

SOLITUD DE INTERÉS DE COLABORACIÓN

25/11/2021

PROYECTO

Análisis de los sistemas informales de orientación para estudiantes universitarios

INTRODUCCIÓN

El equipo docente-investigador del área de Métodos de Investigación y Diagnóstico en Educación (MIDE) de la Universidad de León lleva más de 20 años analizando la orientación educativa y profesional de las instituciones de educación superior, especialmente las universidades. Recientemente, hemos abierto una nueva línea de investigación con el objetivo de analizar los sistemas de apoyo y orientación informales que utilizan los estudiantes, para, en su caso, realizar recomendaciones y propuestas de mejora.

En los últimos años, los estudiantes han comenzado a utilizar sistemas 'informales' basados en la Red 2.0, concretamente foros web y redes sociales, para plantear e intentar dar respuesta a dudas y problemas relativos tanto al acceso a la formación universitaria, como al posterior paso por estos estudios e incluso la incorporación al mercado de trabajo. Precisamente, el carácter de 'informalidad' de estos mensajes hace que las demandas de los estudiantes no puedan ser atrapadas por los sistemas de diagnóstico y evaluación oficiales y, por tanto, se hace más difícil la respuesta institucional a las mismas.

OBJETIVO

Utilizando herramientas de Machine Learning y Procesamiento de Lenguaje Natural (NLP), pretendemos desarrollar un modelo de Inteligencia Artificial (IA) que, por medio del análisis de la información que se transmite en determinados foros web, identifique los problemas que preocupan a los estudiantes y que no son planteados en los canales oficiales.

MÉTODO

1. Crear una base de datos depurada con los mensajes que aparecen en una selección de foros web. El volcado de información se realizará utilizando técnicas de Web Scraping.
2. Una vez depurada esta base de datos, se utilizará el lenguaje de programación Python y algunas de sus librerías de NLP (NLTK, Spacy, ...) o algoritmos tipo LDA para analizar la información de entrada y detectar los tópicos o temas que son planteados.
3. Análisis de la información utilizando técnicas de etiquetado basado en reglas y aprendizaje supervisado para identificar los temas principales.
4. Desarrollo de modelos de aprendizaje no supervisado (o semisupervisado) para generar agrupaciones (clústers) de aquellos mensajes que no hayan podido ser etiquetados en la fase inicial. De esta forma, se podrán identificar, por medio del análisis cualitativo de los mismos, otras demandas de los estudiantes.
5. Elaboración del modelo final de IA con la información de #2 y #3.

ANTECEDENTES

Se ha llevado a cabo ya una descarga inicial de información y se ha configurado una base de datos de 133911 registros que ha sido utilizada para la creación de un lexicon, utilizando el software de análisis cualitativo MAXQDA.

Esta fase de análisis previo nos lleva a identificar dos de las características más definitorias del presente proyecto.

- A. El web scraping nos permite disponer de una gran cantidad de entradas o registros, pero desconocemos qué porcentaje de estos puede contener información útil.
- B. Las propias características formales del input del sistema (información con ambigüedad léxica y semántica, gran cantidad de palabras fuera de los vocabularios, uso de locuciones verbales o neologismos, etc.) hace que sea más complejo desarrollar modelos eficaces desde el punto de vista de su capacidad predictiva.

RECURSOS REQUERIDOS

Estamos buscando expertos en NLP que puedan realizar las tareas #4 y #5 descritas en el método.

Los expertos podrían colaborar con el proyecto de las siguientes formas:

1. Facturando el trabajo.
2. Siendo contratados con cargo al proyecto para realizar las tareas.
3. Facturando horas de formación y asesoramiento al equipo de investigación para que este pueda realizar las tareas.

NECESITAMOS

1. Presupuesto con breve descripción de tareas a realizar (antes del 1 de diciembre).

En el caso de llegar a un acuerdo inicial, se solicitaría una descripción más detallada de las tareas con el fin de incluirla en la planificación del proyecto que se va a solicitar.

CONTACTO

Javier Vidal

620869176

[Grupo ÉVORI](#)

[ResearchGate](#)

[Linkedin](#)

[Orcid](#)

Dpto. Psicología, Sociología y Filosofía
Universidad de León
ESPAÑA