alvarado Becerra Ludwig
Barrios Jiménes Johan Felipe
Lis Cruz Nicolás
Vera Soto Julián David
Arquitectura Empresarial - 20251S

${\bf \acute{I}ndice}$

1. Transcripción de las entrevistas

 $\mathbf{2}$

1. Transcripción de las entrevistas

En la actividad de AVATA del curso Arquitectura Empresarial (figura 1) se dejó un archivo comprimido .rar con la transcripción de las entrevistas junto con unas etiquetas de tiempo tal como se muestra en la figura 2. El objetivo es el de tener estos textos de manera organizada para así poder diferenciar de los entrevistador(es) y el entrevistado. Para lograr esto se van a utilizar dos modelos de inteligencia artificial pre-entrenados; Pyannote.audio[1, 2], encargado de segmentar el audio por speaker capturando en qué instante de tiempo otra persona habla; Whisper[3], encargado de transcribir lo que dicen los speakers.

Evolución de la Ingeniería de Sistemas

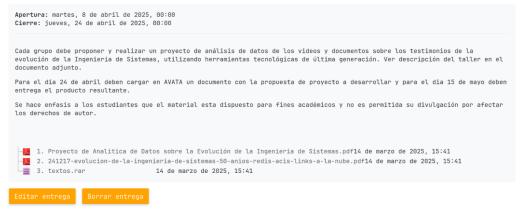


Figura 1: Actividad del curso junto con sus entregables y la transcripción de los textos.

Figura 2: Ejecución del comando bat 1974 - Roberto Pardo (transcript).txt

Referencias

- [1] H. Bredin, "pyannote.audio 2.1 speaker diarization pipeline: principle, benchmark, and recipe," in *Proc. INTERSPEECH 2023*, 2023.
- [2] A. Plaquet and H. Bredin, "Powerset multi-class cross entropy loss for neural speaker diarization," in *Proc. INTERSPEECH* 2023, 2023.

[3] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," 2022.