

Tarea 2: Análisis sentimiento a comentarios de Misprofesores.com

Aldo Daniel Ojeda Rodriguez

30 de julio de 2024

Resumen

1. Introducción

Para el presente proyecto, se desarrolló un código que aplica técnicas de web scraping a datos obtenidos del portal **MisProfesores** el cual es una plataforma ampliamente utilizada por estudiantes de diversas instituciones mexicanas, tanto a nivel bachillerato como universitario, para consultar y evaluar la calidad de los profesores antes de decidir qué materias seleccionar.

El objetivo principal de este proyecto es extraer y analizar las opiniones y calificaciones de los profesores disponibles en **MisProfesores**, con un enfoque especial en el análisis de sentimiento. Esto permitirá identificar patrones en las percepciones de los estudiantes y poder generar un ranking de aquellas materias que contienen comentarios positivos y negativos.

2. Metodología

La metodología empleada para este proyecto incluye la utilización de técnicas avanzadas de web scraping para recopilar datos de manera automatizada de donde se seleccionaron 5 materias impartidas en la Facultad de Ciencias Físico Matemáticas:

geometria analitica	73
calculo integral	72
algebra	63
calculo diferencial	54
fisica basica	46
calculo de varias variables	41

enseguida se aplicó un modelo tipo transformer del tipo **bert-base-multilingual-uncased-sentiment** entrenado para clasificar los comentarios de los estudiantes en una escala del 1 al 5, donde 1 indica una opinión negativa y 5 representa una opinión positiva.

3. Resultados y Discusión de Resultados

Material	1	2	3	4	5
algebra	3.17 %	9.52 %	12.70 %	44.44 %	30.16 %
calculo de varias variables	12.20 %	17.07 %	14.63 %	34.15 %	21.95 %
calculo diferencial	14.81 %	16.67 %	24.07 %	29.63 %	14.81 %
calculo integral	13.89 %	16.67 %	15.28 %	22.22 %	31.94 %
fisica basica	17.39 %	15.22 %	10.87 %	21.74 %	34.78 %
geometria analitica	21.92 %	20.55 %	13.70 %	21.92 %	21.92 %

La tabla muestra la distribución porcentual de opiniones sobre materias, en una escala del 1 al 5 (1 negativo, 5 positivo). En Álgebra predominan opiniones positivas (44.44 % en nivel 4, 30.16 % en nivel 5), con pocas negativas. Cálculo de Varias Variables tiene una distribución más equilibrada, con opiniones positivas y negativas notables. Cálculo Diferencial y Geometría Analítica muestran una mayor dispersión de opiniones por otro lado Cálculo Integral destaca con un 31.94 % de opiniones muy positivas, y Física Básica tiene un alto porcentaje de opiniones positivas (34.78 % en nivel 5) pero también significativas opiniones negativas.

En resumen, Álgebra y Física Básica tienen opiniones predominantemente positivas, mientras que Cálculo de Varias Variables y Geometría Analítica muestran una mayor diversidad en las calificaciones y por otro lado la materia peor calificada seria calculo diferencial.

3.1. Conclusión

Un clasificador de sentimiento nos permite poder interpretar los sentimientos de los usuarios y darles una referencia para comparar o describir ciertos eventos o muestras de datos, en este caso el análisis de sentimiento permitió seleccionar las materias con un sentimiento generalizado positivo y distinguirla de aquellas materias que presentan sentimientos generalizados negativos.