

## Anexo I. Registro del Título del Trabajo Fin de Grado (TFG)

NOMBRE DEL ALUMNO: Nadia de Santiago

PROGRAMA: ADE + RRII

GRUPO: 5ºE6

FECHA:13/10/2024

Director Asignado: Calvo Pascual

Apellidos

, Luis Ángel

Nombre

Título provisional del TFG:

Sentiment analysis sobre empresas familiares en la Comunidad de Madrid.

**ADJUNTAR PROPUESTA (máximo 2 páginas: Índice provisional, objetivos, metodología y bibliografía)**

Firma del estudiante:

Nadia de Santiago Escamilla



Fecha: 14/10/2024

## Índice

---

Agradecimientos

Índice de ilustraciones

Índice de tablas

Índice de anexos

Acrónimos

Resumen ejecutivo y *abstract*.

1. Introducción
  1. Contexto
  2. Motivación
  3. Definición del problema
  4. Objetivos y alcance del proyecto
  5. Metodología y estructura del documento
2. Marco teórico
  1. Empresas familiares en Madrid y su rol en el sector hostelero
  2. Antecedentes
3. Análisis cuantitativo y cualitativo
  1. Análisis de sentimientos
    1. Construcción del dataset
      1. Recopilación de datos
      2. Extracción y limpieza de datos
    2. Transformación del dataset
      1. Preprocesamiento de los textos y transcripciones
      2. Feature selection
  2. Técnicas de Machine Learning aplicadas
    1. Naive Bayes
    2. Support Vector Machines (SVM)
    3. Redes neuronales
    4. Mutual information
    5. Regression Discontinuity
    6. Imbalanced ML
    7. Deep Gaussian Process
4. Resultados y discusión
  1. Interpretación de los resultados
  2. Comparación resultados
  3. Implicaciones y recomendaciones para empresas familiares en hostelería
5. Conclusiones
  1. Resumen de los hallazgos
  2. Problemas y limitaciones del estudio
  3. Próximas líneas de investigación
6. Declaración de uso de herramientas de Inteligencia Artificial Generativa
7. Bibliografía
8. Anexos

## Objetivos

---

Las empresas familiares constituyen el núcleo de la economía nacional y madrileña, representando el 85% del tejido empresarial y generando el 65% del empleo (Cámara Madrid, 20 de marzo de 2024). En concreto, el sector hostelero es uno de los motores fundamentales de riqueza y empleo de Madrid. Las empresas familiares pertenecientes a este sector deben hacer frente a problemas como la falta de digitalización, ausencia de reformas o encarecimiento de la financiación (ADEFAM, 11 de julio de 2024). Sin embargo, muchas de las investigaciones realizadas sobre los obstáculos a los que estas se enfrentan estas empresas están limitados por el uso de métodos tradicionales como encuestas. Por ello, este trabajo de investigación tiene como objetivo principal identificar los problemas reales a los que se

enfrentan estos empresarios a través de un análisis de sentimientos de vídeos y comentarios recopilados de plataformas como Youtube y Twitter. Para alcanzarlo, se plantean los siguientes objetivos específicos:

1. Identificar problemas ocultos o que no son detectados por los métodos tradicionales de estas empresas en la Comunidad de Madrid mediante el análisis de sentimientos y técnicas avanzadas de Machine Learning en comentarios extraídos de redes sociales.
2. Demostrar que los desafíos más graves a los que se enfrentan estas empresas no se han reflejado adecuadamente en estudios anteriores y que, por ende, no se han implementado medidas efectivas que contribuyan a una mejor operatividad y rendimiento de estas empresas.
3. Evaluar el impacto de medidas gubernamentales en la rentabilidad y supervivencia de las empresas familiares a través de la técnica de regresión discontinua.
4. Aplicar algoritmos de ML con el objetivo de clasificar y analizar los datos obtenidos además de generar a partir de ellos predicciones sobre cómo podrían los empresarios familiares hacer frente a futuros desafíos en el sector hostelero.
5. Desarrollar propuestas y estrategias basadas en los hallazgos encontrados para ayudar a las empresas familiares hosteleras a superar los obstáculos identificados.

## Metodología

Se empleará una combinación de técnicas cualitativas y cuantitativas. La primera fase consistirá en la recopilación de datos procedentes de redes sociales (Youtube y Twitter), para capturar comentarios y opiniones de empresarios familiares del sector hostelero. Una vez encontrados los datos, se procederá a hacer una limpieza y transformación de los datos para poder crear una base de datos adecuada para el análisis. En el sentiment analysis realizado se utilizarán algoritmos como Naive Bayes y Support Vector Machines para clasificar los comentarios en aras de identificar posibles patrones o problemas subyacentes. En la siguiente fase, se aplicarán diferentes técnicas de machine learning como Imbalanced ML para corregir posibles desequilibrios, Mutual Information, para distinguir aquellas variables más influyentes en la viabilidad de las empresas familiares, y un modelo de Regresión Discontinua para examinar el impacto de medidas gubernamentales en las empresas de este sector. Por otro lado, también se utilizará un modelo de redes neuronales para captar relaciones más complejas en los comentarios analizados y el Deep Gaussian process para poder predecir posibles escenarios futuros a partir de la información obtenida. Adicionalmente, se complementará el análisis con encuestas o entrevistas a empresarios familiares en el sector de la hostelería para comprobar si los resultados obtenidos en el sentiment analysis coinciden con las experiencias reales de muchos empresarios en el sector.

## Bibliografía

Alhujaili, R. F., & Yafooz, W. M. (2021, March). Sentiment analysis for youtube videos with user comments. In 2021 International conference on artificial intelligence and smart systems (ICAIS) (pp. 814-820). IEEE. Recuperado de [https://ieeexplore.ieee.org/abstract/document/9396049?casa\\_token=Y7P-YbJ3oScAAAAA:wHsRYCnJVpYrKXuFTipJHRLgu8iD54b-WXuBtPDfMFv9npsBxiQULf5rFmUOBHiSLhichpuJug](https://ieeexplore.ieee.org/abstract/document/9396049?casa_token=Y7P-YbJ3oScAAAAA:wHsRYCnJVpYrKXuFTipJHRLgu8iD54b-WXuBtPDfMFv9npsBxiQULf5rFmUOBHiSLhichpuJug)

Chamorro Wolters, A. C. (2019). El impacto de las empresas familiares en la economía de la Comunidad de Madrid. Recuperado de <http://hdl.handle.net/11531/33076>

Claver, E., Rienda, L., & Pertusa, E. M. (2004). Un marco teórico para la continuidad de la empresa familiar desde un punto de vista estratégico. *Esic market*, 35(118). Recuperado de <https://revistasinvestigacion.esic.edu/esicmarket/index.php/esicm/article/view/27>

González Jiménez de la Espada, G. (2013). Evolución y retos del sector hotelero en España. Recuperado de <https://hdl.handle.net/20.500.14468/21171>

Singh, R., & Tiwari, A. (2021). Youtube comments sentiment analysis. *International Journal of Scientific Research in Engineering and Management (IJSREM)*, 5(5), 1-11. Recuperado de [https://www.researchgate.net/profile/Ritika-Singh-38/publication/351351202\\_YOUTUBE\\_COMMENTS\\_SENTIMENT\\_ANALYSIS/links/6092e8b2299bf1ad8d7d79a7/YO-UTUBE-COMMENTS-SENTIMENT-ANALYSIS.pdf](https://www.researchgate.net/profile/Ritika-Singh-38/publication/351351202_YOUTUBE_COMMENTS_SENTIMENT_ANALYSIS/links/6092e8b2299bf1ad8d7d79a7/YO-UTUBE-COMMENTS-SENTIMENT-ANALYSIS.pdf)