

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

NOMBRE DEL ALUMNO:
Carmen Vadillo Caballero

PROGRAMA: E3-Analytics

FECHA: 23 de octubre 2024

Director Asignado: Barcos Redín, Lucía

Título provisional del TFG:

Análisis de los principales riesgos percibidos por las empresas del mercado continuo español detallados en sus informes de gestión utilizando text mining y topic modeling.

ADJUNTAR PROPUESTA (máximo 2 páginas: objetivo, bibliografía, metodología e índice preliminares)

Firma del estudiante: Carmen V

Fecha: 23 de octubre 2024

JUSTIFICACION/MOTIVACIÓN DEL TEMA

En los informes integrados o de gestión que las empresas publican cada año podemos identificar los riesgos y desafíos a los que se enfrentan las organizaciones en los distintos sectores económicos. Ante un entorno tan dinámico, con cambios rápidos en tecnología, situaciones geopolíticas inestables, desafíos medioambientales y presiones sociales, es de suma importancia comprender los riesgos a futuro percibidos por las principales empresas del país. Mediante el uso del text mining y Topic Modeling con el algoritmo LDA, podemos detectar patrones temáticos en estos informes, pudiendo identificar y clasificar riesgos a lo largo del tiempo y por sectores. Aplicar estas herramientas puede resultar muy útil en la toma de decisiones estratégicas, ya que nos puede ofrecer una visión clara de cuáles son los principales riesgos (políticos, medioambientales, tecnológicos o sociales) que perciben las empresas españolas, como están respondiendo a estos retos y que tendencias se pueden prever a futuro.

OBJETIVO GENERAL Y ESPECÍFICOS

El objetivo del TFG es analizar los informes de gestión/integrados de las empresas del mercado continuo español utilizando Topic Modeling con LDA para identificar y clasificar los riesgos empresariales más relevantes en los últimos 5 años, segmentándolos por sectores y tipos de riesgo (tecnológico, geopolítico, financiero, medioambiental, social, de gobierno...).

Objetivos Específicos:

1. Identificar los riesgos mencionados en los informes integrados de cada empresa del mercado continuo español.
2. Determinar cuáles son los riesgos más repetidos por sector empresarial.
3. Analizar la evolución de los riesgos empresariales a lo largo de los últimos 5 años.
4. Establecer patrones sobre los tipos de riesgos más destacados en función de las características del sector y del entorno.
5. Prever tendencias a futuro.

METODOLOGÍA

1. **Recopilación del Corpus:** Reunir los informes de gestión/integrados publicados anualmente por empresas del mercado continuo español durante los últimos 5 años.
2. **Preprocesamiento del Corpus**
3. **Topic Modeling con LDA:** Aplicar el algoritmo LDA, definir cantidad de topics, etiquetar topics.
4. **Análisis Sectorial:** Clasificar las empresas por sectores económicos y analizar los riesgos por sector.
5. **Análisis Temporal:** Comparar la aparición y evolución de tópicos año a año para identificar tendencias y cambios en la percepción del riesgo.
6. **Visualización y Presentación de Resultados:** Usar herramientas de visualización de datos (gráficos de barras, líneas de tiempo..) para representar la frecuencia y evolución de los riesgos por sector.

7. **Interpretación de los resultados en el contexto actual:** considerar factores externos que puedan influir en la percepción del riesgo (crisis, avances tecnológicos, cambios legislativos).

ÍNDICE

1. **Introducción**
2. **Marco Teórico**
 - Informes integrados/gestión de las empresas.
 - Text Mining, Topic Modeling y el algoritmo LDA.
3. **Metodología**
 - Descripción del corpus.
 - Procesamiento de datos y preprocesamiento.
 - Aplicación del LDA.
 - Análisis sectorial y temporal.
4. **Resultados**
 - Descripción y etiquetado de los tópicos encontrados.
 - Análisis de los riesgos por sector.
 - Evolución temporal de los riesgos más relevantes.
 - Visualización de tendencias y patrones.
5. **Análisis**
 - Interpretación de los resultados en relación con el contexto actual.
 - Implicaciones para la toma de decisiones estratégicas.
6. **Conclusiones**
7. **Bibliografía**

BIBLIOGRAFÍA

Galindo-Vargas, L. M., & Delgado-Sánchez, V. P. (2021). Informes integrados: Una perspectiva sostenible para empresas. *Reflexiones Contables*, 4(1), Enero-Junio. <https://revistas.ufps.edu.co/index.php/RC/article/view/3821/5663>

Inditex. (2023). *Informe de Gestión Integrado 2023*. https://static.inditex.com/annual_report_2023/es/Informe_de_Gestion_Integrado_2023.pdf

Kalra, P. y Kumar, L. (2013). *Text Mining: Concepts, Process and Applications*. *Journal of Global Research in Computer Science*. 4(3). https://www.researchgate.net/publication/277160258_TEXT_MINING_CONCEPTS_PROCESS_AND_APPLICATIONS

PWC. (s.f.). *El Informe Integrado: La mejor información para una mayor ventaja competitiva*. <https://www.pwc.es/es/soluciones/auditoria/assets/el-informe-integrado.pdf>

