

Reporte Detallado del Proyecto de Análisis de Sentimientos con RoBERTuito

Introducción

Este proyecto se centra en el análisis de sentimientos utilizando RoBERTuito, un modelo de procesamiento de lenguaje natural en español, para evaluar las opiniones de los estudiantes sobre las empresas donde realizaron sus prácticas. El objetivo es comprender las emociones expresadas en las evaluaciones y extraer información valiosa para mejorar la experiencia de prácticas en las empresas.

Pasos y Procesos del Proyecto

Preparación de Datos:

Archivos de Datos: Se utilizaron varios archivos CSV y XLSX, incluyendo evaluaciones de los estudiantes (Experiencia_2018_verano_2022.xlsx y Experiencia_etiquetado.xlsx), comentarios limpios (comentarios_limpios.csv), y comentarios con sentimiento analizado (comentarios_con_sentimiento.csv). Estos archivos se obtuvieron con los resultados de la limpieza de datos, el formato se decidió así ya que facilita y lleva orden sobre los resultados de cada proceso del RoBERTuito

Limpieza de Datos: Los comentarios fueron preprocesados para eliminar signos de puntuación, números y palabras vacías (utilizando Lista.txt), lo que facilitó un análisis de texto más efectivo.

Implementación de RoBERTuito:

Carga del Modelo: Se cargó RoBERTuito con transformers y pysentimiento.

Análisis de Sentimientos: Los comentarios limpios se procesaron con RoBERTuito para clasificarlos como positivos, negativos o neutrales.

Resultados: Los sentimientos analizados junto con los comentarios se guardaron en comentarios_con_sentimiento.csv.

Análisis y Agrupación por Organización:

Extracción de Sentimientos: Se extrajo información detallada de sentimiento y probabilidad de los comentarios.

Agrupación de Datos: Los comentarios se agruparon por organización para identificar las empresas con mayor número de comentarios positivos, negativos y neutrales.

Resultados y Observaciones:

Se identificaron las organizaciones con mayor número de comentarios positivos y negativos, proporcionando una visión clara de las percepciones de los estudiantes.

Estos resultados pueden usarse para realizar recomendaciones a las organizaciones sobre cómo mejorar la experiencia de las prácticas profesionales.

Resumen de Resultados y Observaciones

Organizaciones con Más Comentarios Positivos

Las organizaciones que recibieron la mayor cantidad de comentarios positivos fueron:

BIRTH GROUP: 12 comentarios positivos.

FAURECIA: 9 comentarios positivos.

FAURECIA PUEBLA: 8 comentarios positivos.

SIALICO: 7 comentarios positivos.

CARBOMEX NANOMATERIALES: 7 comentarios positivos.

Estos resultados sugieren que estas organizaciones brindan experiencias de prácticas profesionales que son bien recibidas por los estudiantes, destacando en áreas como el ambiente laboral, el aprendizaje y el apoyo al estudiante.

Organizaciones con Más Comentarios Negativos

Las organizaciones con el mayor número de comentarios negativos fueron:

ACCURATE MEDIA: 6 comentarios negativos.

BASE 10: 4 comentarios negativos.

OZONO POLARIS: 4 comentarios negativos.

HOSPITAL ANGELES PUEBLA: 4 comentarios negativos.

1ER. BOOT CAMP IBEROAMERICANO DE INNOVACIÓN ABIERTA 2021: 3 comentarios negativos.

Estos hallazgos pueden indicar áreas de mejora para estas organizaciones en términos de la experiencia de prácticas que ofrecen. Los comentarios negativos podrían estar relacionados con aspectos como la falta de apoyo, desafíos en el ambiente laboral o la no correspondencia entre las expectativas y la realidad de las prácticas.

Organizaciones con Más Comentarios Neutros

Las organizaciones con la mayor cantidad de comentarios neutros fueron:

ACCURATE MEDIA: 14 comentarios neutros.

OZONO POLARIS: 9 comentarios neutros.

1ER. BOOT CAMP IBEROAMERICANO DE INNOVACIÓN ABIERTA 2021: 8 comentarios neutros.

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA: 7 comentarios neutros.

FAURECIA: 7 comentarios neutros.

Los comentarios neutrales pueden reflejar experiencias que no fueron particularmente destacadas ni negativas. Esto puede indicar que las prácticas cumplían con los requisitos básicos, pero no ofrecían elementos excepcionales o de especial interés para los estudiantes.

Conclusión y Recomendaciones

Este análisis proporciona una visión valiosa de las experiencias de prácticas profesionales desde la perspectiva de los estudiantes. Las organizaciones pueden utilizar estos datos para entender mejor cómo sus programas de prácticas están siendo percibidos y qué pueden hacer para mejorar.

Para las Organizaciones con Comentarios Mayoritariamente Positivos: Explorar y mantener los aspectos que están funcionando bien, y considerar cómo pueden ser un modelo para otras organizaciones.

Para las Organizaciones con Comentarios Mayoritariamente Negativos: Revisar las áreas de crítica y trabajar para mejorarlas. Esto puede incluir mejorar la comunicación, ofrecer más apoyo o enriquecer las oportunidades de aprendizaje.

Para las Organizaciones con Comentarios Neutrales: Identificar maneras de hacer que las experiencias sean más atractivas y enriquecedoras para los estudiantes.

Este análisis de sentimientos ofrece un medio efectivo para evaluar y mejorar las experiencias de prácticas profesionales, beneficiando tanto a los estudiantes como a las organizaciones.

Limitaciones y Consideraciones Futuras

Validación del Modelo: No se realizó una validación formal del modelo (como una matriz de confusión) debido a la falta de un conjunto de datos etiquetado para comparar.

Interpretación de Resultados: Los resultados dependen de la precisión de RoBERTuito y de la calidad del preprocesamiento de datos.

Referencias:

Pérez, J. M., Furman, D. A., Alemany, L. A., & Luque, F. (2021). Robertuito: a pre-trained language model for social media text in spanish. *arXiv preprint arXiv:2111.09453*.