

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

NOMBRE DEL ALUMNO: Miriam González-Alemán

PROGRAMA: Business Analytics GRUPO: E2+BA A

FECHA: 21/10/2024

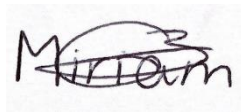
Director Asignado: Coronado Vaca María
Apellidos Nombre

Título provisional del TFG-BA:

Prediction of M&A Participants Using Machine Learning Algorithms and NLP techniques

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

Firma del estudiante:



Fecha: 21/10/2024



Faculty of Economics and Business Administration

Prediction of M&A Participants Using Machine Learning Algorithms and NLP techniques

Author: Miriam González-Alemán

Director: María Coronado Vaca

Abstract

1. Introduction	1
1.1 <i>Objectives</i>	1
1.2 <i>Motivation for the study</i>	1
1.3 <i>Methodology</i>	1
1.4 <i>Structure</i>	3
2. State of the Art.....	3
3. Empiric Analysis	3
3.1 <i>Dataset</i>	3
3.2 <i>Models</i>	3
4. Results and Conclusions	3
5. Bibliography	4
6. Annex.....	6

1. Introduction

1.1 Objectives

The primary objective of this study is to develop predictive models that identify potential targets in mergers and acquisitions (M&A) by leveraging machine learning algorithms for structured numerical data and natural language processing techniques, specifically transformers, for unstructured textual data. This research aims to assess the effectiveness of these methods in predicting which companies are likely to be pursued for acquisition by analyzing financial metrics alongside relevant textual information. By uncovering insights that indicate acquisition potential, the study seeks to provide valuable guidance for stakeholders in the M&A process, ultimately enhancing their ability to identify promising acquisition targets.

1.2 Motivation for the study

Mergers and acquisitions (M&A) are critical events in the corporate landscape, often resulting in significant market shifts. Accurate predictions of potential M&A participants can offer essential insights for investors, financial analysts, and corporate strategists. However, existing models for M&A predictions primarily rely on financial indicators, frequently overlooking the rich information embedded in unstructured data, such as news articles and company statements.

With recent advancements in machine learning (ML) and natural language processing (NLP), particularly through transformer-based models like BERT and GPT, it is now feasible to analyze extensive volumes of textual data to extract insights that traditional models may miss. By integrating quantitative financial data with qualitative textual information, this thesis aims to enhance the precision and reliability of M&A prediction models, providing a more comprehensive understanding of the factors influencing acquisition activities.

1.3 Methodology

This thesis adopts a quantitative approach, combining machine learning for numerical data and natural language processing (NLP) for text analysis to predict M&A participants. Both financial metrics and textual information are analyzed to provide a comprehensive model.

The structured financial data, such as company financials and market indicators, will be collected, cleaned, and preprocessed to handle issues like missing values. Machine learning models, including logistic regression, random forests, and gradient boosting machines (GBMs), will be used to identify patterns in the data. The models will be trained and evaluated using standard performance metrics like accuracy and precision.

In parallel, unstructured textual data from news articles and reports will be processed using transformer-based NLP models like BERT. After preprocessing, sentiment analysis and entity recognition will be applied to extract insights from the text. These insights will be transformed into embeddings, which will serve as additional features for the predictive model.

The final step involves integrating the financial data and textual embeddings into a unified model. This combined approach will improve prediction accuracy and will be evaluated using appropriate testing methods. The analysis will be conducted using Python libraries, ensuring a robust and data-driven methodology.

5. Bibliography

- Aramyan, H. (2021, September 21). Prediction of M&A targets to generate portfolio returns. *LSEG*. <https://developers.lseg.com/en/article-catalog/article/prediction-of-m-and-a-targets-to-generate-portfolio-returns#TheoreticalBackground>
- Aramyan, H. (2022, January 10). Predicting M&A Targets Using ML: Unlocking the potential of NLP based variables. *LSEG*. <https://developers.lseg.com/en/article-catalog/article/predicting-MnA-targets-using-ML-Unlocking-the-potential-of-NLP-variables>
- Hajek, P., and Henriques, R. (2024). Predicting M&A targets using news sentiment and topic detection. *Technological Forecasting and Social Change*, 201, 123270. <https://doi.org/10.1016/j.techfore.2024.123270>
- Katsafados, et al. (2024). Machine Learning in U.S. Bank Merger Prediction: A Text-Based Approach. *European Journal of Operational Research*, 312(2), 783–797. <https://doi.org/10.1016/j.ejor.2023.07.039>
- Katsafados, et al. (2021). Using textual analysis to identify merger participants: Evidence from the U.S. banking industry. *Finance Research Letters*, 101949. <https://doi.org/10.1016/j.frl.2021.101949>.
- Moriarty, R., Ly, H., Lan, E., & McIntosh, S. (2019). Deal or No Deal: Predicting Mergers and Acquisitions at Scale. 2019 IEEE International Conference on Big Data (Big Data), 5552-5558. <https://doi.org/10.1109/BigData47090.2019.9006015>
- Nicholas Center for Corporate Finance & Investment Banking (2019, December 15). From ML to M&A. Ten M&A Target Predictions through a Machine Learning Model. Wisconsin School of Business

Rodrigues, B. D., & Stevenson, M. J. (2013). Takeover prediction using forecast combinations. *International Journal of Forecasting*, 29(4), 628-641.

Routledge, B. R., Sacchetto, S., & Smith, N. A. (2013). Predicting merger targets and acquirers from text. In Carnegie Mellon University working paper.