Cogent Business & Management

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/oabm20

Fine-tuning transformer models for M&A target prediction in the U.S. ENERGY sector

Inés Rodríguez-Muñoz-de-Baena , María Coronado-Vaca & Esther Vaquero-Lafuente

To cite this article: Inés Rodríguez-Muñoz-de-Baena , María Coronado-Vaca & Esther Vaquero-Lafuente (2025) Fine-tuning transformer models for M&A target prediction in the U.S. ENERGY sector, Cogent Business & Management, 12:1, 2487219, DOI: 10.1080/23311975.2025.2487219

To link to this article: https://doi.org/10.1080/23311975.2025.2487219

Published online: 09 Apr 2025.

COGENT BUSINESS & MANAGEMENT
2025, VOL. 12, NO. 1, 2487219
https://doi.org/10.1080/23311975.2025.2487219

BANKING & FINANCE | RESEARCH ARTICLE
OPEN ACCESS
Check for updates

Fine-tuning transformer models for M&A target prediction in the U.S. ENERGY sector

Inés Rodríguez-Muñoz-de-Baena[a], María Coronado-Vaca[a] and Esther Vaquero-Lafuente[a,b]

[a]ICADE School of Economics and Business Administration, Universidad Pontificia Comillas, Madrid, Spain; [b]Instituto de Investigación Tecnológica (IIT), Universidad Pontificia Comillas, Madrid, Spain

ABSTRACT
This study explores the application of transformer models directly for classification in predicting mergers and acquisitions (M&A) targets within the U.S. energy sector. The primary objective is to evaluate the capability and performance of various transformer-based models in directly predicting M&A target companies, while the secondary objective investigates the relationship between target companies and renewable energy terminology in their annual reports. We present a novel approach to predicting M&A targets by utilizing cutting-edge Natural Language Processing (NLP) techniques, such as fine-tuned transformer LLMs (Large Language Models) for direct classification. We analyze textual data from 200 publicly-listed US energy companies' SEC-filings and employ FinBERT, ALBERT, and GPT-3-babage-002 as predictive models of M&A targets. We provide empirical evidence on LLMs' capability in the direct classification of M&A target companies, with FinBERT utilizing oversampling, being the top-performing model due to its high precision and minimized false positives, critical for precise financial decision-making. Additionally, while the study revealed key differences in target and non-target report characteristics, it finds no significant evidence that M&A target companies use more renewable energy-related terminology. It is the first paper applying fine-tuned transformer-LLMs to predict M&A targets, effectively showcasing their capability for this task of direct classification as predictive models.

ARTICLE HISTORY
Received 4 November 2024
Revised 28 February 2025
Accepted 24 March 2025

KEYWORDS
Mergers and acquisitions (M&A); renewable energy; takeover target prediction; green M&A; natural language processing (NLP); transformer models; large language models (LLM)

SUBJECTS
Environmental Economics; Business, Management and Accounting; Finance; Artificial Intelligence; Algorithms & Complexity; Computer Science (General); Corporate Finance

1. Introduction

In recent decades, the energy sector has undergone significant transformations (Bhutada, 2022). By the year 2000, fossil fuels (coal, oil, and gas) accounted for 77% of the energy mix (Bhutada, 2022; IEA, 2020). Between 1995 and 2015, energy demand soared from 8589 to 13,147 million tonnes of oil equivalent (Mtoe), exerting increased pressure on the environment through heightened CO_2 emissions (Ahmad & Zhang, 2020). Emissions are led by China, followed by the United States, India, and the European Union (Global Carbon Project, 2023). This has triggered significant sustainability concerns, particularly in the EU and USA, where emissions have seen a notable decline of 28% and 18%, respectively, between 2000 and 2023 (Global Carbon Project, 2023). In contrast, China and India have both witnessed emissions growth rates soar over 200% during the same period, fueled by economic expansion and industrialization.

The surge of sustainability-focused initiatives has yielded significant tangible outcomes within the energy sector, encouraging the adoption of renewable energy sources and redirecting investments and consumption patterns toward these sources. The landmark signing of the Paris Agreement, an international treaty aimed at limiting global warming to 2 °C over pre-industrial levels, by numerous countries including the United States, coupled with the adoption of the 2030 Agenda for Sustainable Development in 2015, has laid the foundation for sustainable development (Ahmad & Zhang, 2020; Andriuškevičius &

CONTACT María Coronado-Vaca mcoronado@comillas.edu ICADE School of Economics and Business Administration, Universidad Pontificia Comillas, Alberto Aguilera 23, Madrid, Spain

Štreimikienė, 2021). This joint effort for decarbonization, bolstered by government backing, constitutes a major tailwind accelerating renewable energy adoption (van Nieuwenhuijzen et al., 2023).

Recent geopolitical instability derived from the Ukraine and Israel-Hamas wars has contributed to oil and gas price inflation and volatility. This, combined with increasing general environmental awareness of end-consumers, as well as the demonopolization and liberalization of energy markets, has further catalyzed the consumption of electricity and raised concerns over energy security (Energy Institute, 2024). Additionally, technological advancements, including improved efficiency, storage solutions, and digitalization through smart grids, have made renewable energy more cost-effective and accessible (Andriuškevičius & Štreimikienė, 2021; Bouchard et al., 2023). As a result of this market environment, rapidly expanding renewable energies such as solar, wind, and hydropower are projected to overtake oil as the primary source of energy by 2050 (EIA, 2021).

As industry trends evolve and companies aim to stay updated and competitive, so do the M&A trends in this industry and the characteristics of selected M&A targets. Some of the most relevant external factors influencing M&A in the energy sector specifically include globalization, technological developments, global geopolitical turmoil, and the global energy transition (Niemczyk et al., 2022). Shen et al. (2021) found that geopolitical risk promotes the M&As of listed companies in the energy and electric power industries. There are also internal motives driving M&A activity and consolidation in the energy sector, such as acquiring technological capabilities, increasing market share, incorporating new products into your offering, entering a different market, or eliminating a competitor (Andriuškevičius & Štreimikienė, 2021, Galperina & Klen, 2017).

With climate change and sustainability taking centre stage of global concerns, investment, and corporate strategies are increasingly aligning with Environmental, Social, and Governance (ESG) principles (Barros et al., 2022; Ding et al., 2024). As a result, M&A motives in the energy sector have undergone a significant shift, transitioning from market positioning to resource diversification objectives. Notably, there has been a discernible move away from traditional oil-focused targets towards green energy investments, making the green energy transition the most common reason for M&A in the sector in recent years (Niemczyk et al., 2022). Hawkes et al. (2023) empirical analysis of a twenty-year history of M&A in the energy sector found that international oil companies have significant investment potential and are today exploring the renewables space and appear to be valuing innovation based on renewables on a subset of their business.

In this rapidly evolving context, inorganic growth is typically preferred as it allows for faster agility in adapting, resulting in renewable energy companies emerging as the primary targets for M&A transactions (Andriuškevičius & Štreimikienė, 2021). This is reflected in valuation multiples of companies operating in the utilities sector, which saw an increase from 10.5x EV/EBITDA in December 2020 to 11.7x in December 2021 (KPMG, 2022). In stark contrast, oil and gas valuation multiples experienced a decline from 8.4x to 6.9x during the same period (KPMG, 2022). Moreover, within the utilities sector, renewable energy companies are trading at higher multiples, reaching 30.9x EV/EBITDA by April 2023 according to Deloitte's Q1 2023 Energy Quarterly Update (Deloitte, 2023). This trend reflects growing investor interest and confidence in the long-term prospects of renewable energy investments.

In 2022, 1.6k M&A deals were completed in the USA, achieving a combined value of $1.4 trillion, of which 222 were energy-related transactions (Kaske et al., 2023). This figure translates to approximately 5% of the nation's total gross domestic product (GDP), estimated at $28 trillion, underlining the substantial contribution of M&A activity to the country's economy (IMF, 2024). According to the Institute of Mergers, Acquisitions & Alliances (IMAA, 2023), the energy and power industry has played a prominent role in global M&A activity, accounting for 7.2% of all deals since 1985 and ranking 7th among all industries. Impressively, in terms of total deal value, this sector surpasses all others except for financial institutions, representing 13.7% of total M&A value (IMAA, 2023). Moreover, EY has identified energy as 'the sector to watch in 2024' regarding M&A activity, indicating its strategic importance and potential for growth (Kaske et al., 2023).

As a result of the situation described above, predicting M&A targets in the US energy sector is of significant relevance for many stakeholders. Managers can leverage this timely information to identify appropriate potential targets in their strategic corporate investments toward green transition and green innovation, and even to implement an appropriate takeover defense strategy if necessary. Regulators can benefit from this timely prediction to assess antitrust compliance, investigate suspected insider trading, and promote the economy's green transition. Additionally, accurately identifying potential acquisition targets can provide a competitive edge for investors, as the announcement of M&A activities often leads to an increase in the target company's stock price (Adnan & Hossain, 2016; Bhagat et al., 2005; Campa & Hernando, 2004; DeLong, 2001; Goergen & Renneboog, 2004; Jarrell et al., 1988, Dodd & Ruback, 1977; Leledakis & Pyrgiotakis, 2022; Teti & Tului, 2020). There is also evidence of the wealth effects of M&A on target firm bondholders experiencing significantly positive announcement period returns (Billett et al., 2004; Chen et al., 2020). Therefore, being able to predict such events can result in substantial financial gains and form the basis of a value-based trading strategy (Easterwood et al., 2024; Kedia & Zhou, 2014).

Despite this relevance of the M&A target prediction, Tunyi (2021), in his historical review of fifty years of voluminous research on takeover target prediction that partitioned prior studies in the area into four distinct eras, found a significant decline in the number of papers exploring issues around takeover prediction in their identified fourth and last era (2010–2018), making a call for new research in this area. In addition, in recent decades, a growing literature has emerged that relies on Machine Learning (ML) for the identification and prediction of M&A targets in different sectors (Brar et al., 2009; Espahbodi & Espahbodi, 2003; Meghouar & Ibrahimi, 2020, among others; Ouzounis et al., 2009; Pasiouras & Tanna, 2010; Zanakis & Zopounidis, 1997). Existing studies to predict M&A targets using ML models primarily use traditional quantitative financial data/variables (Tunyi, 2021) as inputs for models such as decision trees-DT, random forest-RF and ensemble models (Aramyan, 2021; Bourne et al., 2019; Espahbodi & Espahbodi, 2003; Wei et al., 2008), support vector machine-SVM (Ouzounis et al., 2009), or neural networks (Sen & Gibbs 1994; Liu et al., 2007; Ouzounis et al., 2009; Rodrigues & Stevenson, 2013; Bourne et al., 2019; Anagnostopoulos & Rizeq, 2019;2021) among many others. Nevertheless, over the last five years, the latest AI advancements for this task increasingly involve text mining and Natural Language Processing (NLP) techniques, which can analyze unstructured text to uncover insights that are not evident from quantitative data alone. It is precise since the fourth and last era identified by Tunyi (2021), since 2019, over the last five years, that this growing literature has appeared. However, this takeover prediction literature based on textual analysis, although growing, is still scarce and underdeveloped, as will be shown in the literature review section, with only six recent studies (Aramyan, 2022; Hajek & Henriques, 2024; Katsadados et al., 2024; Katsafados et al., 2021; Moriarty et al., 2019; Routledge et al., 2017) to the best of our knowledge. None of those six papers focus specifically on the U.S. energy sector, and all of them except Moriarty et al. (2019) combine textual data with financial variables. Moreover, what is more relevant, only two of these six papers (Aramyan, 2022; Hajek & Henriques, 2024) use LLM-transformer models but they are not used as direct predictive models of M&A targets, but for previous sentiment analysis to construct sentiment-based variables that the authors include in their prior 'traditional' ML predictive models of M&A targets (which baseline form used only quantitative financial variables) and analyze the contribution of these sentiment-based variables obtained through LLM-transformers to the performance of their prior 'traditional' ML predictive models. Thus, no prior study exists in the literature applying LLM-transformers for direct classification of M&A target identification.

Driven by this motivation, and to respond to this research gap in the literature, in our work, we introduce a novel approach and apply transformer models directly for classification in predicting M&A targets within the U.S. energy sector. We assess the capability and performance of various transformer-based models in directly predicting M&A target companies. To this end, we analyze textual data from 200 publicly listed US energy companies' SEC filings (specifically, we use the disclosure in the firm's Management Discussion and Analysis (MD&A) section of the annual Form 10-K filing) and employ FinBERT, ALBERT, and GPT-3-babage-002 as predictive models of M&A targets. Moreover, this study explores the terminology of U.S. energy target companies' annual reports. We try to ascertain whether target companies use renewable energy-related terminology more frequently than non-target ones. For this purpose, we use exploratory data analysis (EDA) and data visualisation techniques like word-clouds and frequency bars to study term usage patterns.

1.1. Contributions

Our paper has important implications for both academics and practitioners (firm managers, individual or institutional investors, policymakers, and other stakeholders). This study's academic contributions are threefold: a) It contributes to the underdeveloped literature on the use of information extracted from company-related documents in the realm of M&A target prediction as we add to this emerging literature of only six previous studies but with promising results. Our research novelty entails, to the best of our knowledge, the first-ever application of transformer-LLM models for direct classification of M&A target identification. Our predictive model for takeover targets is directly the transformer-LLM model, effectively showcasing their capability for this task of direct classification as predictive models. b) This work opens the door to many new research lines for improved LLMs training and application to directly predict with enhanced accuracy in M&A scenarios, not only targets but also successful M&As, for example, and many other strands of research in the area of M&As such as detecting illegal corporate insider trading, uncovering the dynamics of market convergence through M&A, or measuring M&A performance, amongst others. c) Additionally, as our paper increases the knowledge about the prediction of M&A targets in the energy sector, it also contributes to the growing literature on Green M&A (GM&A). GM&A refers to merger and acquisition enterprises engaging in acquiring green resources and improving environmental image (Lu, 2021; Sun et al., 2023; Yang & Chi, 2023), which incorporates green concepts into traditional acquisition. Moreover, our paper's contributions to practitioners are also threefold: by predicting M&A targets on time, d) we help investors leverage the opportunity of considerable gains that occur by the dates of the announcement deal; e) policy-makers and regulators, can better assess antitrust compliance and investigate suspected insider trading; f) we help energy companies (not only the possible acquirers, but also the possible targets such as startup companies to assess their possibility of being acquired or merged) and heavy-polluting companies to better identify M&A targets, expediting a quicker transition to cleaner energies.

1.2. Organization of the paper

The rest of the paper is organised as follows. Section 2 reviews the literature on M&A target prediction with ML and NLP techniques, showing no prior studies have applied transformer models to directly predict M&As. Section 3 details the empirical study, both our sample collection and our methodology. We first delineate the process of data acquisition, cleaning, and preprocessing. Then, we describe the approach for fine-tuning the transformers and their application as predictive models for direct classification. Section 4 reports and discusses our empirical results. Finally, we close our paper with the conclusions and further work in section 5.

2. Literature review

The study of M&A target prediction has been an area of active research for many years with voluminous literature. Tunyi (2021) reviews fifty years (1968–2018) of research on takeover target prediction from a historical perspective, partitioning prior studies in the area into four distinct eras: initial studies on takeover target prediction focused on financial ratios. However, it was later concluded (1986–2002) that it was impossible to build a successful investment strategy based on such a target prediction, and only the subsequent use of alternative modelling techniques, such as ML methods, achieved a significant improvement in prediction performance, although still with limited accuracy (2003–2009). In the fourth and last era (2010–2018), he found a decline in the literature on M&A target prediction but these studies provided some evidence that takeover target prediction can lead to abnormal returns when combined with appropriate screening strategies. Although we add some other prior studies not included in Tunyis's 2021 review and carefully extend it for studies over the last five years (2019–2024), we refer to it as a very interesting and complete review (Tunyi, 2021). Our review of prior literature on M&A target prediction until the last five years is summarized in Table 1 and shows that these studies have been mainly based on a) quantitative financial variables and b) 'traditional' statistical and ML techniques such as logistic regression (Anagnastopoulos & Rizeq, 2019, 2021; Aramyan, 2021; Pasiouras & Tanna, 2010; Rodrigues & Stevenson, 2013); support vector machine-SVM (Ouzounis et al., 2009); decision trees-DT, random forest-RF and ensemble models (Bourne et al., 2019; Espahbodi & Espahbodi, 2003; Wei et al., 2008), neural networks-NN (Anagnastopoulos & Rizeq, 2019, 2021; Bourne et al., 2019; Liu et al., 2007; Ouzounis et al., 2009; Rodrigues & Stevenson, 2013; Sen & Gibbs, 1994), etc. Regarding the quantitative financial variables

Table 1. Summary of traditional statistical models and ML techniques, and financial and managerial variables used in some prior studies on M&A target prediction.

Authors	Methods	Financial variables/measures	Major limitations
Sen and Gibbs (1994)	NN	Inefficient management (ROE, sales turnover). Undervaluation (market/book value, P/E, dividend payout). Demographic factors (firm size).	This study did not achieve additional performance over LR models due to irrelevant input data.
Zanakis and Zopounidis (1997)	LR	Profitability: (EBIT/TA, CF/TA, NI/NW, gross profit/TA). Managerial performance: (fixed assets/TA, NW/TA, (LTD+CL)/WC, inventory/WC, inventory/TA, cash/TA). Solvency: ((LTD+CL)/TA, (LTD+CL)/CF, NW/(NW+LTD), quick assets/CL, CA/CL, WC/TA).	The data's heterogeneity, due to the inclusion of different sectors, and the limited historical occurrence of takeovers in Greece vs larger M&A markets such as the U.S., reduces predictive power.
Espahbodi and Espahbodi (2003)	DT	Firm size, growth-resource mismatch, leverage, investment opportunity, intangible assets, free cash flow, dividend policy.	Significant drop in accuracy in the validation set, raising robustness concerns.
Liu et al. (2007)	NN	Profitability (profit margin, ROA). Growth (increase ROA, increase ROE, increase ratio of sales). Asset management (asset turnover; inventory turnover; account receivables turnover). Solvency (Quick ratio, Current ratio, asset/liability ratio).	This study only includes data from the Chinese market between 2004 and 2006, which will probably limit generalizability to other periods of time. Also, the Hopfield network achieves moderate accuracy (80.69% for targets and 61.33% for non-targets), which is not ideal.
Brar et al. (2009)	LR	Firm size; inefficient management; growth-resource mismatch; undervaluation; financial structure; sales; firm age; # of employees; market barrier.	Correct classification rate is moderate, indicating challenges in accurately identifying takeover targets (only 42 out of 99 targets were identified).
Wei et al. (2008)	Ensemble models	Technological variables derived from patent analysis: technological quantity (number of patents, number of recent patents); technological quality (impact of patent, technology strength); technological innovation (link to science, Technology Cycle Time); technological diversity (concentration rate); technological compatibility (Compatibility of Technological Fields, Relative Strength of Technological Quantity, Relative Strength of Technological Quality, Relative Strength of Technological Innovation, Relative Strength of Technological Diversity).	This study focuses adds value for technology intensive industries but is not directly applicable to the energy sector.
Ouzounis et al. (2009)	SVM, NN (MLP), UTADIS	Growth-resource mismatch, inefficient management, firm size, undervaluation, free cash flow, price-earnings, dividend policy.	This study excluded financial companies from the dataset due to distinct accounting practices, reducing dataset diversity. Also, the matching of target firms with similar non-target firms may have introduced limitations due to reliance on predefined criteria.
Pasiouras and Tanna (2010)	LR	Profitability (NIM, NIM/TA, ROA, ROE); firm size (TA); capital strength (equity/assets, equity/loans, equity/liabilities, equity/short-term funding); expenses management; liquidity; market share	There is a complete reliance on financial data, potentially missing relevant information, as well as considerable levels of misclassification.
Rodrigues and Stevenson (2013)	LR, NN and combinations	Inefficient management. Company undervaluation. P/E. Growth-resource mismatch. Dividend payout. Inefficient financial structure. Firm size.	Some limitations include its inability to individually assess the significance of each input, the need for a large dataset, and that results might be influenced by the number of non-targets within predictions.
Anagnastopoulos and Rizeq (2019)	NN (MLP), LR	Inefficient management: (ROE). Company undervaluation: (P/E). Firm size: (Market capitalization). Leverage: (Debt/equity). Liquidity: (Current ratio). Growth: (Rate of change of annual revenues).	This study focuses exclusively on the tech sector in the US, limiting its generalizability to other industries or region.
Bourne et al. (2019)	RF, NN, and ensemble models	Profitability, growth & returns. Capital structure/Liquidity. Trading multiples. Executive compensation & demographics. Board ownership. Patent profile	The model's effectiveness is hindered by outdated predictions and limited access to alternative or proprietary data sources to, for example, include data for private companies. The study is also limited by the lack of industry-specific metrics, challenges in variable selection, and the long prediction durations.

(*Continued*)

Table 1. Continued.

Authors	Methods	Financial variables/measures	Major limitations
Meghouar and Ibrahimi (2021)	LR	Growth-resource mismatch, undervaluation, firm size, firm performance, dividend policy, free cash flow, growth, ownership structure	A key limitation of this study lies in the restricted sample size, as it focuses solely on large firms with deals exceeding $100 million.
Aramyan (2021)	LR and clustering	Inefficient Management: (ROE, ROC, Return on Sales, Profit Margin, Gross Profit Margin, Profit to Capital, Sales Growth(3y)). Company undervaluation: (P/E, EV to Sales, EV to EBITDA, Market to Book). Growth-resource Imbalance: (Growth-Resource Mismatch, Free Cash Flow to Sales, Operating Cash Flow to Total Assets). Leverage: (Total Debt to Equity, Debt to EV, Net Debt per Share, Net Debt to Total Capital). Liquidity: (Current Ratio, Cash to Capital, Working Capital to Total Assets).	This study is limited by the sole use of logistic regression, which may not capture all complexities of M&A dynamics. Additionally, evolving market conditions and incomplete variable selection highlight the need for further data sources and advanced modeling techniques to enhance reliability and relevance.
Anagnastopoulos and Rizeq (2021)	NN, LR	Inefficient management: (ROE, ROA, EBITDA, Asset turnover). Company undervaluation: (P/E, Market/book value). Firm size: (Market capitalization, TA). Leverage: (Debt/equity, equity multiplier). Liquidity: (Current ratio, net WC). Growth: (Rate of change of annual revenues).	Limitations include unclear cause-and-effect relationships between variables, the use of a single industrial setting (IT market), and the absence of external validation with independent data.

LR: logistic regression; DT: decision tree; SVM: support vector machine; NN: neural network; MLP: multilayer perceptron; UTADIS: utilités additives discriminantes; RF: random forest; TA: total assets; CF: cash-flow; NI: net income; NW: net worth; WC: working capital; LTD: long-term debt; CL: current liabilities. NIM: net interest margin; P/E: price to earnings.
Source: Own elaboration.

(Table 1), researchers typically have based their ML predictive models on factors that increase the likelihood of a company becoming an M&A target, many of them grounded on the popular main hypotheses introduced by the financial literature to identify a takeover target such as: • the inefficient management hypothesis (Palepu, 1986; Healy et al., 1992)) typically characterized by profitability ratios (EBITDA margin ROE, ROCE, ROA and/or asset turnover); • the growth-resource mismatch hypothesis (Palepu, 1986, Aramyan, 2021) measured by Free Cash Flow to Sales, Operating Cash Flow to Total Assets, Dividend Payout; • the undervaluation hypothesis (Bradley et al., 1988; Dietrich & Sorensen 1984, Pound, 1988; Dong et al., 2006; Healy et al., 1992; Powell, 1997) characterized by market valuation ratios (market-to-book and price-earnings ratios); • the firm-size hypothesis (Dietrich & Sorensen 1984; Levine & Aaronovitch, 1981; Palepu, 1986; Powell, 1997), measured by market capitalization and total assets; • the leverage, liquidity and growth variables, are also takeover determinants (Aramyan, 2021; Palepu, 1986; Powell & Yawson, 2007; Rodrigues & Stevenson, 2013), measured by debt to equity, current ratio and growth in revenues.

Our review of prior literature on M&A target prediction until the last five years also highlights the claims some of these authors have made over time about the need to incorporate nonfinancial variables for improving the accuracy of these financial-based M&A target prediction models, even the ML ones (Zanakis & Zopounidis, 1997; Pasiouras & Tanna, 2010). Thus, taking a step further in the literature review, it is precise since the fourth and last era identified by Tunyi (2021), since 2019, over the last five years, that the latest AI advancements for M&A target prediction increasingly involve text mining approaches and Natural Language Processing (NLP) techniques, which can analyze unstructured text to uncover insights that are not evident from quantitative data alone. However, this M&A target prediction literature based on textual analysis, although growing, is still scarce and underdeveloped, with only six recent studies (Aramyan, 2022; Hajek & Henriques, 2024; Katsadados et al., 2024; Katsafados et al., 2021; Moriarty et al., 2019; Routledge et al., 2017) to the best of our knowledge (Table 2). None of those six papers focus specifically on the U.S. energy sector, and all of them except Moriarty et al. (2019) combine textual data with financial variables. Moreover, what is more relevant, only two of these six papers (Aramyan, 2022; Hajek & Henriques, 2024) use LLM-transformer models but they are not used as direct predictive models of M&A targets, but for previous sentiment analysis to construct sentiment-based variables that the authors include in their prior 'traditional' ML predictive models of M&A targets (which baseline form used only quantitative financial variables) and analyze the contribution of these sentiment-based variables obtained through LLM-transformers to the performance of their prior 'traditional' ML predictive models. Thus, no prior study exists in the literature applying LLM-transformers for direct classification of M&A target identification. Additionally, the evolution of the different NLP techniques used can be noticed through the review of these six only existing papers (Table 2) that apply a range of these NLP techniques, from the most traditional ones such as the 'Bag-of-Words' (BoW) approaches (Katsafados et al., 2024), to the most advanced ones such as LLMs (Aramyan, 2022; Hajek & Henriques, 2024). BoW has the significant drawback of assuming that the words in the text are independent of one another, which is evidently false. Other studies have relied on word embedding technology (Katsafados et al., 2024), and, even though they are an improvement on the use of BoW, this NLP technique does not offer a completely satisfactory answer either, because the vector representation does not take into account the context of the sentence and therefore negations are not considered by this approach. As a result, both methods are very far from understanding sentences in their full extent. However, the field is experiencing an enormous revolution since the implementation of transformer models has been possible due to the rise of computing. Transformers are LLMs that capture the dependencies between words, that are encoded in word embeddings whose space represents the meaning of the words. Specifically, transformers' empirical results dramatically outperform the classical pipeline of machine learning models with a bag of-words representation of the most common and relevant words of the texts according to algorithms such as TF-IDF (Term Frequency-Inverse Document Frequency) (Garrido-Merchán et al., 2023). TF-IDF is used by three of these six studies (Katsafados et al., 2021, 2024; Moriarty et al., 2019). But transformers can only be accurately estimated by supercomputing centers or organizations that have lots of computing power, being impossible to be trained from zero by small organizations or businesses for small supervised learning tasks. To avoid this limitation in its application, transfer learning can be used to fine-tune a transformer trained in a similar task to the one that needs to be solved (Yang et al., 2020). The fine-tuning process adapts the behavior of the transformer to the particular task to be solved and it is cheap in computational terms, thus allowing any organization to perform it. Moreover, the results that a fine-tuned transformer can deliver outperform classical methodologies, or small models trained from zero (Garrido-Merchán et al., 2023). For this reason, in our paper, we propose fine-tuning three pre-trained transformers (FinBERT, ALBERT and GPT-3 babbage-002).

Routledge et al. (2017) examined the impact of incorporating textual variables on their M&A target predictive model (logistic regression) with only financial variables and found that combining the two kinds of information, financial variables, and text, resulted in higher performance. They focused on U.S. companies and included text from the 'Management's Discussion and Analysis' (MD&A) section of the SEC's 10-K form, a common practice in similar papers. Additionally, Moriarty et al. (2019) have also included the 'Business Description' section in their study, applying different NLP techniques (TF-IDF and latent Dirichlet allocation (LDA)) and then using logistic regressions and clustering to predict both targets and bidders M&A across all industries in the USA. Katsafados et al. (2021) explored M&A target and acquirer prediction in the US banking sector by combining logistic regression and sentiment analysis. They specifically utilized Loughran and McDonald's list of positive and negative words specifically tailored for texts in a corporate or financial context (Loughran & McDonald, 2011), to classify annual report's text sentiment as positive or negative, integrating this sentiment classification into their logistic regression model as another variable. The paper concluded that banks with a more negative tone in their annual reports were more likely to become M&A targets (Katsafados et al., 2021). In subsequent research for US bank M&A target prediction these same researchers (Katsafados et al., 2024) use textual information along with financial variables as inputs to several ML models (LR, SVM, RF and NN) and found that: (1) when textual information is used as a single type of input, the predictive accuracy of their ML predictive models is similar, or even better, compared to their ML predictive models using only financial variables as inputs, and (2) when they jointly use textual information and financial variables as inputs, the predictive accuracy of their ML predictive models is substantially improved compared to the same models using a single type of input. Finally, the only two papers that use transformer-LLMs in the context of M&A target predictions (Aramyan, 2022; Hajek, & Henriques, 2024), concretely BertRNA, BERTopic, and Finbert models, do not use them as direct predictive models of M&A targets, but for previous sentiment analysis to construct sentiment-based variables that the authors include in their prior 'traditional' ML

Table 2. Summary of 6 only prior papers on M&A target prediction that use textual variables and NLP techniques (other than LLM-transformers).

Authors	Use LLM-transformer?	Use LLM-transformer for directly predicting M&A targets?	Text mining or the NLP technique used	Type of variables used	'Traditional' ML predictive models used
Routledge et al. (2017)	No	–	Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network	• Financial variables (market/book value, PPE (the book value of property plant, and equipment), log of cash balance, the size of leverage (book value of debt over book value of assets), size (market value of equity), and ROA). • Textual Data from annual 10-k fillings (MD&A section)	LR
Moriarty et al. (2019)	No	–	TF-IDF LDA (latent Dirichlet allocation)	• Financial variables • Textual Data from annual 10-k fillings (MD&A and Business Description sections)	LR clustering
Katsafados et al. (2021)	No	–	TF-IDF (term frequency-inverse document frequency)	• Bank-specific financial variables (cost efficiency, ROA, firm size, capital strength, loans/TA, loan loss provisions, non-interest income, • Textual Data from annual reports	LR
Katsadados et al. (2024)	No	–	Bag of Words (BoW) TF-IDF Word embeddings	• Bank-specific financial variables (cost efficiency, ROA, firm size, capital strength, loans/TA, market power, asset quality, non-interest income, deposits) • Textual Data from annual reports	LR, SVM, RF and NN
Aramyan (2022)	Yes	No	Use FinBERT and BERT-RNA but only for prior sentiment extraction on the news preceding M&A announcements. These sentiment variables are then included in the ML models together with the financial variables	• Financial Indicators: the same as in Aramyan (2021) (see Table 1) • Text-based: (news sentiment)	LR RF XGBoost
Hajek and Henriques (2024)	Yes	No	Use FinBERT but only for prior sentiment analysis and topic detection (BERTopic) on company-specific news articles. These sentiment variables are then included in the ML models together with the financial variables.	• Financial Indicators: Inefficient management; Growth-resource mismatch; undervaluation; profitability; leverage; firm size and age; liquidity; dividend policy; profitability; P/E; ownership structure. • Text-based: (news sentiment and topic detection)	LR LDA DT SVM Ensemble methods: *Bagging-based: -Bagging -RF -ExtraTrees *Boosting-based: -AdaBoost -GBoost -XGBoost

Source: Own elaboration.

predictive models of M&A targets (which baseline form used only quantitative financial variables) and analyze the contribution of these sentiment-based variables obtained through LLM-transformers to the performance of their prior 'traditional' ML predictive models. Thus, they do not analyze the capacity of these transformer-LLMs to directly predict M&A targets, but they 'only' analyze the capacity to leverage the text-based variables (obtained through transformer-LLMs) to improve their prior ML predictive models (LR, RF and XGBoost in Aramyan, 2022; and LR, LDA, DT, SVM, Bagging, RF, ExtraTrees, AdaBoost, GBoost, and XGBoost in the case of Hajek, & Henriques, 2024). Our approach applies fine-tuned transformer-LLMs for direct classification of M&A target identification. Our predictive models for M& target prediction are directly the fine-tuned transformer models (concretely, in our case, we fine-tune FinBERT, ALBERT, and GPT-3-babage-002). For instance, while FinBERT is trained originally (and in the works of Aramyan, 2022; Hajek, & Henriques, 2024) to effectively identify sentiment in a given text as 'positive', 'negative' or 'neutral', thereby predicting three classes, in our study, the model has been fine-tuned for a binary classification task, labelling texts as '1' (target) or '0' (non-target). This adaptation allows the model to focus on distinguishing between the two specific classes relevant to our research.

To complete this review, two more papers must be mentioned. Ma et al. (2017) present a case study for China's computer numerical control machine tools industry in which they propose a detailed stepwise strategy-based methodology based on patent textual analysis to identify, filter, select, and evaluate target firms for a specific acquirer in this sector, additionally using clustering techniques and a qualitative assessment of technical experts at this industry-related companies and Government Commissions. Thus, although they apply advanced text mining techniques (not transformers) in an M&A target context, they do not develop an M&A target predictive model. They rather identify and select the companies that are more likely to be successfully completed in a case study. Parungao et al. (2022) use qualitative textual information from target firms' letters to shareholders to describe the attractiveness of firms as M&A targets, but they do not predict M&A targets; they predict the completion of the said deal, the success of this M&A. That is, they develop an M&A completion prediction model (via decision tree), not an M&A target prediction model. Moreover, they do not use transformers in their textual analysis, nor advanced NLP techniques.

Thus, to the best of our knowledge, there are no prior studies in the literature applying LLM-transformers for direct classification of M&A target identification and our study is the first to propose it. Hence, based on the context of the research outlined in the introduction section and on the literature review discussed so far, we address two research questions in this paper:

RQ1: *How do transformer-LLM models perform when applied for direct classification of M&A target identification in the U.S. energy sector?*

RQ2: *Do U.S. energy target companies use renewable energy-related terminology more frequently than non-target ones?*

There exists also growing literature on applying ML techniques to different areas of M&As, other than predicting merger participants, which are beyond the scope of this paper, but we just want to point out them as strands of M&A literature that could leverage the application of our paper's proposed methodology (LLM-transformer) for further research (as stated in the suggested further research section). A non-exhaustive review of prior studies on these other M&A literature streams includes examples such as predicting M&A failures (Lee et al., 2020) or success (Branch et al., 2008; Ma et al., 2017; Morgan, 2018; Parungao et al., 2022; Zhang et al., 2012); detecting illegal corporate insider trading (Esen et al., 2019); uncovering the dynamics of market convergence through M&A (Aaldering et al., 2019); or measuring M&A performance (An et al., 2006); amongst others.

Finally, as our paper increases the knowledge about the prediction of M&A targets in the energy sector, it also contributes to the growing literature on Green M&A (GM&A). GM&A refers to merger and acquisition enterprises engaging in acquiring green resources and improving environmental image (Lu, 2021; Sun et al., 2023; Yang & Chi, 2023), which incorporates green concepts into traditional acquisition. Prior literature (mostly in Chinese firms) has documented that GM&As significantly impact on a wide spectrum of variables: reducing illegal pollution discharge (Lu et al., 2023); promoting corporate green transition, i.e. corporate environmental investment (Lu, 2021; Sun et al., 2023); enhancing green innovation (Liang et al., 2022; Zhang et al., 2023; Zhu et al., 2024) and ESG performance and sustainable development (Li & Lu, 2023; Zhang et al., 2024); promoting energy efficiency in Chinese listed high energy-consuming companies (Lu et al., 2024); enhancing firms' environmental responsibility (Shi & Huang, 2024). Although the GM&A literature is mostly focused on China, Qiao et al. (2023) show the M&A behavior of US renewable energy firms has a significantly higher effect on technological innovation than that of Chinese and European renewable energy firms; and second, the effect of cross-border M&A of Chinese and European renewable energy firms on innovation is significantly higher than the effect of domestic M&A on innovation, while the effect of cross-border M&A of US renewable energy firms on innovation is not significantly different from the effect of domestic M&A on innovation.

3. Methodology

This section deals with the materials and methods that have been used to provide an answer to our research questions. We start the section with the data that has been used as we believe that it is the most critical information for practitioners, delineating the process of data acquisition, cleaning, and pre-processing. Then, we describe the approach for fine-tuning the transformers and their application as predictive models for direct classification.

3.1. Dataset

The Global Industry Classification Standard (GICS) defines the energy sector as encompassing oil, gas, and coal companies, while companies operating in electricity generation (including renewable and nuclear sources) are classified as utilities (MSCI, 2023). However, for the purposes of this paper, the term 'energy sector' is used to refer to both GICS sectors.

3.1.1. Data sourcing

The final dataset is a compilation of multiple data sources. First, a Bloomberg dataset for energy M&A transactions in the United States since 2015 was downloaded directly from the Bloomberg terminal. For this purpose, the Bloomberg search was narrowed down by applying the following filters:

- Country: United States (both for target and acquirer)
- Announcement Date: 01/01/2015 - 31/04/2024
- Deal Type: M&A
- Deal Status: Completed
- Sector: Energy (target)
- Target: Public
- Deal Value: Over $500 million

The initial search resulted in an Excel file containing over 700 transactions. Basic cleaning was performed by removing companies with missing information, those whose tickers did not end in 'US' and targets incorrectly classified as energy companies by Bloomberg with SIC (Standard Industrial Classification) codes distinctly different from energy, such as 'advertising', 'malt beverages', or 'book publishing'. From this clean dataset, 20 transactions were randomly selected to form the 20 targets for the final data. Additionally, 180 random non-targets for different years were chosen from a list of 51 non-target companies, which included 22 companies from the S&P500 Energy index and 29 from the S&P500 Utilities index. This process resulted in an Excel file listing 200 companies from 2015 to 2023, with approximately 22 companies per year. Therefore, 10% of the final sample are target companies.

This percentage means there is an overrepresentation of non-target companies, however, this has been done on purpose to obtain a dataset that closely approximates a stratified sample. This means that the sample maintains the same proportion of positives in the entire population. In this case, 10% represents a realistic percentage of companies subject to M&As in the market, as it aligns with the rounded average (9.5%) of datasets used in five previous academic papers (Aramyan, 2021; Brar et al., 2009: Routledge et al., 2017; Hajek & Henriques, 2024; Katsafados et al., 2021).

The imbalanced nature of the dataset (a typical characteristic of whichever M&A target prediction task and thus, it is present in all such prior studies in the literature), results in underrepresentation bias, meaning there is an unequal distribution of classes in the training data and the model might struggle to predict certain minority classes, hindering its performance (Cherepanova et al., 2023). To address this issue, the model's accuracy must surpass the majority rule, which in this case means it should be greater than 90%. This is because if the model were to always predict the majority class, the resulting accuracy would be the same as the percentage of that class in the sample. Therefore, the model must exceed this percentage to provide meaningful predictive value.

The analysis has been focused on the past 10 years, a period long enough to encompass diverse economic scenarios, including growth, recession, and volatility. This allows us to incorporate the influence of various contexts on M&A activities into the model. Moreover, the energy sector has experienced significant developments during this time, including technological advancements, regulatory changes, and energy price fluctuations. Starting the analysis in 2015 is particularly relevant due to the signing of the Paris Agreement, which has spurred investment in renewable energy and sustainable practices, influencing M&A in the sector.

The final step in constructing the dataset involved extracting the respective MD&A sections from 10-k filings. These filings are mandatory annual reports that American companies must submit to the United States Security and Exchange Commission (SEC), the authority overseeing the country's securities market. These reports are composed of five sections: (i) Business Description; (ii) Risk factors; (iii) Selected financial data; (iv) MD&A; and (v) Financial Statements and Supplementary Data. For this work, only the text from the fourth section has been extracted, as it provides detailed insight into management's view of the company's performance.

The advantage of standardized sections in filings is that the extraction process can be automated. A script in Python programming language is then developed for an automatic process and executed using Google Colab and Jupyter Notebook. Additionally, the SEC provides an API for the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system, which houses all public SEC documentation since 1994, facilitating easy access to these datasets. Using this API, the sec-api.io API we extract textual data from annual reports for these 200 target companies and non-target peers in the S&P500 Energy and Utilities sectors across their corresponding years and exported into an Excel file using Python.

Finally, the resulting dataset consists of 5 columns (Year, Ticker, Company, MD&A Text, and Target) and 200 observations. It is important to mention that the 'Year' variable refers to the year prior to the Announcement Date obtained from Bloomberg, as this represents the data that acquirers considered when making their transaction decisions. The 'Target' variable is a dummy variable where a '1' indicates the company was a target of M&A and a '0' indicates the opposite. This will be the dependent variable the models aim to predict. Table 3 illustrates the first five rows of the dataset at this stage.

3.1.2. Text preprocessing

In this second phase, the MD&A text undergoes preprocessing or cleaning to eliminate any potential noise that could interfere with the algorithm's performance. This task is critical because the quality of the corpus, serving as the model's sole input, directly impacts its effectiveness. The Python library 'nltk' is utilized for this purpose, removing extra whitespaces, numbers, and non-alphanumeric characters such as punctuation marks, special characters, or line breaks. Additionally, all text is converted to lowercase, considering evidence that algorithms may yield different predictions for the same word based on its case. Lastly, stop words (articles, conjunctions, prepositions, pronouns, and common verbs) are eliminated as they contribute minimal value to the text.

Another typical step in preprocessing textual data is lemmatization, which involves reducing words to their root form or lemma. For instance, 'ate' would be transformed into 'eat'. However, lemmatization may sometimes restrict the depth of the text and could potentially interfere with capturing specific nuances, particularly in neural-based algorithms (Camacho-Collados & Pilehvar, 2018). Hence, lemmatization was omitted in this study. Also, while some studies opt to eliminate words or tokens with low frequency (Katsafados et al., 2024), this study focuses on predicting a rare event, so removing infrequent words might result in the loss of valuable information.

Observations containing less than 50 words were excluded. This step was deemed necessary due to the existence of observations where texts comprised brief references, such as 'The information required by Item 7 is set forth under Management's Discussion and Analysis of Financial Condition and Results of Operations; in the Annual Report, on pages 2 through 81'. Consequently, some observations were removed resulting in a dataset comprising 183 companies, of which 20 were identified as targets. Despite the reduction, the dataset still maintains a realistic representation with 11% of companies classified as targets.

3.1.3. Exploratory data analysis

Once all text had been processed, a text length analysis was conducted to analyze any possible differences between targets and non-targets, as depicted in Table 4. It shows that texts associated with target companies (labelled as '1') have a mean length of 2366 words, while non-targets (labelled as '0') have a slightly lower average word count of 2094. Additionally, the standard deviation of non-target texts (931.7) is significantly above that of targets (316.4), suggesting there is greater variability in length in the non-target group. This could imply that reports for targets are more detailed and consistent in length.

A preliminary sentiment analysis was conducted using the TextBlob library on Python. To interpret the output, it should be noted that '0' implies a neutral sentiment polarity, while scores above this indicate a positive sentiment and lower scores imply a negative sentiment. Also, the further from '0' the score is the stronger the sentiment. The results, as shown in Table 5, reveal that both groups have very close to neutral sentiment scores. This could be a result of management avoiding negative implications in reports, a practice known as impression management (Caserio et al., 2019). Nonetheless, target average and median sentiment scores were almost half of non-target scores, implying a lower sentiment. Also, target sentiment presented lower standard deviations compared to non-targets. This could indicate that targets have a lower sentiment than non-targets and this sentiment is more consistent.

Finally, differences in content were analyzed on a high level. The word frequency bar charts (Figure 1) illustrate the occurrence of specific words in each group per report, as resulting word counts have been divided by the total number of observations for each group. In the sustainability category, terms like 'renewable', 'clean', 'green', 'sustainable', and 'neutral' are examined. Interestingly, non-target companies use all terms, except for 'neutral', more frequently than targets. For cleaner energy sources the same happens, as the words included are more common in non-target company reports. In the fossil fuels category, target companies mention 'oil' and 'gas' significantly more often than non-target companies, which is surprising. The dominance of these terms in target companies' reports is contrary to the initial idea that target companies could be related to renewable energies.

Examining terms related to sentiment, target companies use 'growth' and 'opportunity' more frequently. This could reflect an optimistic outlook and potential growth opportunities in the future that might attract M&A. The term 'risk' is also more prevalent in target companies, which is consistent with the idea that target 10-Ks would have negative sentiment. In the M&A category, target companies use 'merger', 'acquisition' and 'partnership' more frequently. This might be because companies that have engaged in previous M&A activities tend to have a better understanding of the process and are seen as more capable of achieving synergies and cost savings, making them attractive targets. For general business terms,

Table 3. Dataset structure.

	Year	Ticker	Company	MD&A text	Target
0	2015	APA	APA Corp	Item 7. Management & #146; s Discussion and Anal…	0
1	2015	CVX	Chevron Corp	Item 7. Management & #8217; s Discussion and Ana…	0
2	2015	EQT	EQT Corp	Item 7. & #160; Mangement & #8217; s Discu…	0
3	2015	HES	Hess Corp	Item 7. Management & #8217; s Discussion and Ana…	0
4	2015	MRO	Marathon Oil Corp	Item 7. Mangement & #8217; s Discussion and Ana…	0

Source: Own elaboration.

Table 4. Text length analysis.

mean	std	min	25%	50%	75%	max
2094.255556	931.702179	17.0	2034.25	2503.5	2682.25	2916.0
2365.550000	316.403386	1692.0	2235.75	2472.0	2571.00	2819.0

Source: Own elaboration.

target companies mention debt much more frequently, which aligns with the idea that target companies exhibit higher leverage (Aramyan, 2021; Palepu, 1986; Powell & Yawson, 2007; Rodrigues & Stevenson, 2013). Also, 'liquidity', 'profit', 'loss' and 'management' appear more frequently in target's reports, while 'investment' is more common in non-targets.

The word frequency analysis reveals distinct differences in the language used by target and non-target companies. Non-target companies emphasize sustainability and cleaner energy terms, while target companies focus on traditional energy sources, growth opportunities, and M&A activities. These insights provide valuable context for understanding the different communication styles of each group and these patterns will form the basis for the algorithm's learning process.

Also, two word clouds displaying the 40 most frequent words for target and non-target companies have been generated, as depicted in Figures 2 and 3, respectively. Several terms such as 'crude oil', 'natural gas', 'cash flow', 'cost', 'financial statement', 'revenue', or 'commodity price' are prevalent in both groups, reflecting their relevance within the industry context. However, notable distinctions emerge, with certain words appearing more frequently in one group compared to the other. For instance, 'adjusted EBITDA', 'production', 'project', 'capital expenditure', and 'acquisition' feature prominently in the target word cloud but are absent in the non-targets' depiction. Conversely, terms like 'income tax', 'net income', 'per share', and 'impact' are more prevalent among non-targets. The anticipated outcome is for the model to discern these variations and identify their significance in classifying a company as a target or non-target.

This EDA, using data visualisation techniques such as word clouds and frequency bars to study term usage patterns, will also let us explore our second research question.

3.2. Transformer models

This section details the organization of the dataset and the training of transformer models. The process starts with dividing it into training and test sets, and preparing the dataset by tokenizing the text, splitting it into manageable chunks, and choosing hyperparameters. The discussion then moves to the fine-tuned transformer models trained (FinBERT, ALBERT and GPT-3 babbage-002), highlighting their configurations and training processes. Finally, key aspects of the models' architecture and optimization techniques are examined to provide a clear understanding of their setup.

Vaswani et al. (2017) introduced the Transformer, a novel neural network architecture, that processes the whole text at once, rather than in order like Recurrent Neural Networks (RNNs) do, allows for parallelization, making its training much more efficient in terms of memory, as well as faster than previous methods. Moreover, it can be more easily scaled than previous algorithms to analyze large datasets. These improvements have resulted in the ideation of the Transformer architecture to be regarded as a milestone in the progress of both NLP and AI. Transformer Architecture forms the pillars of some of the innovative and popular models. Such is the case of the notorious GPT model (Generative Pre-trained Transformers), as well as BERT (Bidirectional Encoder Representations from Transformers), Copilot, and BART, among others. In this context, we can finetune large language models, like transformers such as BERT or the GPT family, to efficiently solve a task that is similar to the one that they were trained on, which is precisely the methodology that we follow in our paper. In this paper, we use three transformer models, two of which are based on the BERT model (FinBERT and ALBERT) and the third one belongs to the GPT family (GPT-3 babbage-002).

GPT models have been developed by the company OpenAI, launching their first version, GPT-1, in 2018 but truly becoming a benchmark for NLP models due to the high quality of the text produced with their second version, GPT-2, in 2019 (Wu et al., 2023). However, it did not get worldwide recognition until the 2020 inauguration of GPT-3, the most extensive language model trained at the time. This behemoth

Table 5. TextBlob sentiment analysis.

Target	mean	std	min	25%	50%	75%	max
0	0.023064	0.044531	−0.108370	0.000000	0.026360	0.052879	0.116764
1	0.012721	0.036496	−0.072259	−0.003454	0.016456	0.039497	0.075831

Source: Own elaboration.

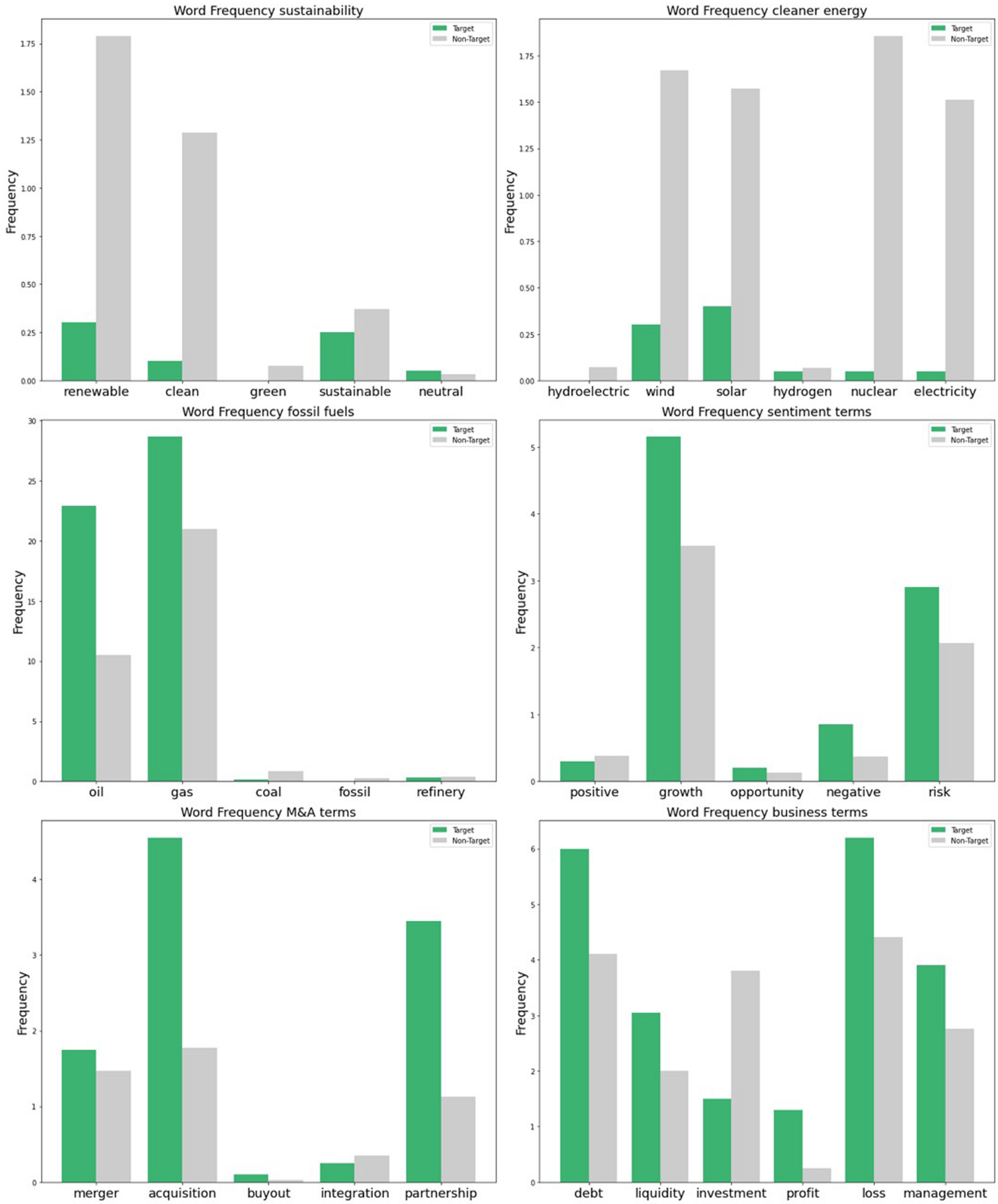

Figure 1. Word frequency analysis bar charts.
Source: Own elaboration.

NLP model was trained with 175 billion parameters, making it ten times bigger in size than the closest competing models created by Google and Microsoft according to Dale (2021).

Another leading model in the field is BERT, developed by Google Research in 2018. This model was trained with a BookCorpus formed by 800 million words (the same one as GPT-1) plus a Wikipedia dataset containing 2500 million words (Devlin et al., 2018). This model is the foundation of newer models such as ALBERT (Lan et al., 2019) and FinBERT among others, the latter a state-of-the-art LLM that adapts to the finance domain, developed by Huang et al. (2023).

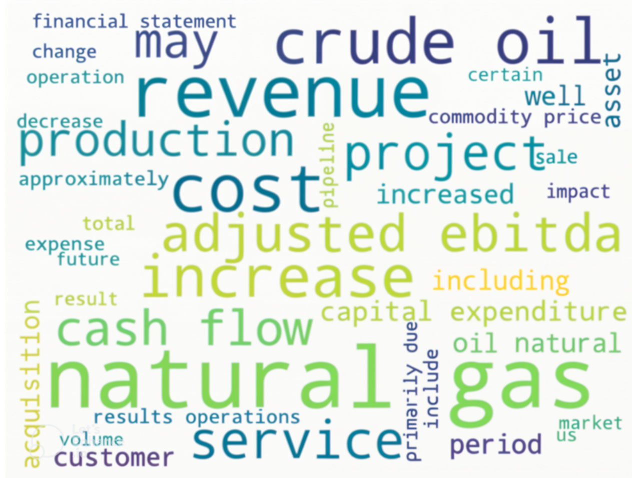

Figure 2. Word-cloud for target companies.
Source: Own elaboration.

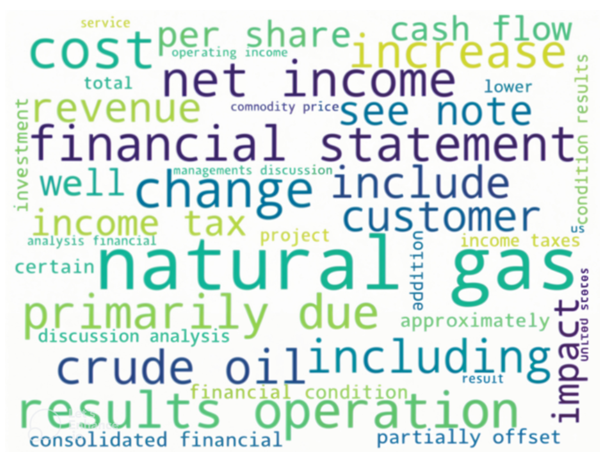

Figure 3. Word-cloud for non-target companies.
Source: Own elaboration.

3.2.1. Model input construction

The first step was to segment the dataset into training and test sets. Following the common practice of previous studies (Moriarty et al., 2019, Katsafados et al., 2024; Routledge et al, 2017), an 80/20 split was used. The split is based on date values rather than being done randomly, so the 80% corresponding to the training set is composed of the first 145 observations, while the rest of the sample containing the 36 most recent observations is used for testing. As the model is expected to predict future events, it is better to test it in a future period (Espahbodi & Espahbodi, 2003; Katsafados et al., 2024).

Transformer models have architectural constraints with sequence length limits due to the computational limitations of self-attention mechanisms. For example, BERT has a maximum sequence length of 512 tokens. To accommodate this limit, the tokenizer incorporates truncation, which stops observations at the 512-token limit. However, this derives in a loss of context, as any text beyond the 512th word is disregarded. Various techniques have been proposed to handle longer texts, but the 512-token limit remains a significant constraint in practical applications.

To address this issue, a chunking approach has been used by splitting texts into chunks of a maximum of 512 tokens. An overlap between one chunk and the next has been included to maintain context and ensure smooth transitions, which is especially relevant in sequence-to-sequence models such as transformers (Jaiswal & Milios, 2023). Additionally, the tokenizer includes a padding option, where texts that are less than 512 tokens long are filled with zeros ('0's) to occupy the full space of 512 tokens. This is done to standardize sequence sizes (Varis & Bojar, 2021).

Next, these chunks are tokenized. Tokenisation is a critical step in NLP, where text is converted into a numeric format that models can understand. The string of text is converted into individual tokens, which will take different formats depending on the method used. Different NLP models require different tokenizers that align with the model's architecture. For example, the Bag of Words approach is a basic tokenization method where each token is an individual word. However, this method is quite limited as it does not consider word order and does not capture contextual information.

In the case of BERT, it uses a WordPiece tokenizer, which is a subword tokenizer. This splits certain words into subword units to make training more efficient and handle uncommon and compound words (Minixhofer et al., 2023). For instance, the word 'deleverage' would be transformed into ['del', '##ever', '##age'], where hashtags indicate that the token is part of a larger word. The tokenizer returns a dictionary that represents the encoded input for the model. This dictionary contains three key-value pairs (see Figure 4 for an illustration of encodings created on the dataset):

- Input_ids: each token is mapped to a unique numerical ID based on a pre-built vocabulary that the model was trained on.
- Token_type_ids: used in tasks like question answering, which is not our case.
- Attention_mask: differentiates between actual text (labelled as 1) and padding tokens (labelled as 0).

To enhance the efficiency of the training process and optimize the dataset for accelerated computing via a Graphics Processing Unit (GPU), a data loader is employed. The data loader facilitates the batching and shuffling of input data for the transformer model. Batching enables the processing of smaller groups of data before updating parameters, leading to faster computation and more stable training (Ioushua et al., 2023). Shuffling randomizes data at the start of each epoch, preventing the model from learning order patterns and reducing the risk of overfitting. However, in this study, shuffling has not been used because the MD&A texts have been split in order to meet the token limit and the chunks corresponding to one original text should be analyzed sequentially. The data loader also supports parallelization and the use of pinned memory, which speeds up data transfer to the GPU.

After the data is preprocessed and efficiently loaded, hyperparameters that control the learning process are chosen. Key hyperparameters include:

- **Batch Size**: A batch size of 8 is used to balance memory usage and efficiency.
- **Number of Workers**: 4 workers have been used to leverage parallel data loading.
- **Learning Rate**: it is the step size in error adjustment at each iteration looking for a minimum of the loss function, which controls how much the model's weights are adjusted. If it is too big, it might skip the minimum error, and if it is too small, it might incur overfitting. A value of 0.00002 or 2e-5 is commonly used for fine-tuning transformers.
- **Number of Epochs**: this is the number of times that the algorithm goes through the entire training set. Training for 3 epochs aims to ensure sufficient learning while preventing overfitting.

3.2.2. FinBERT

FinBERT is a domain-specific version of the BERT model, pre-trained on financial data. Like BERT, FinBERT uses a transformer architecture that employs self-attention mechanisms to understand contextual relationships between words in a sentence. The model processes text by converting words into embeddings,

```
input_ids: [2, 1097, 18, 5460, 2495, 1553, 2874, 1736, 1311, 17140, 2103, 8338,
token_type_ids: [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
attention_mask: [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
```

Figure 4. Train encodings resulting from AlbertTokenizer.
Source: Own elaboration.

which are then analyzed bidirectionally to capture their meaning in the given context. By leveraging this bidirectional understanding, FinBERT can interpret complex financial language and extract insights such as sentiment or classifications. When fine-tuned, the model can be updated to recognize patterns and nuances specific to the task at hand, such as M&A target classification in this study.

We first fine-tune the FinBERT model (https://huggingface.co/ProsusAI/finbert). Originally, FinBERT has been trained on a diverse set of financial data, including Reuters financial news and a sentiment-labeled financial phrase bank, making it domain-specific to finance (Malo et al., 2014). This specialisation makes this model especially interesting for this study. Although it was originally designed for sentiment analysis tasks, this model can still be used for various other applications through fine-tuning.

Fine-tuning involves adjusting the internal parameters in a pre-trained model for a specific task. This process leverages FinBERT's pre-existing knowledge of word meanings and relationships in financial contexts while tailoring it to the requirements of the new task. In this study, the fine-tuning process involves using a dataset specific to our target task, ensuring the model adapts well to the nuances of the new application.

For instance, FinBERT was originally trained to effectively identify sentiment in a given text as 'positive', 'negative' or 'neutral', thereby predicting three classes. However, in this study, the model has been fine-tuned for a binary classification task, labelling texts as '1' (target) or '0' (non-target). This adaptation allows the model to focus on distinguishing between the two specific classes relevant to our research. We have trained this model on Google Colab which allows users to connect CUDA (NVIDIA's parallel computing architecture), reducing training time significantly compared to fine-tuning on a personal computer. The model was fine-tuned once on a personal computer and took around 5 hours to finish the task, while using Google Colab models took between 5 and 10 minutes to fine-tune.

As mentioned earlier, the dataset suffers from data imbalance, where only 11% of the dataset consists of target companies. The models have been trained on this imbalanced data, however, the model has also been fine-tuned on balanced data. This has only been done for BERT models (FinBERT and ALBERT) because training is free and, hence, more convenient to test different percentages of the minority class. Oversampling and undersampling are two common techniques used to address class imbalance in datasets. Oversampling involves increasing the number of instances in the minority class (target companies) by replicating them, which helps the model learn more from these examples. In contrast, undersampling reduces the number of instances in the majority class (non-targets) to match the minority class, leading to a smaller dataset.

In this case, undersampling was not used because it would result in a smaller sample from an already small dataset, potentially causing the model to miss important patterns in the data. With only 20 of the observations being target companies, reducing the non-target class would have left too few samples to train a model. Therefore, oversampling was deemed more suitable, as it allowed for maintaining a larger dataset while addressing the class imbalance. For example, creating 5 duplicates of each target observation (6 total copies) results in the distribution shown in Figure 5, where the sample has 44% of target companies and is much more balanced than the original. With 7 copies, there are 48% of targets, and 5 copies result in 40% of targets.

3.2.3. Albert

Additionally, we fine-tune ALBERT (https://github.com/google-research/ALBERT), another version of the original BERT model. ALBERT, which stands for 'A Lite BERT', is a variant of BERT optimized for efficiency and performance. It achieves this by employing techniques that significantly reduce memory usage and computational cost without sacrificing performance (Gao et al., 2023). This is done by implementing two innovations to the BERT model:

- Factorized Embedding Parameterization: This technique addresses one of the key inefficiencies in BERT, where word embeddings take up a significant portion of model size. It is done by separating the size of hidden layers (dimension of final word representation in the model) from the vocabulary embedding size (number of words the model can recognize), significantly reducing parameter redundancy.

- Cross-layer Parameter Sharing: in BERT, each transformer layer has its own set of parameters. However, ALBERT reuses the same parameters across all transformer layers, reducing the number of parameters while maintaining model performance.

These optimizations allow ALBERT to process large amounts of text faster and with lower computational costs, while still analyzing text bidirectionally. Like all BERT models, it uses self-attention to identify relationships between words and can be fine-tuned to adapt to specific tasks, such as binary classification for target prediction.

This makes ALBERT particularly interesting for this task, as it offers the robust language understanding capabilities of BERT while being more resource-efficient. This efficiency is especially beneficial for tasks that require processing large amounts of data or need to be deployed in environments with limited computational resources. Considering that our 200 observations average around 2000 words each, the model will need to process over 400,000 words, which translates to an even larger number of tokens.

The model, like FinBERT, has been fine-tuned on Google Colab to speed up the process and benefit from the parallel processing capabilities of available GPU and it has also been fine-tuned with both an imbalanced and an oversampled dataset. ALBERT's efficiency ensures that it can handle this substantial processing load faster than FinBERT. However, leveraging the power of GPU the difference was not substantial, with ALBERT taking around 30 seconds less to fine-tune than FinBERT.

3.2.4. GPT-3 babbage-002

GPT-3 is a powerful LLM developed by OpenAI. It's known for its exceptional capabilities in generating human-quality text, translating languages, writing creative content, and answering questions. It utilizes a classical transformer encoder-decoder architecture and has been pre-trained on massive amounts of text. However, unlike BERT, the exact training data size and source remain undisclosed.

A GPT-3 model has been fine-tuned to complement this study with a different type of transformer model. OpenAI offers a user-friendly API that has been used to access and leverage the capabilities of GPT-3. A significant advantage of the OpenAI API is its pay-per-use model. Users only pay for the resources consumed, making it a cost-effective option for exploring GPT's capabilities. Also, it offers a selection of models with varying capabilities and costs. More specifically, the 'babbage-002' version of GPT-3 has been fine-tuned for the task. This is because it is the most affordable option available, costing $0.4/1 M tokens, making it more attractive for an experimental stage. However, no testing for oversampling has been carried out with the GPT model as each try would increase the cost of the study.

Babbage-002 is a variant of the GPT-3 model, optimized for cost-efficiency and performance balance. It uses the transformer architecture, processing text in an autoregressive manner—predicting the next word based on prior context, utilizing a decoder-only design that focuses on generating text rather than understanding bidirectional relationships. Compared to larger GPT-3 models like Davinci, Babbage-002 is lighter, faster, and more affordable, making it ideal for moderate-complexity NLP tasks. While it lacks the depth of more advanced GPT models, it offers a practical trade-off between computational efficiency and output quality, particularly in budget-conscious applications such as this study.

GPT-3 has a maximum token limit of 4096 tokens, which comfortably exceeds the token count of our data's longest single observation. Consequently, there was no need to implement text chunking for this dataset, as each observation falls well within the model's token capacity. This allowed the training process to proceed smoothly without the need for additional preprocessing steps to divide the text into smaller segments. By staying within GPT-3's token limit, the model can fully utilize each complete observation, ensuring that no critical information is lost due to truncation. The OpenAI interface shows that 1,150,026 tokens were trained.

```
Class distribution after oversampling: Counter({0: 128, 1: 102})
```

Figure 5. Resulting class distribution for oversampling with 6 total copies.
Source: Own elaboration.

While using a more updated version of the GPT model like GPT-4 or using ChatGPT is tempting, fine-tuning the model will tailor the model to the specific task at hand. Prompting through ChatGPT's free interface is easier. However, it relies only on the model's pre-trained knowledge, which might not capture the domain-specific details required for M&A target prediction in this case.

4. Results and discussion

When evaluating the performance of a classification model, it is crucial to use a variety of metrics to gain a comprehensive understanding of how well the model is performing. Each metric offers insights into different aspects of the model's effectiveness and helps in making informed decisions about improvements. The metrics that have been evaluated are accuracy, precision, recall, the confusion matrix, and types of errors (see Table 6). By considering all these metrics, the strengths and weaknesses of the different models can be examined.

The confusion matrix shows the number of true positive, true negative, false positive, and false negative predictions made by the model and provides a detailed breakdown of the model's performance, allowing us to see not just the overall accuracy but also the types of errors the model is making. In the confusion matrix in Table 6, columns represent predictions and rows represent actual values, with non-targets on the left column and top row and targets on the right column and bottom row. For instance, FinBERT (imbalanced)'s confusion matrix shows that 262 observations have been predicted as '0' correctly, and 14 predicted '0's were actually '1's. Also, this model predicted 11 '1's correctly, but 10 observations predicted as targets were actually non-targets.

In the context of predicting M&A targets, understanding Type I and Type II errors is crucial. A Type I error, or false positive, occurs when the model incorrectly identifies a company as an M&A target when it is not. This can lead to unnecessary investing in a company based on the erroneous belief that it is a likely target, potentially resulting in wasted resources and financial losses. Conversely, a Type II error, or false negative, happens when the model fails to identify a true M&A target, thus missing out on valuable opportunities. This can result in significant opportunity costs, as investors or companies might overlook profitable investments. Balancing these errors is essential, as the goal is to minimize the costs associated with false positives while also capturing as many true positives as possible to maximize strategic advantages.

Accuracy measures the proportion of correctly classified instances out of the total instances. It is particularly useful when the classes in the dataset are balanced, as it gives a clear picture of how often the model makes correct predictions. However, when the dataset is imbalanced, accuracy can be misleading as a model that always predicts the majority class will achieve high accuracy but fails to capture the minority class effectively. This is the case of the ALBERT (with an imbalanced dataset) and GPT models, where accuracy is 92% and 84%, respectively, but the models do not classify any test observations as targets. The highest accuracy was achieved by the FinBERT model with an oversampling of 7 copies, reaching 95.29%, followed by the ALBERT model with an oversampling of 6 copies (93.62%). The worst accuracy by far was that of the ALBERT with an oversampling of 7 copies (83.90%), which is due to the model predicting all observations as targets. This model, when fed 7 or more copies fails to capture the differences between classes and predicts all as '1'.

Precision is the proportion of true positive predictions out of all positive predictions, which is particularly important in situations where the cost of false positives is high. For instance, in M&A target prediction, a false positive means investing in a company that is predicted to become a target but does not. The highest precision is achieved by the FinBERT model with 7 oversampling copies, which is correct 87% of the times it predicts a positive instance, which is more than 20 basis points higher than the next best model. The worst-performing models are ALBERT and GPT on the imbalanced dataset, which are incapable of predicting target companies and, therefore, have a score of 0%.

Finally, recall or sensitivity, measures the proportion of true positive predictions out of all actual positive instances. High recall indicates that the model is effective at identifying positive instances, even if it means producing more false positives. Recall ensures that the model captures as many relevant instances as possible, maximising the targets correctly identified. In this sense, the best-performing model is ALBERT with 6 oversampling copies, reaching 64% recall, followed by FinBERT with a 52% score.

Table 6. Performance metrics for all models.

Model	Accuracy	Precision	Recall	Conf. Matrix
FinBERT (imbalanced)	91.92%	52.38%	44.00%	262 10 14 11
FinBERT oversampling (7 copies)	95.29%	86.67%	52.00%	270 2 12 13
FinBERT oversampling (6 copies)	91.92%	53.33%	32.00%	265 7 17 8
FinBERT oversampling (5 copies)	92.93%	62.50%	40.00%	266 6 15 10
ALBERT (imbalanced)	91.61%	0.00%	0.00%	273 0 25 0
ALBERT oversampling (7 copies)	83.90%	100.00%	100.00%	0 273 0 25
ALBERT oversampling (6 copies)	93.62%	61.54%	64.00%	263 10 9 16
ALBERT oversampling (5 copies)	84.23%	17.65%	24.00%	245 28 19 6
GPT-3 babbage-002 (imbalanced)	83.59%	0.00%	0.00%	32 0 3 0

Source: Own elaboration.

Again, ALBERT and GPT fine-tuned on imbalanced datasets achieved the worst scores with 0% sensitivity.

Regarding the fine-tuned models, the FinBERT model with oversampling of 7 copies emerged as the top performer, correctly identifying M&A target companies with high confidence and minimal false positives, making it suitable for applications where precision is critical, such as financial investments. False positives can lead to misguided investment decisions, potentially resulting in financial losses when a company is incorrectly identified as an acquisition target. Therefore, the superior precision of FinBERT underscores its robustness in accurately identifying true M&A targets while mitigating the risk of false positives. However, one might argue that this model misses significant investment opportunities and could prefer the ALBERT model with an oversampling of 6 copies, which achieves a higher recall and is able to capture more opportunities. This represents a trade-off and the choice of model would depend on the investor's risk aversion, being an investor with a high-risk aversion more likely to choose FinBERT and vice versa.

Conversely, the ALBERT and GPT models fine-tuned with imbalanced data performed quite poorly in predicting M&A targets, particularly due to their inability to classify any test observations as targets despite achieving high accuracy scores of 92% and 84%, respectively. This highlights a significant limitation when using these models in scenarios involving imbalanced datasets and specific domain contexts such as finance. With only 200 observations and an 11% target class, the data available in this study was insufficient to provide a robust training ground for these models. This limited data likely prevented the model from learning meaningful patterns and accurately distinguishing between target and non-target companies. The imbalance between target and non-target classes further exacerbates this issue, making it difficult for the model to generalize beyond the training data. Also, ALBERT's lack of pre-training on financial data and GPT's primary suitability for generative tasks rather than classification are probable reasons behind their underperformance in this study. Therefore, it can be concluded that it is better to fine-tune models on balanced datasets, even if this is not representative of actual M&A target companies as a percentage of all companies in the market.

Since no prior study exists in the literature applying LLM-transformers for direct classification of M&A target identification, we cannot directly compare our research results to previous comparable studies. However, we will compare them in the context of the only six prior papers on M&A target prediction that use textual variables and NLP techniques (other than LLM-transformers), shown in Table 2, that provide some kind of comparable performance metric. Hajek and Henriques (2024) also found the best-performing model is the random forest with an accuracy score of 93.9%, below our FinBERT model with an oversampling of 7 copies (95.29%). Aramyan (2022), found a precision of 26% for both Random Forest and XGBoost models with sentiment-based variables obtained by applying FinBERT, far below the precision reached by all our FinBERT models (52%, 44%, 40% and 32%) and even further below the precision of our ALBERT model with an oversampling of 7 copies (100%) and with an oversampling of 6 copies (61.54%). He also found a recall of 51% for the RF (far below the recall achieved by our ALBERT models with an oversampling of 7 copies (100%) and with an oversampling of 6 copies (64%), and slightly below the recall of our FinBERT with oversampling of 7 copies (52%)), and a recall of 53% for the XGBoost (far below the recall achieved by our ALBERT models with an oversampling of 7 copies (100%) and with an oversampling of 6 copies (64%), and slightly higher than the recall of our FinBERT with oversampling of 7 copies (52%)). Finally, Moriarty et al. (2019) and Katsafados et al. (2024) applied less advanced NLP techniques than LLM-transformers (TF-IDF and LDA in Moriarty et al., and TF-IDF, BoW, and word embeddings in Katsafados et al. (2024), in their target prediction task, found the best-performing model is the random forest (using both textual features based on the bag of word approach and financial variables), with an accuracy score of 89.7%, below the accuracy achieved by all our FinBERT models (95.29%, 92.93% and 91.92%), and also below our ALBERT model with an oversampling of 6 copies (93.62%). Moriarty et al. (2019) found that it is easier to predict acquirers with a high degree of accuracy than it is to predict targets since their best result for target precision was only 7.6%.

Related to the second research question of this study, although the EDA revealed key differences between target and non-target companies, no significant evidence was found to indicate that target companies use more renewable-related words, contrary to what could have been expected from previous findings by Hawkes et al. (2023), Niemczyk et al. (2022), or Andriuškevičius and Štreimikienė (2021). However, it was found that target company reports tend to be longer and more consistent. Sentiment analysis showed that while both groups had neutral scores, target companies exhibited slightly lower and more consistent sentiment. The content analysis highlighted that non-target companies frequently used sustainability-related terms, whereas target companies focused more on traditional energy sources, growth opportunities, risks, and M&A activities. These patterns suggest distinct characteristics for both groups that could be learned by a model.

One possible reason for the lack of significant evidence supporting the hypothesis that target companies use more renewable-related terminology could be the lower level of pressure in the clean energy transition in the United States. As it has been less aggressive and slower in promoting sustainability and clean energy adoption than Europe, this difference in regulatory and market pressures could explain why the study did not find significant renewable energy-related terminology in the MD&A sections of US target companies (Lantiainen & Song, 2014).

The classification task has been moderately successful, with several models achieving high performance suggesting transformer models can effectively learn to identify target companies. However, the study did not uncover enough evidence to affirm that target companies use renewable energy-related terminology more frequently than non-target companies.

This study faces several limitations and barriers, making the task non-trivial and challenging. One significant limitation is the difficulty in accessing data for free. Resources like Open Corporate require payment for comprehensive data access, and while the SEC API was used, it only allows for 100 free requests. This constraint limited the amount of data available for this study, adding complexity to the task by resulting in a relatively small dataset. Transformer models require substantial amounts of data to train effectively and achieve high performance. This study's dataset consisted of only 200 observations, with just 20 being target companies. This small sample size is significantly below the typical requirements for training transformer models, which often involve thousands or even millions of examples to accurately capture the complexity and nuances of language patterns. Even for fine-tuning tasks, extensive datasets are usually necessary to adapt the model effectively to specific contexts.

Due to the easier accessibility of data, this study focuses exclusively on large, publicly traded companies, meaning that the unique dynamics and patterns of M&A activity involving smaller startups are not captured. Andriuškevičius and Štreimikienė (2021) highlighted that companies often prefer growth through the acquisition of startup companies. Publicly traded companies included in the dataset are much bigger and their acquisition probably results from different strategies. This limitation could impact the findings and their applicability to a broader range of M&A scenarios in the energy sector. Difficulty in data sourcing is further exacerbated in the European market, where the study was initially aimed to focus on, as it is hard to find easily extractable data and where filings differ not only in format but also in language.

5. Conclusions and suggested future lines of research

This paper assesses two research questions. First, it analyses the predictive utility and performance of three transformer-based LLMs in directly predicting M&A target companies, while the secondary research question investigates the relationship between target companies and renewable energy terminology in their annual reports.

In summary, the study effectively showcased transformer models' capability in identifying M&A target companies, with FinBERT utilizing oversampling, being the top-performing model due to its high precision and minimized false positives, critical for precise financial decision-making. While the study uncovered notable differences in target and non-target report characteristics, it found no significant evidence that M&A target companies use more renewable energy-related terminology, suggesting other factors may be more influential in M&A activity within the energy sector in the USA.

This research provides explicit findings that showcase the effectiveness of transformer models, particularly FinBERT, in predicting M&A targets in the U.S. energy sector. Among the models tested, FinBERT emerged as the top performer, achieving high precision (86.67% with oversampling), which is critical for minimizing false positives in financial decision-making. ALBERT, while slightly less precise, demonstrated higher recall in certain configurations, capturing more opportunities to identify targets. However, GPT-3 babbage-002 underperformed in this domain, especially with imbalanced datasets, highlighting its limitations in classification tasks tailored for finance. This suggests that encoder-decoder architectures may be more suitable than decoder-only for such task The analysis also revealed distinct characteristics of target companies, such as longer and more consistent MD&A reports, a slightly lower sentiment score, and a focus on growth opportunities and M&A-related terms. Despite these insights, the study found no significant evidence that target companies use renewable energy-related terminology more frequently, suggesting other factors may play a larger role in M&A activity within this sector.

This study also addresses several gaps in the existing literature. It is the first to apply fine-tuned transformer models, such as FinBERT and ALBERT, for direct classification of M&A targets, offering a novel methodological contribution to the field. Unlike prior studies that relied on traditional machine learning models or sentiment-based variables derived from textual data, this study integrates unstructured text from SEC filings to uncover new insights. Additionally, it focuses on the U.S. energy sector, a domain largely unexplored in M&A prediction research, providing a sector-specific lens on M&A dynamics. By tackling class imbalance using oversampling techniques and examining Green M&A considerations, the study paves the way for future research into how environmental and non-financial factors shape M&A activity, while also demonstrating the potential of transformer models in predictive financial applications.

Moreover, our paper's contributions to practitioners are also threefold: by predicting M&A targets on time, we help investors leverage the opportunity of considerable gains that occur by the dates of the announcement deal; policy-makers and regulators, can better assess antitrust compliance and investigate suspected insider trading; and, finally, we help energy companies (not only the possible acquirers, but also the possible targets such as startup companies to assess their possibility of being acquired or merged) and heavy-polluting companies to better identify M&A targets, expediting a quicker transition to cleaner energies.

These findings emphasize the ongoing need for improved LLMs model training and application to enhance predictive accuracy in M&A scenarios.

Future research should focus on expanding the dataset significantly. Collecting a larger number of observations with a higher proportion of target companies will provide a more comprehensive foundation for training and evaluating transformer models. Additionally, exploring other techniques to handle class imbalance, such as advanced oversampling methods or generating synthetic data, could improve model performance. Furthermore, gaining access to more advanced LLM models (due to their recent rapid development) and computational resources will likely enhance the effectiveness and efficiency of the models used.

Another promising direction for future research is to extend the analysis to European markets and perform a comparative study. The regulatory environment, market dynamics, and cultural factors differ significantly between the US and Europe, which could lead to different patterns and predictors of M&A activity.

Also, experimenting with different chunk sizes when processing textual data might improve model performance. For instance, using ChunkBERT, a model designed to handle longer texts by splitting them into more manageable chunks, could be a valuable approach (Jaiswal & Milios, 2023). This technique could allow for the capture of more contextual information within lengthy documents, thereby potentially improving the accuracy and robustness of M&A target predictions.

Finally, it would also be interesting to explore using transformer-LLMs on alternative prediction tasks in other strands of M&A prior research, such as predicting M&A failures (Lee et al., 2020) or success (Branch et al., 2008; Ma et al., 2017; Morgan, 2018; Parungao et al., 2022; Zhang et al., 2012); detecting illegal corporate insider trading (Esen et al., 2019); uncovering the dynamics of market convergence through M&A (Aaldering et al., 2019); or measuring M&A performance (An et al., 2006); amongst others.

Acknowledgments

Inés Rodríguez-Muñoz-de-Baena: conceptualization, design, data curation, and analysis, writing of the first draft. María Coronado-Vaca: conceptualization, design, methodology and analysis, writing and revision of the first draft challenging it obtaining a critical revision to improve the intellectual content and academic attractiveness. Esther Vaquero-Lafuente: conceptualization, revising the drafts, including a critical contribution to improve content and academic attractiveness. All authors have approved the final manuscript.

Authors' contributions

CRediT: **Inés Rodríguez-Muñoz-de-Baena**: Conceptualization, Data curation, Formal analysis, Software, Writing – original draft, Writing – review & editing; **María Coronado-Vaca**: Conceptualization, Formal analysis, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing; **Esther Vaquero-Lafuente**: Conceptualization, Writing – review & editing.

Disclosure statement

The authors have no relevant financial or non-financial interests to disclose.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

About the authors

Inés Rodríguez-Muñoz-de-Baena holds a double degree in Business Administration and Business Analytics from ICADE (Universidad Pontificia Comillas, Madrid) and works as an Associate at Boston Consulting Group (BCG). Her research interests include Mergers and Acquisitions, Private Equity and Natural Language Processing.

María Coronado-Vaca holds a PhD in Business Administration and Economics (Universidad Pontificia Comillas-ICADE), a Quantitative Finance MSc. and is a Certified Financial Analyst by EFFAS (European Federation of Financial Analyst Associations). She is a professor in the Faculty of Business and Economics (ICADE) at Universidad Pontificia Comillas in Madrid, where she has also been Director of the Department of Financial Management, Director of the Master's Degree in Supervision of Credit Institutions for the Bank of Spain and Coordinator of the Master's Degree in Finance. She has focused her research on green finance, banking, climate-induced financial innovation, financial risk management and measurement, IA in finance, and published several articles, book chapters, and books.

Esther Vaquero-Lafuente is a Full Professor at Universidad Pontificia Comillas, ICADE, Madrid, where she teaches Corporate Finance and Portfolio Management. She holds a PhD in Business and Economics. She is the Director of the Sustainable Finance Observatory at Universidad Pontificia Comillas in collaboration with CaixaBank AM. Her research interests include Sustainable Finance and Behavioral Finance.

ORCID

María Coronado-Vaca http://orcid.org/0000-0002-0315-3753
Esther Vaquero-Lafuente http://orcid.org/0000-0001-5196-3462

Data availability statement

The data that support the findings of this study (corporates' reports) are available from Bloomberg and the US SEC (Securities and Exchange Commission). Restrictions apply to the availability of Bloomberg data, which were used under license for the current study, and so are not publicly available. Data are, however, available on reasonable request from the corresponding author.

References

Aaldering, L. J., Leker, J., & Song, C. H. (2019). Uncovering the dynamics of market convergence through M&A. *Technological Forecasting and Social Change*, *138*, 95–114. https://doi.org/10.1016/j.techfore.2018

Adnan, A., & Hossain, A. (2016). Impact of M&A announcement on acquiring and target firm's stock price: An event analysis approach. *International Journal of Finance and Accounting*, *5*(5), 228–232. https://doi.org/10.5923/j.ijfa.20160505.02

Ahmad, T., & Zhang, D. (2020). A critical review of comparative global historical energy consumption and future demand: The story told so far. *Energy Reports*, *6*, 1973–1991. https://doi.org/10.1016/j.egyr.2020.07.020

An, S., He, Y., Zhao, Z., & Sun, J. (2006). Measurement of merger and acquisition performance based on artificial neural network. *Proceedings of the 5th IEEE International Conference on Cognitive Informatics*, 17-19 July 2006, Beijing, China (pp. 502–- 506). Institute of Electrical and Electronics Engineers (IEEE); Curran Associates, Inc. https://doi.org/10.1109/COGINF.2006.365537

Anagnostopoulos, I., & Rizeq, A. (2019). Confining value from neural networks: A sectoral study prediction of takeover targets in the US technology sector. *Managerial Finance*, *45*(10/11), 1433–1457. https://doi.org/10.1108/MF-12-2017-0523

Anagnostopoulos, I., & Rizeq, A. (2021). Conventional and neural network target-matching methods dynamics: The information technology mergers and acquisitions market in the USA. *Intelligent Systems in Accounting, Finance and Management*, *28*(2), 97–118. https://doi.org/10.1002/isaf.1492

Andriuškevičius, K., & Štreimikienė, D. (2021). Developments and trends of mergers and acquisitions in the energy industry. *Energies*, *14*(8), 2158. https://doi.org/10.3390/en14082158

Aramyan, H. (2021). *Prediction of M&A targets to generate portfolio returns*. 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

Aramyan, H. (2022). *Predicting M&A targets using ML: Unlocking the potential of NLP based variables*. LSEG Refinitiv. https://developers.lseg.com/en/article-catalog/article/predicting-MnA-targets-using-ML-Unlocking-the-potential-of-NLP-variables

Barros, V., Verga Matos, P., Miranda Sarmento, J., & Rino Vieira, P. (2022). M&A activity as a driver for better ESG performance. *Technological Forecasting and Social Change*, *175*, 121338. https://doi.org/10.1016/j.techfore.2021.121338

Bhagat, S., Dong, M., Hirshleifer, D., & Noah, R. (2005). Do tender offers create value? New methods and evidence. *Journal of Financial Economics*, *76*(1), 3–60. https://doi.org/10.1016/j.jfineco.2004.05.002

Bhutada, G. (2022). *The 200-year history of mankind's energy transitions*. World Economic Forum. https://www.weforum.org/agenda/2022/04/visualizing-the-history-of-energy-transitions/

Billett, M. T., King, T.-H. D., & Mauer, D. C. (2004). Bondholder wealth effects in mergers and acquisitions: New evidence from the 1980s and 1990s. *The Journal of Finance*, *59*(1), 107–135. https://doi.org/10.1111/j.1540-6261.2004.00628.x

Bouchard, P., Voß, S., Heilig, L., & Shi, X. (2023). A case study on smart grid technologies with renewable energy for central parts of Hamburg. *Sustainability*, *15*(22), 15834. https://doi.org/10.3390/su152215834

Bourne, M., Dircks, T., Kardatzke, J., Miller, D., Roberts, M., & White, C. (2019). *From ML to M&A Ten M&A Target Predictions through a Machine Learning Model*. Research Report, Nicholas Center for Corporate Finance & Investment Banking, University of Wisconsin. https://business.wisc.edu/centers/nicholas/blog/nicholas-centers-first-ever-machine-learning-consulting-project-research-report-on-predicting-ma-targets/

Bradley, M., Desai, A., & Kim, E. H. (1988). Synergistic gains from corporate acquisitions and their division between the stockholders of target and acquiring firms. *Journal of Financial Economics*, *21*(1), 3–40. https://doi.org/10.1016/0304-405X(88)90030-X

Branch, B., Wang, J., & Yang, T. (2008). A note on takeover success prediction. *International Review of Financial Analysis*, *17*(5), 1186–1193. https://doi.org/10.1016/j.irfa.2007.07.003

Brar, G., Giamouridis, D., & Liodakis, M. (2009). Predicting European takeover targets. *European Financial Management*, *15*(2), 430–450. https://doi.org/10.1111/j.1468-036X.2007.00423.x

Camacho-Collados, J., & Pilehvar, M. T. (2018). On the role of text preprocessing in neural network architectures: An evaluation study on text categorization and sentiment analysis. *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP, pp*. 40–-46. https://doi.org/10.18653/v1/w18-5406

Campa, J. M., & Hernando, I. (2004). Shareholder value creation in European M&As. *European Financial Management*, *10*(1), 47–81. https://doi.org/10.1111/j.1468-036X.2004.00240

Caserio, C., Panaro, D., & Trucco, S. (2019). Management discussion and analysis: A tone analysis on US financial listed companies. *Management Decision*, *58*(3), 510–525. https://doi.org/10.1108/MD-10-2018-1155

Chen, F., Ramaya, K., & Wu, W. (2020). The wealth effects of merger and acquisition announcements on bondholders: New evidence from the over-the-counter market. *Journal of Economics and Business*, *107*(, 105862. https://doi.org/10.1016/j.jeconbus.2019.105862

Cherepanova, V., Reich, S., Dooley, S., Souri, H., Dickerson, J., Goldblum, M., & Goldstein, T. (2023). A deep dive into dataset imbalance and bias in face identification. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, August 08–10, 2023, Montréal, QC, Canada, pp. 229—247. Association for Computing Machinery (ACM). https://doi.org/10.1145/3600211.3604691

Dale, R. (2021). GPT-3: What's it good for? *Natural Language Engineering*, *27*(1), 113–118. https://doi.org/10.1017/S1351324920000601

Deloitte. (2023). Energy Quarterly Report Q1 2023. https://www2.deloitte.com/content/dam/Deloitte/xa/Documents/corporate-finance/us-dcf-2023-q1-energy-report-final.pdf

DeLong, G. L. (2001). Stockholder gains from focusing versus diversifying bank mergers. *Journal of Financial Economics*, *59*(2), 221–252. https://doi.org/10.1016/S0304-405X(00)00086-6

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. https://doi.org/10.48550/arXiv.1810.04805

Dietrich, J. K., & Sorensen, E. (1984). An application of logit analysis to prediction of merger targets. *Journal of Business Research*, *12*(3), 393–402. https://doi.org/10.1016/0148-2963(84)90020-1

Ding, Y., Li, J., & Tian, Y. (2024). The short and long-term effects of cross-border M&A network on Chinese enterprises' green innovation. *Economic Modelling*, *134*, 106704. https://doi.org/10.1016/j.econmod.2024.106704

Dodd, P., & Ruback, R. (1977). Tender offers and stock holder returns. An empirical analysis. *Journal of Financial Economics*, *5*(3), 351–373. https://doi.org/10.1016/0304-405X(77)90043-5

Dong, M., Hirshleifer, D., Richardson, S., & Hong Teoh, S. (2006). Does investor misvaluation drive the takeover market? *The Journal of Finance*, *61*(2), 725–762. https://doi.org/10.1111/j.1540-6261.2006.00853.x

Easterwood, S., Netter, J., Paye, B., & Stegemoller, M. (2024). Taking over the size effect: Asset pricing implications of merger activity. *Journal of Financial and Quantitative Analysis*, *59*(2), 690–726. https://doi.org/10.1017/S0022109023000030

EIA (U.S. Energy Information Administration). (2021). *EIA projects a nearly 50% increase in world energy use by 2050, led by growth in renewables*. https://www.eia.gov/international/data/world

Energy Institute (EI). (2024). *73rd Statistical Review of World Energy*.

Esen, M. F., Bilgic, E., & Basdas, U. (2019). How to detect illegal corporate insider trading? A data mining approach for detecting suspicious insider transactions. *Intelligent Systems in Accounting, Finance and Management*, *26*(2), 60–70. https://doi.org/10.1002/isaf.1446

Espahbodi, H., & Espahbodi, P. (2003). Binary choice models and corporate takeover. *Journal of Banking & Finance*, *27*(4), 549–574. https://doi.org/10.1016/S0378-4266(01)00258-8

Galperina, L., & Klen, Y. (2017). Global trends of international mergers and acquisitions in the energy sector. *International Economic Policy*, *14*(8), 21–58. https://doi.org/10.3390/en14082158

Gao, X., Ding, G., Liu, C., & Xu, J. (2023). Research on high precision Chinese text sentiment classification based on Albert Optimization. *2023 15th International Conference on Advanced Computational Intelligence (ICACI)*, pp. 1–-6. https://doi.org/10.1109/icaci58115.2023.10146191

Garrido-Merchan, E. C., Gozalo-Brizuela, R., & Gonzalez-Carvajal, S. (2023). Comparing BERT against traditional machine learning text classification. *Journal of Computational and Cognitive Engineering*, *2*(4), 352–356. https://doi.org/10.47852/bonviewJCCE3202838

Global Carbon Project. (2023). Global carbon budget 2023 [Data set]. https://www.globalcarbonproject.org/

Goergen, M., & Renneboog, L. (2004). Shareholder wealth effects of European domestic and cross-border takeover bids. *European Financial Management*, *10*(1), 9–45. https://doi.org/10.1111/j.1468-036X.2004.00239.x

Hajek, P., & 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

Hawkes, A., Muûls, M., & Hamilton, J. W. (2023). Big oil and the energy transition: Evidence from M&A. *Energy Policy*, *183*, 113762. https://doi.org/10.1016/j.enpol.2023.113762

Healy, P. M., Palepu, K. G., & Ruback, R. S. (1992). Does corporate performance improve after mergers? *Journal of Financial Economics*, *31*(2), 135–175. https://doi.org/10.1016/0304-405X(92)90002-F

Huang, A. H., Wang, H., & Yang, Y. (2023). FinBERT: A large language model for extracting information from financial text. *contemporary Accounting Research*, *40*(2), 806–841. https://doi.org/10.1111/1911-3846.12832

Institute of Mergers, Acquisitions & Alliances (IMAA). (2023). *M&A statistics: Transactions and activity by year. M&A trends*. https://imaa-institute.org/mergers-and-acquisitions-statistics/

International Energy Agency (IEA). (2020). Global primary energy demand by fuel, 1925-2019. *IEA*. https://www.iea.org/data-and-statistics/charts/global-primary-energy-demand-by-fuel-1925-2019

International Monetary Fund (IMF). (2024). IMF Datamapper World Economic Outlook: GDP, current prices [Dataset]. Retrieved from https://imf.org/external/datamapper/NGDPD@WEO/USA

Ioushua, S. S., Hasidim, I., Shayevitz, O., & Feder, M. (2023). Batches stabilize the minimum norm risk in high dimensional overparameterized linear regression. *Cornell University* https://doi.org/10.48550/arVix.2306.08432

Jaiswal, A., & Milios, E. (2023). Breaking the token barrier: Chunking and convolution for efficient long text classification with BERT. *Cornell University* https://doi.org/10.48550/arXiv.2310.20558

Jarrell, G. A., Brickley, J. A., & Netter, J. M. (1988). The market for corporate control: The empirical evidence since 1980. *Journal of Economic Perspectives*, *2*(1), 49–68. https://doi.org/10.1257/jep.2.1.49

Kaske, E., Chesnick, S., Brundage, J., Carroll, R., & Khan, K. (2023). *M&A activity: Sector of the year*. EY. https://www.ey.com/en_us/insights/mergers-acquisitions/m-and-a-activity-report

Katsafados, A. G., Androutsopoulos, I., Chalkidis, I., Fergadiotis, E., Leledakis, G. N., & Pyrgiotakis, E. G. (2021). Using textual analysis to identify merger participants: Evidence from the U.S. Banking Industry. *Finance Research Letters*, *42*, 101949. https://doi.org/10.1016/j.frl.2021.101949

Katsafados, A. G., Leledakis, G. N., Pyrgiotakis, E. G., Androutsopoulos, I., & Fergadiotis, M. (2024). Machine learning in 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

Kedia, S., & Zhou, X. (2014). Informed trading around acquisitions: Evidence from corporate bonds. *Journal of Financial Markets*, *18*, 182–205. https://doi.org/10.1016/j.finmar.2013.07.002

KPMG. (2022). Quarterly Brief: Renewable energy valuation in the global energy transition. https://assets.kpmg.com/content/dam/kpmg/sk/pdf/2022/17th-edition-international-valuation-newsletter.pdf

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., & Soricut, R. (2019). ALBERT: A lite BERT for self-supervised learning of language representations. arXiv:1909.11942. https://doi.org/10.48550/arXiv.1909.11942

Lantiainen, S., & Song, N. (2014). Review and Comparative Analysis of US and EU public policies promoting Wood Energy. In: Dell, G., Egger, C. (Eds.), *World Sustainable Energy Days Next 2014*., pp. 89–94. Springer Vieweg. https://doi.org/10.1007/978-3-658-04355-1_11

Lee, K., Joo, S., Baik, H., Han, S., & In, J. (2020). Unbalanced data, type II error, and nonlinearity in predicting M&A failure. *Journal of Business Research*, *109*, 271–287. https://doi.org/10.1016/j.jbusres.2019.11.083

Leledakis, G. N., & Pyrgiotakis, E. G. (2022). U.S. bank M&As in the post-Dodd–Frank Act era: Do they create value? *Journal of Banking & Finance*, *135*, 105576. https://doi.org/10.1016/j.jbankfin.2019.06.008

Levine, P., & Aaronovitch, S. (1981). The financial characteristics of firms and theories of merger activity. *The Journal of Industrial Economics*, *30*(2), 149–172. https://doi.org/10.2307/2098200

Li, H., & Lu, J. (2023). Can low-carbon mergers and acquisitions reduce carbon emissions? Evidence from China's energy companies. *Energy*, *283*, 129175. https://doi.org/10.1016/j.energy.2023.129175

Liang, X., Li, S., Luo, P., & Li, Z. (2022). Green mergers and acquisitions and green innovation: An empirical study on heavily polluting enterprises. *Environmental Science and Pollution Research International*, *29*(32), 48937–48952. https://doi.org/10.1007/s11356-022-19270-3

Liu, H., Chen, H., & Hu, Y. (2007). Financial characteristics and prediction on targets of M&A based on SOM-Hopfield neural network. In *Proceedings of the 2007 IEEE International Conference on Industrial Engineering and Engineering Management*, 2-5 December 2007, Singapore (pp. 80–-84). Institute of Electrical and Electronic Engineers (IEEE); Currant Associates, Inc. https://doi.org/10.1109/IEEM.2007.4419155

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Lu, J. (2021). Can the green merger and acquisition strategy improve the environmental protection investment of listed company? *Environmental Impact Assessment Review*, *86*, 106470. https://doi.org/10.1016/j.eiar.2020.106470

Lu, J., Li, H., & Wang, G. (2023). The impact of green mergers and acquisitions on illegal pollution discharge of heavy polluting firms: Mechanism, heterogeneity and spillover effects. *Journal of Environmental Management*, *340*, 117973. https://doi.org/10.1016/j.jenvman.2023.117973

Lu, J., Li, H., & Guo, F. (2024). Low-carbon mergers and acquisitions as a driver for higher energy efficiency: Evidence from China's high energy-consuming companies. *Energy*, *290*, 130116. https://doi.org/10.1016/j.energy.2023.130116

Ma, T., Zhang, Y., Huang, L., Shang, L., Wang, K., Yu, H., & Zhu, D. (2017). Text mining to gain technical intelligence for acquired target selection: A case study for China's computer numerical control machine tools industry. *Technological Forecasting and Social Change*, *116*, 162–180. https://doi.org/10.1016/j.techfore.2016.10.061

Malo, P., Sinha, A., Ko-Rhonen, P. J., Wallenius, J., & Takala, P. (2014). Good debt or bad debt: Detecting semantic orientations in economic texts. *JASIST*, *65*(4), 782–796. https://doi.org/10.48550/arXiv.1307.5336

Meghouar, H., & Ibrahimi, M. (2020). Financial characteristics of takeover targets: A French empirical evidence. *EuroMed Journal of Business*, *16*(1), 69–85. https://doi.org/10.1108/EMJB-06-2019-0088

Minixhofer, B., Pfeiffer, J., & Vulić, I. (2023). Compoundpiece: Evaluating and improving decompounding performance of language models. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (EMNLP), 6-10 December 2023, Singapore. Association for Computational Linguistics (ACL); Currant Associates Inc. https://doi.org/10.18653/v1/2023.emnlp-main.24

Morgan, P. E. (2018). Predictive power? Textual analysis in mergers & acquisitions. *Marriott Student Review*, *2*(2), 8. https://scholarsarchive.byu.edu/marriottstudentreview/vol2/iss2/8

Moriarty, R., Ly, H., Lan, E., & McIntosh, S. K. (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

MSCI. (2023). Global industry classification sector (GICS): Definitions of GICS Sectors effective close of March 17, 2023. https://www.msci.com/documents/1296102/11185224/GICS+Sector+Definitions+2023.pdf/822305c6-f821-3d65-1984-6615ded81473?t=1679088764288#:~:text=Energy%20Sector%3A%20The%20Energy%20Sector,oil%20%26%20gas%20equipment%20and%20services

Niemczyk, J., Sus, A., Borowski, K., Jasiński, B., & Jasińska, K. (2022). The dominant motives of mergers and acquisitions in the energy sector in Western Europe from the perspective of Green Economy. *Energies*, *15*(3), 1065. https://doi.org/10.3390/en15031065

Ouzounis, G., Gaganis, C., & Zopounidis, C. (2009). Prediction of acquisitions and portfolio returns. *International Journal of Banking, Accounting and Finance*, *1*(4), 381–406. https://doi.org/10.1504/IJBAAF.2009.023151

Palepu, K. G. (1986). Predicting takeover targets: A methodological and empirical analysis. *Journal of Accounting and Economics*, *8*(1), 3–35. https://doi.org/10.1016/0165-4101(86)90008-X

Parungao, M. E., Galido, A., Suazo, M. L., & Parungao, L. A. (2022). Exploring qualitative data as predictors for M&A: Empirical analysis of target firms' letters to shareholders. *Cogent Business & Management*, *9*(1), 2084970. https://doi.org/10.1080/23311975.2022.2084970

Pasiouras, F., & Tanna, S. (2010). The prediction of bank acquisition targets with discriminant and logit analyses: Methodological issues and empirical evidence. *Research in International Business and Finance*, *24*(1), 39–61. https://doi.org/10.1016/j.ribaf.2009.01.004

Pound, J. (1988). The information effects of takeover bids and resistance. *Journal of Financial Economics*, *22*(2), 207–227. https://doi.org/10.1016/0304-405X(88)90069-4

Powell, R. G. (1997). Modelling takeover likelihood. *Journal of Business Finance & Accounting*, *24*(7–8), 1009–1030. https://doi.org/10.1111/1468-5957.00148

Powell, R., & Yawson, A. (2007). Are corporate restructuring events driven by common factors? Implications for takeover prediction. *Journal of Business Finance & Accounting*, *34*(7–8), 1169–1192. https://doi.org/10.1111/j.1468-5957.2007.02028.x

Rodrigues, B. D., & Stevenson, M. J. (2013). Takeover prediction using forecast combinations. *International Journal of Forecasting*, *29*(4), 628–641. https://doi.org/10.1016/j.ijforecast.2013.01.008

Qiao, L., Dong, W., & Lv, X. (2023). The heterogeneous impacts of M&As on renewable energy firms' innovation: Comparative analysis of China, the US and EU. *International Review of Economics & Finance*, *87*, 306–323. https://doi.org/10.1016/j.iref.2023.04.015

Routledge, B. R., Sacchetto, S., & Smith, N. A. (2017). *Predicting Merger Targets and Acquirers from Text*. Working Paper Carnegie Mellon University. https://sulawesi.tepper.cmu.edu/pdf/ma_ste_latest.pdf

Sen, T. K., & Gibbs, A. M. (1994). An evaluation of the corporate takeover model using neural Networks. *Intelligent Systems in Accounting, Finance and Management*, *3*(4), 279–292. https://doi.org/10.1002/j.1099-1174.1994.tb00071.x

Shen, H., Liang, Y., Li, H., Liu, J., & Lu, G. (2021). Does geopolitical risk promote mergers and acquisitions of listed companies in energy and electric power industries. *Energy Economics*, *95*, 105115. https://doi.org/10.1016/j.eneco.2021.105115

Shi, P., & Huang, Q. (2024). Green mergers and acquisitions and corporate environmental responsibility: Substantial transformation or strategic arbitrage? *Economic Analysis and Policy*, *83*, 1023–1040. https://doi.org/10.1016/j.eap.2024.08.013

Sun, Z., Sun, X., Wang, W., & Wang, W. (2023). Source reduction strategy or end-of-pipe solution? The impact of green merger and acquisition on environmental investment strategy of Chinese heavily polluting enterprises. *Journal of Cleaner Production*, *413*, 137530. https://doi.org/10.1016/j.jclepro.2023.137530

Teti, E., & Tului, S. (2020). Do mergers and acquisitions create shareholder value in the infrastructure and utility sectors? Analysis of market perceptions. *Utilities Policy*, *64*, 101053. https://doi.org/10.1016/j.jup.2020.101053

Tunyi, A. (2021). Fifty years of research on takeover target prediction: A historical perspective. *Qualitative Research in Financial Markets*, *13*(4), 482–502. https://doi.org/10.1108/QRFM-08-2020-0169

Van Nieuwenhuijzen, V., Willner, M., Reinders, S., & Utkarsh, A. (2023). *The US inflation reduction act is driving clean-energy investment one year in*. Goldman Sachs Asset Management. https://www.gsam.com/content/gsam/us/en/institutions/market-insights/ gsam-insights/perspectives/2023/us-inflation-reduction-act-is-driving-clean-energy-investment-one-year-in.html

Varis, D., & Bojar, O. (2021). Sequence length is a domain: Length-based overfitting in Transformer models. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 8246—8257. https://doi.org/10.18653/v1/2021.emnlp-main.650

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30* https://doi.org/10.48550/arXiv.1706.03762

Wei, C., Jiang, Y., & Yang, C. (2008). Patent analysis for supporting merger and acquisition (M&A) prediction: A data mining approach. Markets, services, and networks. In C. Weinhardt, S. Luckner, & J. Stößer (Eds.), *Designing e-business systems. Markets, services, and networks*. WEB 2008 (Vol. 22). Lecture Notes in Business Information Processing (pp. 187–200). Springer. https://doi.org/10.1007/978-3-642-01256-3_16

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L., & Tang, Y. (2023). A brief overview of CHATGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, *10*(5), 1122–1136. https://doi.org/10.1109/JAS.2023.123618

Yang, Q., Zhang, Y., Dai, W., & Pan, S. J. (2020). *Transfer Learning*. Cambridge University Press. https://doi.org/10.1017/9781139061773

Yang, Y., & Chi, Y. (2023). Path selection for enterprises' green transition: Green innovation and green mergers and acquisitions. *Journal of Cleaner Production*, *412*, 137397. https://doi.org/10.1016/j.jclepro.2023.137397

Zanakis, S. H., & Zopounidis, C. (1997). Prediction of Greek company takeovers via multivariate analysis of financial ratios. *Journal of the Operational Research Society*, *48*(7), 678–687. https://doi.org/10.1057/palgrave.jors.2600401

Zhang, M., Johnson, G., & Wang, J. (2012). Predicting takeover success using machine learning techniques. *Journal of Business & Economics Research* , *10*(10), 547–552. https://doi.org/10.19030/jber.v10i10.7264

Zhang, Y., Sun, Z., & Zhou, Y. (2023). Green merger and acquisition and green technology innovation: Stimulating quantity or quality? *Environmental Impact Assessment Review*, *103*, 107265. https://doi.org/10.1016/j.eiar.2023.107265

Zhang, Y., Sun, Z., Sheng, A., Zhang, L., & Kan, Y. (2024). Can green technology mergers and acquisitions enhance sustainable development? Evidence from ESG ratings. *Sustainable Development*, *32*(6), 6072–6087. https://doi.org/10.1002/sd.3004

Zhu, L., Shen, J., & Yeerken, A. (2024). Impact analysis of mergers and acquisitions on the performance of China's new energy industries. *Energy Economics*, *129*, 107189. https://doi.org/10.1016/j.eneco.2023.107189