

Mejoras en Accesibilidad del Repositorio de la UNED basadas en IA

Resumen

La incorporación de modelos de IA a procesos de subtítulo y transcripción ha traído consigo una nueva forma de abordar problemas que de forma manual serían muy complejos y conllevarían un alto coste económico, además ha permitido mejorar en términos de Accesibilidad los contenidos de un repositorio audiovisual muy extenso, el de la UNED, en aras de favorecer la inclusión de toda la comunidad universitaria, y en especial a aquellas personas con dificultades.

Desarrollo del proyecto

La educación superior es fundamental para garantizar el cumplimiento de los derechos de las personas con discapacidad, sin embargo, las cifras sobre la presencia de estudiantes con discapacidad que cursan estudios superiores siguen siendo muy baja. Poco más de 23.000 personas con discapacidad están actualmente recibiendo algún tipo de educación superior, casi el 40% elige la UNED, lo que supone que aproximadamente, 9.000 de nuestros estudiantes declaran tener algún tipo de discapacidad.

La UNED, por tanto, se perfila como una opción para los estudiantes con discapacidad no solo por la semi presencialidad que caracteriza su forma de trabajar, sino también porque llegamos a lugares pequeños, dado el despliegue de Centros Asociados a lo largo de todo el territorio nacional, y también fuera de España, que conforman la UNED. Lugares donde una persona con discapacidad no tendría opciones de formación. De hecho, los datos indican que, en los entornos rurales, el número de estudiantes con discapacidad que acceden a los estudios superiores es muy bajo, decreciendo de forma proporcional a la densidad de la zona en cuestión. Solo la UNED da solución real a estas situaciones, dada su estructura y filosofía. Además, en la UNED, a través de UNIDIS, el Centro de Atención a Universitarios con Discapacidad, se trabaja para conseguir la flexibilización de los tiempos y las metodologías docentes y crear planes de concienciación y sensibilización que permita a nuestras y nuestros profesores, así como al resto de miembros de nuestra comunidad educativa, entender y reconocer las necesidades educativas especiales, que suelen quedar desatendidas. Parte de esa metodología flexible se soporta gracias al material audiovisual con el que trabajamos.

La UNED cuenta con un gran repositorio audiovisual donde se recogen todos los contenidos audiovisuales correspondientes a su amplia oferta académica. Durante el año 2023 se generaron más de 100.000 grabaciones con una duración de más de 82.000 horas. Todo este contenido generado a lo largo de los años y los años venideros era necesario adaptarlos adecuadamente para permitir ponerlo al alcance de todos en igualdad de condiciones.

Dado el volumen de datos manejados era impensable procesar toda esta información manualmente y los intentos de automatización previos no habían dado buen fruto.

Con el auge de la IA, y la liberación de modelos LLM, en junio de 2023, comenzamos a experimentar cómo estas nuevas técnicas podrían ayudarnos en nuestro cometido. El trabajo que presentamos a continuación es una prueba de que la IA ha venido para ayudarnos a mejorar, en este caso en el terreno de la Accesibilidad y actualmente sus resultados se han puesto al alcance de toda nuestra comunidad universitaria con muy buena aceptación.

Objetivos del proyecto

El objetivo principal de este proyecto es dotar de subtítulo y transcripción las videoclases que se generan en la UNED en el día a día, para mejorar su accesibilidad y ponerlas a disposición del alumnado con dificultades, pero también para aquellos que aun sin tener dificultades, prefieren disponer del texto para estudiar, elaborar sus apuntes u otras actividades relacionadas.

Llevando a cabo este objetivo, además, logramos acercarnos a otro importante objetivo, una responsabilidad de todos, cumplir el reglamento y las leyes vigentes en materia de accesibilidad.

Actividades llevadas a cabo

Con este objetivo y con la liberación de modelos de IA de conversión de voz a texto, se iniciaron pruebas con algunos contenidos para ver si la IA podía generar subtítulos y transcripciones de calidad. Se utilizaron modelos en los que la tasa de error era inferior al 3% utilizando audios en español, y los resultados fueron muy satisfactorios. Se analizaron gracias a la ayuda de una red de voluntarios en los que participaron docentes y alumnos de UNIDIS y sus respuestas nos animaron a continuar y a desplegar la solución para hacerla visible a toda la comunidad universitaria.

Soluciones tecnológicas implementadas

Una vez que se obtuvieron resultados satisfactorios, el principal problema que nos encontramos fue la capacidad de cómputo necesaria para llevar a cabo la inferencia. Utilizando cómputo CPU no podíamos dar cobertura a todos los contenidos que debíamos gestionar, por lo que tuvimos que pasar al uso de cómputo GPU, en nuestro caso utilizando 2 tarjetas NVIDIA L4, que nos permiten en la actualidad gestionar unas 180 horas al día de subtítulo.

Dado que actualmente se generan al día en la UNED más horas de grabación de las que podemos procesar, teniendo picos algunos días de 800 horas/día, se desarrolló un interfaz que permitiera a los usuarios solicitar las grabaciones en las que está interesado disponer de subtítulo y transcripción. De este modo se priorizan las peticiones y se da cobertura a las necesidades del alumnado con la capacidad de cómputo que tenemos.

Este interfaz se introdujo tanto en la web como en las aplicaciones móviles de la UNED de modo que cualquier usuario que acceda al contenido audiovisual, puede solicitar el subtítulo/transcripción si éste no está disponible y se pondrá a su disposición en un plazo no superior a 24-48 horas, informándole vía mail de que ya puede consultarlo.

Estos desarrollos se han llevado a cabo con frameworks como Ionic, base de datos Mysql Innodb Cluster y API Rest Laravel, manejando también Python y scripting para la parte de IA. Siempre se ha priorizado el uso de herramientas OpenSource, así como modelos IA también OpenSource. Debido al volumen de datos que manejamos las opciones de pago por uso no era una opción viable para este proyecto.

A partir de los resultados que proporcionan los modelos de IA, se han introducido mejoras sustanciales. Una vez obtenida la salida de los modelos de IA, se realizan varios postprocesos para eliminar alucinaciones en grabaciones con silencios largos o de mala calidad, detecciones incorrectas del idioma, generación de textos con variables fuera de rango, etc. y aunque aun nos queda margen de mejora, se revisan solamente el 15% de las grabaciones, el resto es un proceso totalmente automático.

Recursos utilizados para el proyecto

En el proyecto han participado activamente, UNIDIS, gestionando los materiales para las pruebas y su red de voluntarios para las revisiones. Vicerrectorado de Tecnología de la UNED e INTECCA encargados de analizar los diferentes modelos de IA, realizar las adaptaciones necesarias en dichos modelos y diferentes desarrollos para las integraciones en el portal web del repositorio audiovisual de la UNED y sus aplicaciones móviles.

En cuanto al hardware para el cómputo GPU se está utilizando un servidor PRIMERGY RX2530M6 con 2 tarjetas NVIDIA L4. Asimismo, se hace uso del resto de componentes integrados en el CPD que la UNED tiene en INTECCA (almacenamiento, servidores de producción, electrónica de red...)

Conclusiones

Principales aprendizajes de la experiencia

A pesar de que hay ciertas reticencias en el uso de la IA, con este trabajo queremos poner en relieve que, utilizada adecuadamente, puede mejorar mucho la calidad y eficiencia de ciertos procesos que, sin su ayuda, serían inabordables.

El feedback de los usuarios nos ha sorprendido gratamente. Inicialmente dimos por supuesto que lo iban a valorar más positivamente aquellos usuarios con dificultades, pero con el paso del tiempo nos hemos dado cuenta de que lo utilizan usuarios de todo tipo, con lo que estamos todavía más satisfechos del alcance logrado.

Próximos pasos. Objetivos próximos

Teniendo en cuenta la velocidad a la que se generan nuevos modelos de IA, se nos abren un montón de oportunidades para mejorar nuestros procesos. Se mencionan a continuación, algunas de las líneas de investigación que podremos abordar en el corto plazo:

- Aplicar nuevos algoritmos de IA para generar subtítulo y transcripción en multidioma de las grabaciones
- Aplicar nuevos algoritmos de IA para generar resúmenes y palabras clave de las grabaciones
- Implementar herramienta para gestión de feedback de usuarios
- Implementar subtítulo en directo en Webconferencia AVIP

Resultados alcanzados

La mejor forma de ver los resultados obtenidos y nuestro indicador para este proyecto es el número de solicitudes y por tanto subtítulos y transcripciones generadas desde la puesta en marcha en producción en diciembre de 2023. Actualmente ya se superan las 15.000 grabaciones y más de 16.500 horas subtituladas y transcritas.



Más de 15.000 solicitudes desde diciembre 2023

Figura 1: Evolución de solicitudes de grabaciones

Cabe mencionar que con las herramientas disponibles antes de poner en marcha este proyecto se realizaban alrededor de 300 adaptaciones de subtulado al año y que actualmente esperamos superar las 45.000 adaptaciones/año.

Videos Relacionados

1. El valor de UNIDIS

https://www.intecca.uned.es/portal/grabacion?ID_Sala=3&ID_Grabacion=808608&hashData=08bacf356412077d3e39a9e4549b5cfe¶msToCheck=SURfR3JhYmFjaW9uLEIEX1NhbGEs

2. Objetivo Bierzo, IA para transformar la Uned

<https://youtu.be/JWtUcoABbbE?si=mhvXe-zsCVuEKqTR>