PSCS184: Financial Causality Detection

A PROJECT REPORT

Submitted by,

MEDHA JEENOOR	**20211CSE0209**
V SAIPRIYA DIPIKA	**20211CSE0178**
MADIHA AZIZ	**20211CSE0075**
AVIJIT SAMANTRAYA	**20211CSE0010**

Under the guidance of,

Dr. Sandeep Albert Mathias

in partial fulfilment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

DECEMBER 2024

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING

CERTIFICATE

This is to certify that the Project report **"Financial Causality Detection"** being submitted by "MEDHA JEENOOR: 20211CSE0209", "V SAIPRIYA DIPIKA: 20211CSE0178", "MADIHA AZIZ: 20211CSE0075", and "AVIJIT SAMANTRAYA: 20211CSE0010", in partial fulfilment of the requirement for the award of the degree of Bachelor of Technology in Computer Science and Engineering is a Bonafide work carried out under my supervision.

Dr. SANDEEP ALBERT MATHIAS
Assistant Professor
School of CSE&IS
Presidency University

Dr. ASIF MOHAMMED H.B
Assistant Professor (SG) & HoD
School of CSE&IS
Presidency University

Dr. L. SHAKKEERA
Associate Dean
School of CSE
Presidency University

Dr. MYDHILI NAIR
Associate Dean
School of CSE
Presidency University

Dr. SAMEERUDDIN KHAN
Pro-VC School of Engineering
Dean -School of CSE&IS
Presidency University

ii

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled **Financial Causality Detection** in partial fulfilment for the award of Degree of **Bachelor of Technology** in **Computer Science and Engineering**, is a record of our own investigations carried under the guidance of **Dr. Sandeep Albert Mathias, Assistant Professor, School of Computer Science Engineering & Information Science, Presidency University, Bengaluru.**

We have not submitted the matter presented in this report anywhere for the award of any other Degree.

MEDHA JEENOOR	**20211CSE0209**
V SAIPRIYA DIPIKA	**20211CSE0178**
MADIHA AZIZ	**20211CSE0075**
AVIJIT SAMANTRAYA	**20211CSE0010**

ABSTRACT

Transformer-based multilingual question answering models are used to detect causality in financial text data. This study employs BERT (Devlin, et al. 2019) for English text and XLM-RoBERTa (Conneau, et al. 2020) for Spanish data, which were fine-tuned on the SQuAD datasets (Rajpur, et al. 2016) (Rajpurkar, Jia and Liang 2018). These pre-trained models are used to extract answers to the targeted questions. We design a system using these pre-trained models to answer questions, based on the given context. The results validate the effectiveness of the systems in understanding nuanced financial language and offers a tool for multi-lingual text analysis. Our system is able to achieve SAS scores of 0.75 in Spanish and 0.82 in English.

iv

ACKNOWLEDGEMENT

First of all, we are indebted to the **GOD ALMIGHTY** for giving us an opportunity to excel in our efforts to complete this project on time. We express our sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Pro-VC, School of Engineering and Dean, School of Computer Science Engineering & Information Science, Presidency University for getting us permission to undergo the project.

We express our heartfelt gratitude to our beloved Associate Deans **Dr. Shakkeera L and Dr. Mydhili Nair,** School of Computer Science Engineering & Information Science, Presidency University, and **Dr. Asif Mohammed H.B**, Head of the Department, School of Computer Science Engineering & Information Science, Presidency University, for rendering timely help in completing this project successfully.

We are greatly indebted to our guide **Dr. Sandeep Albert Mathias, Assistant Professor** and Reviewer **Dr. Vishwanath Y, Professor**, School of Computer Science Engineering & Information Science, Presidency University for their inspirational guidance, valuable suggestions, and for providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We would like to convey our gratitude and heartfelt thanks to the PIP2001 Capstone Project Coordinators **Dr. Sampath A K, Dr. Abdul Khadar A and Mr. Md Zia Ur Rahman,** department Project Coordinators **Mr. Amarnath J L and Dr. Jayanthi Kamalasekaran** and Git hub coordinator **Mr. Muthuraj.**

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

MEDHA JEENOOR
V SAIPRIYA DIPIKA
MADIHA AZIZ
AVIJIT SAMANTRAYA

v

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

ix

x

CHAPTER-1
INTRODUCTION

Finance is important in shaping global economies and influencing decision-making across industries, governments, and individuals. It comprises a broad spectrum of activities, from managing investments and risks to analysing market trends and understanding economic shifts. In such a field, uncovering hidden patterns and relationships within financial data is critical for informed decision-making and strategic planning. Among these patterns, causality is a vital concept, revealing how one financial event or factor can directly influence another. Understanding causal relationships is essential for assessing risks, forecasting trends, and evaluating the impact of decisions. Financial causality deepens our comprehension of economic dynamics and empowers stakeholders to make informed, data-driven choices, fostering transparency and stability in financial ecosystems.

Causality detection, especially in finance, is crucial for understanding events by identifying cause-effect relationships in sources like news articles, earnings reports, and market analyses. In finance, even minor events can trigger significant outcomes, making it essential for analysts to uncover hidden drivers behind market trends, corporate actions, and economic changes. However, traditional Natural Language Processing (NLP) models struggle to detect subtle or implicit causal links, as financial language often involves technical jargon, nuanced phrasing, and indirect references, such as speculative statements. For example, while a central bank policy change may affect stock prices, this connection might not be explicitly stated, requiring deeper analysis to infer the relationship.

The **Financial Causality Detection** project aims to address these challenges by combining causal inference with advanced NLP models, improving the identification of cause-effect relationships in financial documents. Traditional rule-based models rely on explicit cues like "because" or "due to," limiting their ability to capture implicit links spread across complex sentences or multiple documents. Also, the complexity of understanding financial jargon, speculative statements, and the nuanced phrasing of financial texts make this task all the more challenging. The project utilises specialised datasets in English and Spanish to overcome this. The English dataset has been sourced from various 2017 UK financial annual reports from the corpus made available by UCREL at Lancaster University. The Spanish dataset has been extracted from a corpus of Spanish financial annual reports from 2014 to 2018 (further details of the dataset can be found here: https://www.lllf.uam.es/wordpress/fincausal-25/fnp-2025/). These datasets reflect examples of real-world context interpretation challenges like polysemy, implied causality, and industry-specific variability to ensure that the models capture the complexities of financial narratives.

To handle the nuanced nature of financial texts—such as hypothetical or conditional statements ("an increase in inflation could lead to higher interest rates")—the project utilises context embeddings and attention mechanisms for maximum clarity and understanding. In the future, Transfer learning methods could also be used, enabling models to learn from large datasets and adapt to domain-specific contexts. This flexibility ensures that the models can accurately interpret financial scenarios across industries and market conditions.

The project uses advanced transformer-based models, such as BERT (Devlin, et al. 2019) and XLM-RoBERTa (Conneau, et al. 2020), fine-tuning them with financial texts to detect subtle, causal patterns that traditional models might miss. Attention mechanisms will highlight key parts of the text, enhancing the interpretability of causal reasoning for analysts and decision-makers. Contextual Embeddings give a clearer and more appropriate meaning to words depending on the nature of their surrounding sentence, thereby improving the model's ability to detect latent links. This approach will allow the models to identify complex causal chains, such as how geopolitical events ripple through markets, providing deeper insights into financial dynamics.

XLM-RoBERTa (Unsupervised Cross-lingual Language Model Representation or XLM-R), a multilingual variant of RoBERTa, is designed to handle diverse languages with robust natural language understanding capabilities. Financial texts in Spanish often present unique challenges due to the use of domain-specific terminology, complex sentence structures, and regional linguistic variations. By utilising XLM-RoBERTa, this model achieves an accurate extraction of cause-effect relationships in Spanish financial documents. This approach ensures an improved understanding of causal factors in financial reports and handles the added complexity of context preservation of texts across other languages.

By enabling the detection of causal relationships from a range of multilingual financial texts, the system improves the efficiency of financial analysis and opens up new possibilities for more informed, data-driven decision-making across the global financial landscape.

CHAPTER-2
LITERATURE SURVEY

2.1 Introduction

Understanding causality in financial systems is essential for explaining events, forecasting trends, and making decisions in volatile markets. This survey examines existing methodologies for financial causality detection, focusing on text mining, machine learning, and deep learning approaches. These methods are applied to both structured and unstructured financial datasets, such as financial news and reports, to uncover cause-effect relationships.

2.2 Analysis of Existing Methods

2.2.1 Text Mining and Sentiment Analysis

Methodology: Text mining and sentiment analysis involve extracting useful information from large amounts of text data, such as financial news, and categorizing sentiment (positive, negative, or neutral). The approaches used machine learning models like Support Vector Machines (SVM) combined with lexicons for sentiment detection and stock market prediction.

- **Advantages**: Automates sentiment classification, reduces manual effort, and links sentiment with market trends.
- **Disadvantages**: Struggles with nuanced language (e.g., sarcasm), relies on high-quality input data, and is computationally intensive for large datasets.

Evaluation:
(Feder, et al. 2022) discussed how sentiment analysis can provide predictive insights into market trends. However, these methods, while powerful, often fail to uncover underlying causal triggers. For instance, while sentiment might influence stock prices, it cannot explain the root causes such as regulatory changes or market shocks.

2.2.2 Causal Inference in Machine Learning

Methodology: Causal inference methods like Granger Causality and Bayesian Networks are employed to detect causal relationships between financial events and their outcomes. These methods examine time series data to predict future financial behavior based on past patterns.

- **Advantages**: Helps in discovering cause-effect relationships, supports better decision-making, and handles complex interactions.
- **Disadvantages**: Requires large datasets and is assumption-driven, which can lead to errors if assumptions are violated.

Evaluation:
(Kumar, et al. 2023) highlighted how Granger Causality and Bayesian Networks can model the impact of past financial events on future outcomes. However, their application is limited when it comes to detecting nonlinear relationships and working with unstructured data, which deep learning models can address.

2.2.3 Causality in Natural Language Processing (NLP)

Methodology: NLP models extract causal relationships from financial texts using rule-based systems, machine learning, and deep learning. Recent advancements have incorporated Causal Graphical Models and counterfactual reasoning to estimate causal effects.

- **Advantages**: Extracts meaning from unstructured text, improves interpretability by explaining causal triggers, and processes large datasets efficiently.
- **Disadvantages**: High dimensionality of text data requires powerful models, and NLP models can struggle with implicit causality and language ambiguity.

Evaluation:
(Yang, Han and Poon 2021) discusses how NLP can detect causal relationships from financial reports, but the challenge remains in identifying implicit causality that is not explicitly stated in text.

2.2.4 Deep Learning for Financial Sentiment and Causality Detection

Methodology: Deep learning models, including Recurrent Neural Networks (RNNs) and transformers, detect sentiment and causality in financial documents. These models predict financial outcomes and identify causal triggers from news and reports.

- **Advantages**: High accuracy, multilingual capabilities, and automated detection of complex patterns.
- **Disadvantages**: Data-hungry models are prone to overfitting, and their “black-box” nature makes interpretation challenging.

Evaluation:
(Day and Lee 2016) demonstrated that deep learning significantly improves the detection of causal triggers in financial news. However, their high computational cost and limited interpretability remain major challenges.

2.2.5 Knowledge-Based Approaches

Methodology: Knowledge-based methods use manually crafted rules and predefined syntactic or semantic patterns (e.g., causal connectives like "because" or "since") to identify causal relationships.

- **Advantages**: Transparent, interpretable, and precise when applied to domain-specific texts.
- **Disadvantages**: Scalability is an issue, and the rules become inflexible with changing language patterns.

Evaluation:
(Sakaji and Izumi 2023) showed that knowledge-based systems could precisely capture causal relationships in financial reports, but they are not adaptable to new or large-scale datasets. This limitation is why hybrid systems that combine rule-based and machine learning approaches are often favored.

2.2.6 Statistical Machine Learning (ML)-Based Approaches

Methodology: These approaches rely on features extracted from text (e.g., syntactic structures, word embeddings) and use classifiers like SVM or Naïve Bayes (NB) to detect causal relationships.

- **Advantages**: Automates feature extraction and is adaptable to various data types and languages.
- **Disadvantages**: Requires domain-specific feature engineering, which may still require expertise.

Evaluation:
(Man, Nguyen and Nguyen 2022) found that machine learning methods like SVM could be trained on diverse datasets for accurate causality detection. However, these methods require human intervention for feature extraction and are often opaque in how they arrive at decisions.

2.2.7 Deep Learning Approaches

Methodology: These methods, including transformer models like BERT (Devlin, et al. 2019) and T5, utilize pre-trained language models and attention mechanisms to capture complex, contextual dependencies in text.

- **Advantages**: High performance, multilingual support, and the ability to detect nested causalities.
- **Disadvantages**: “Black-box” models with high computational costs and data dependencies.

Evaluation:
(Scaramozzino, Cerchiello and Aste 2021) demonstrated that deep learning models like BERT significantly improved causality detection, but their complexity and lack of interpretability limit practical applications in financial systems.

2.2.8 Granger Causality

Methodology: Granger causality tests whether historical values of one variable predict future values of another, analyzing time-series data to identify directional relationships.

- **Advantages**: Simple to apply, ideal for linear data relationships, and helps in forecasting financial outcomes.
- **Disadvantages**: Limited to linear relationships and unsuitable for real-time decision-making.

Evaluation:
(Hong 2009) used Granger causality to analyze financial contagion, but its linear nature makes it inadequate for capturing more complex financial systems.

2.2.9 Transfer Entropy

Methodology: Transfer entropy measures the amount of information flow between variables, capturing both linear and nonlinear dependencies.

- **Advantages**: Captures nonlinear relationships, bidirectional causality, and is resilient to moderate noise.
- **Disadvantages**: Computationally expensive and sensitive to high noise levels.

Evaluation:
(Zaremba and Aste 2014) highlighted transfer entropy's ability to capture complex causal relationships between market sentiment and stock prices, though it still faces challenges with noisy data.

2.2.10 Causal Link Detection Between Financial News and Stock Movements

Methodology: This method uses text and time-series metrics like TF-IDF for news and Pearson correlation for stock prices to detect causal links.

- **Advantages**: Offers predictive capabilities and integrates different data types.
- **Disadvantages**: Challenges with time-lag issues and limited insight into true causality.

Evaluation:
(Qu and Kazakov 2019) applied this approach effectively for predicting stock movements from news data, but they noted that this model cannot always distinguish between correlation and true causality.

2.2.11 Deep Learning for Causality Detection (FinCausal Task)

Methodology: Deep learning models like BERT (Devlin, et al. 2019) and RoBERTa (Liu 2019) detect both explicit and implicit causal links in financial documents, even when causality is spread across multiple sentences.

- **Advantages**: Processes complex, unstructured text and captures implicit relationships.
- **Disadvantages**: Requires large training datasets and computational power.

Evaluation:
(Mariko, et al. 2020) applied these models in the FinCausal task to detect complex causal relationships in financial documents. The challenge remains in model transparency and scalability.

2.2.12 Kernel-Based Methods

Methodology: Kernel-based methods extend Granger causality by using kernel functions to detect nonlinear relationships in financial data.

- **Advantages**: Effective for complex, nonlinear financial systems.
- **Disadvantages**: Requires extensive tuning and high computational resources.

Evaluation:
(Sakaji and Izumi 2023) discussed how kernel-based methods can identify subtle nonlinear relationships, though the complexity of kernel tuning remains a significant drawback.

2.2.13 Deep Learning Models (Causal-BERT)

Methodology: Causal-BERT uses transformer-based models to detect cause-effect relationships from unstructured financial text.

- **Advantages**: Identifies implicit and explicit causal patterns, highly effective for financial reports.
- **Disadvantages**: Lack of interpretability and requires large, labeled datasets.

Evaluation:
(Khetan, et al. 2020) demonstrated that Causal-BERT is highly effective in identifying causal links in financial texts, though it still faces challenges with model transparency.

2.3 Conclusion

This review illustrates the broad range of methods used in financial causality detection, each with its own strengths and limitations. While traditional methods like Granger causality remain foundational, deep learning approaches such as Causal-BERT are pushing the boundaries by handling complex, unstructured datasets. Despite advancements, challenges such as computational demands, model interpretability, and data dependency persist. Future research should focus on integrating these methods, leveraging their strengths to create more robust financial causality detection frameworks.

CHAPTER-3
RESEARCH GAPS OF EXISTING METHODS

3.1. Introduction

Advancements in sentiment analysis, causality detection, and machine learning have revolutionized how data is interpreted in fields such as finance, healthcare, and natural language processing (NLP). The ability to understand sentiment and establish causal relationships is critical for tasks like predicting stock movements, evaluating policy impacts, and automating decision-making processes.

Despite these advancements, the methodologies employed face significant limitations. This chapter systematically analyzes the gaps in existing methods, highlighting their inadequacies in addressing the dynamic and complex nature of real-world problems. The insights gathered establish the foundation for further research to overcome these limitations.

3.2. Detailed Analysis of Methods and Research Gaps

3.2.1 Text Mining and Sentiment Analysis

- **Strengths**: This method efficiently processes vast amounts of textual data to extract sentiment information, identifying trends and patterns in unstructured data.
- **Weaknesses**:
    - Limited to correlation-based insights; lacks mechanisms to determine causality.
    - Struggles with domain-specific sentiment expressions, such as in financial or medical texts.
- **Research Gap**:
    - There is a pressing need to integrate causality detection frameworks into text mining to bridge the gap between sentiment trends and their causal impacts.

3.2.2 Causal Inference in Machine Learning

- **Strengths**: Provides a robust statistical framework for identifying cause-effect relationships, particularly useful for structured datasets.

- **Weaknesses**:
  - Poor scalability when applied to high-dimensional, unstructured data such as text or images.
  - Assumes predefined causal structures, which may not hold in dynamic environments.
- **Research Gap**:
  - Future research must focus on enhancing causal inference techniques for unstructured datasets and automating causal structure discovery.

3.2.3 Causality in Natural Language Processing (NLP)

- **Strengths**: This method contextualizes causality within text, helping to understand cause-effect relationships embedded in language.
- **Weaknesses**:
  - Highly dependent on annotated datasets, which are scarce and expensive to create.
  - Struggles to disambiguate causality in linguistically complex sentences.
- **Research Gap**:
  - Developing self-supervised or unsupervised learning techniques to reduce dependence on labeled data is essential.

3.2.4 Deep Learning for Financial Sentiment and Causality Detection

- **Strengths**: Achieves high accuracy in identifying sentiment and causal relationships in financial narratives when sufficient training data is available.
- **Weaknesses**:
  - Overfits to domain-specific datasets, limiting its ability to generalize across financial contexts.
  - Computationally expensive, making real-time applications challenging.
- **Research Gap**:
  - There is a need for lightweight, generalizable deep learning models that can adapt to various financial scenarios in real time.

3.2.5 Knowledge-Based Approaches

- **Strengths**: Leverages domain expertise and pre-defined knowledge bases to enhance understanding of causality and sentiment.
- **Weaknesses**:
  - Highly reliant on the quality and comprehensiveness of the knowledge base.
  - Cannot adapt dynamically to new or evolving scenarios.
- **Research Gap**:
  - Future research must integrate dynamic learning mechanisms to update knowledge bases in real-time.

3.2.6 Statistical Machine Learning (ML)-Based Approaches

- **Strengths**: Well-suited for structured datasets with predefined relationships.
- **Weaknesses**:
  - Assumes linearity, which is often unrealistic in complex systems like finance.
  - Struggles with multi-modal data integration (e.g., combining text and numerical data).
- **Research Gap**:
  - Addressing non-linearities and integrating multi-modal data for robust causal analysis remains an open challenge.

3.2.7 Deep Learning Approaches

- **Strengths**: Achieves state-of-the-art results in sentiment analysis and causal detection for diverse datasets.
- **Weaknesses**:
  - Black-box nature limits interpretability, which is crucial in high-stakes applications like finance.
  - Requires extensive labeled data for training, leading to resource constraints.
- **Research Gap**:
  - Developing interpretable and resource-efficient deep learning models is an urgent priority.

3.2.8 Granger Causality

- **Strengths**: Provides a robust statistical framework for analyzing time-series data.
- **Weaknesses**:
    - Assumes linear relationships and stationarity, which may not hold in financial contexts.
    - Inadequate for analyzing non-linear, high-dimensional data.
- **Research Gap**:
    - Extending Granger causality to handle non-linear and non-stationary time-series data is a critical area of research.

3.2.9 Transfer Entropy

- **Strengths**: Captures non-linear relationships in time-series data.
- **Weaknesses**:
    - Computationally intensive, making it unsuitable for real-time applications.
    - Requires large datasets for reliable estimates.
- **Research Gap**:
    - Research should focus on optimizing transfer entropy for real-time, resource-constrained environments.

3.2.10 Causal Link Detection Between Financial News and Stock Movements

- **Strengths**: Effective at correlating financial narratives with market trends.
- **Weaknesses**:
    - Often assumes causality without rigorous testing, leading to potential overfitting.
    - Limited applicability to diverse financial instruments or markets.
- **Research Gap**:
    - Incorporating robust causal testing frameworks to validate detected links is essential.

3.2.11 Deep Learning for Causality Detection (FinCausal Task)

- **Strengths**: Provides task-specific models for causal detection in financial text.

- **Weaknesses**:
  - Lacks generalization beyond the specific dataset it was trained on.
  - Computationally demanding for large-scale applications.
- **Research Gap**:
  - Developing domain-agnostic models with reduced computational costs is an important area of focus.

3.2.12 Kernel-Based Methods

- **Strengths**: Flexible and robust for detecting relationships in high-dimensional data.
- **Weaknesses**:
  - Requires careful parameter tuning, which can be computationally expensive.
  - Sensitive to noise in datasets.
- **Research Gap**:
  - Research should focus on improving robustness to noise and automating parameter optimization.

3.2.13 Deep Learning Models (Causal-BERT)

- **Strengths**: Incorporates transformer-based architectures for improved causality detection in text.
- **Weaknesses**:
  - Computationally expensive and requires large-scale annotated datasets.
  - Black-box nature limits interpretability.
- **Research Gap**:
  - Enhancing Causal-BERT with interpretability and efficiency features is essential for practical adoption.

3.3. Cross-Method Gaps

Across the analyzed methods, recurring gaps appear:

1. **Scalability and Efficiency**: Many methods face challenges in scaling to real-world, high-dimensional datasets or adapting to new domains.
2. **Generalization**: Methods often overfit to specific datasets, limiting their broader applicability.

3. **Interpretable Models**: The need for transparent models is critical, particularly in fields like finance and healthcare.
4. **Integration of Causality and Sentiment**: Few methods effectively combine sentiment analysis with rigorous causality detection.

3.4. Conclusion

The analysis of these methods reveals significant gaps that hinder their effectiveness in addressing complex, real-world challenges. Many of the approaches struggle to balance scalability, interpretability, and efficiency, which are critical for practical applications. Scalability remains a key concern, as methods often falter when applied to high-dimensional, dynamic datasets or large-scale real-time environments. Similarly, interpretability is a persistent issue, especially with deep learning-based approaches, where the "black-box" nature of models limits their utility in fields like finance, where understanding the reasoning behind predictions is crucial.

Furthermore, the integration of sentiment analysis and causality detection remains an underexplored area. While some methods excel in one aspect, they often fall short in providing a comprehensive framework that combines both. This disconnect highlights the need for hybrid approaches that can effectively bridge these domains, offering richer insights and actionable outcomes.

Efficiency is another challenge, particularly for computationally intensive models such as deep learning. These methods require substantial resources, making them less accessible for smaller organizations or real-time applications. Optimizing these methods for resource efficiency without compromising performance is an area that demands immediate attention.

CHAPTER-4
PROPOSED METHODOLOGY

4.1 Resources Used

The resources used for the question-answering tasks in our project involve:

- Transformer-based pre-trained models (E.g. BERT (Devlin, et al. 2019) and XLM-RoBERTa (Conneau, et al. 2020)) for generating the answers for the provided context-question pairs.
- Python libraries for input and output data processing in CSV format.
- Transformers libraries from Hugging Face (Wolf, et al. 2020) for accessing and executing the QA pipelines.

Language	**Pre-trained Model Used**
Spanish	deepset/xlm-roberta-large-squad2
English	bert-large-uncased-whole-word-masking-finetuned-squad

Table 1: Specific resources used for the Spanish and English datasets

Different resources were used due to linguistic requirements as shown in Table 1.
For Spanish, deepset/xlm-roberta-large-squad2 was utilized which is a large variant of the XLM-RoBERTa model (Conneau, et al. 2020). It is pre-trained with data in 100 languages. It is fine-tuned on SQuAD2.0 (Rajpurkar, Jia and Liang 2018) including questions which may not have answerable contexts, thereby enhancing its ability to handle such challenging situations.
For English, bert-large-uncased-whole-word-masking-finetuned-squad was utilized which is a pre-trained BERT (Devlin, et al. 2019) model on SQuAD (Rajpur, et al. 2016).

4.2 Parameter Settings

- To ensure compatibility with resource-constrained environments, the Spanish QA model was executed on a CPU (device = -1)
- To work with the available hardware capabilities, the default device was used for the English QA model.

4.3 Workflow

- **Data Ingestion**: Reading and parsing the input CSV files is done using Python's csv library. To ensure completeness, the rows with less than three columns are identified and skipped. The three columns are namely "ID", "Text" and "Question".
- **Question-Answering Pipeline**: The question-context pair for each valid row is extracted and passed though the QA pipelines XLM-RoBERTa (Conneau, et al. 2020) and BERT (Devlin, et al. 2019) for Spanish and English respectively.
- **Answer Extraction**: The most probable answer was produced by the pipeline for the question, based on the context provided.
- **Output Generation**: The answer is appended to the valid row as a new column "Answer". The resulting row was stored in new CSV file.
- This is an **iterative process**.

For example, consider that we have the following row from the English dataset: "1882.b;Underlying Group EBITDA declined by 10.1% to £10.0m (2016: £11.2m). This decline has been driven by an increase in UK overheads of £1.0m (5.6%) due to investment in support of our strategic initiatives and well-publicised cost headwinds.;What has motivated the increase in UK overheads by £1.0 million or 5.6%?".

Our system will generate the line: "1882.b;Underlying Group EBITDA declined by 10.1% to £10.0m (2016: £11.2m). This decline has been driven by an increase in UK overheads of £1.0m (5.6%) due to investment in support of our strategic initiatives and well-publicised cost headwinds.;What has motivated the increase in UK overheads by £1.0 million or 5.6%?;investment in support of our strategic initiatives."
The workflow has been depicted in Figure 1.

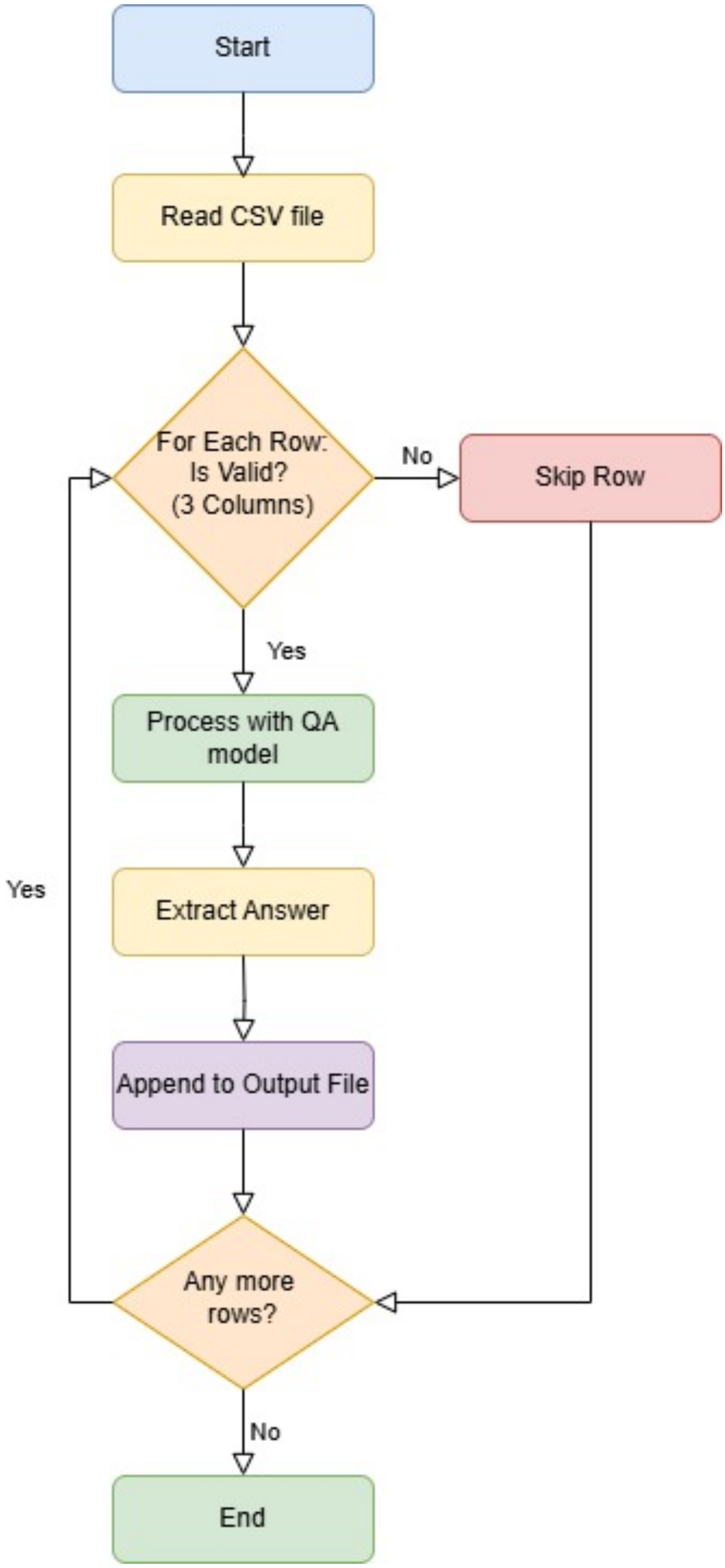


Figure 1: Workflow of our system

CHAPTER-5
OBJECTIVES

5.1 Enhancing Causal Extraction Using Universal Dependencies and Hint Expressions

By using Universal Dependencies and hint expressions to capture syntactic and semantic information while maintaining interpretability, the objective is to improve causal extraction. The method, which addresses issues like context sensitivity, polysemy, and implicit causation, will be tested on a variety of datasets, including news and financial texts, with applications in event analysis and finance.
(Man, Nguyen and Nguyen 2022), (Sakaji and Izumi 2023), (Yang, Han and Poon 2021), (Mariko, et al. 2020)

5.2 Improving Detection of Implicit Cause-Effect Interactions in NLP Models

By tackling the difficulties in detecting latent cause-effect links, improving models such as BERT, and utilizing datasets and methods that better capture these complexities, this project seeks to improve the identification of implicit causal interactions.
(Khetan, et al. 2020)

5.3 Advancing Financial Text Analysis through Causality Detection and Inference

By identifying and elucidating cause-effect links, improving comprehension of financial results, and developing NLP with specialized causality detection tools, this research focuses on applying causal inference to financial texts.
(Feder, et al. 2022), (Scaramozzino, Cerchiello and Aste 2021), (Mariko, et al. 2020), (Qu and Kazakov 2019)

5.4 Exploring Causal Inference in Financial Markets and Policy Impact Analysis

To explore the potential of **causal inference** in financial sectors like banking and insurance, assessing the impact of economic policies and external shocks on financial markets, and aim to provide clearer insights into complex financial interactions and support informed decision-making.

(Kumar, et al. 2023) **,** (Zaremba and Aste 2014)

5.5 Developing Generative AI Systems for Financial Causality Detection and Stock Market Prediction

To develop a generative AI system for **Financial Causality Detection**, identifying and explaining cause-effect relationships in English and Spanish datasets, and aim to refine methods for detecting causal links between financial news and stock market performance, moving beyond correlation-based models to predict stock behavior.

(Feder, et al. 2022) **,** (Day and Lee 2016)**,** (Kumar, et al. 2023)**,** (Qu and Kazakov 2019)

CHAPTER-6
SYSTEM DESIGN & IMPLEMENTATION

6.1 Introduction

This chapter provides a comprehensive overview of the architecture, components, and implementation of the Financial Causality Detection project. The system is thoroughly designed to identify causal relationships within financial documents by leveraging advanced natural language processing (NLP) techniques and transformer-based models. With its ability to process multilingual financial datasets, the system extracts significant cause-effect relationships, thereby enhancing financial decision-making and risk assessment processes.

6.2 System Architecture

The system is designed with modular components for scalability and efficiency, comprising the following key elements:

6.2.1 Data Ingestion

- Sources include English and Spanish financial datasets sourced from UCREL and LLLF corpora.
- **Process:** Data collection from specified repositories.

6.2.2 Modelling Layer

- **Transformer Models:** Fine-tuning BERT (Devlin, et al. 2019) for English and XLM-RoBERTa (Conneau, et al. 2020) for Spanish texts.
- **Attention Mechanism:** Highlights context-relevant sections for causal link detection.
- **Embedding Techniques:** Contextual embeddings ensure nuanced understanding of text.

6.2.3 Causal Inference Module

- Combines linguistic patterns with statistical inference for accurate detection of implicit and explicit causal relationships.
- Outputs: CSV file with the predicted answers.

6.2.4 Evaluation

- Metrics like SAS (Risch, et al. 2021) is used for model evaluation.

6.3 Hardware and Software Requirements

6.3.1 Hardware

- **CPU:** For faster processing of the models and for handling file operations.
- **Memory (RAM):** For loading and inference.
- **Storage:** For input and output CSV files.

6.3.2 Software

- Programming Language: Python 3.9.
- Libraries:
  - Transformers (Hugging Face (Wolf, et al. 2020) for BERT/XLM-RoBERTa).
  - CSV Module
  - File I/O Utilities.

6.4 System Workflow

6.4.1 Data Ingestion

- **Objective:** Load and preprocess input data for analysis.
- **Steps:**
  1. **Load Input CSV File:** Read the financial dataset containing columns such as ID, Context, and Question.
  2. **Extract Relevant Rows:** Filter rows required for the question-answering process.

6.4.2 Question-Answering Pipeline

- **Objective:** Utilize a pre-trained BERT (Devlin, et al. 2019) model for causal relationship detection.

- **Steps:**
    1. **Load the Model and Tokenizer:** Use the Transformers library to initialize the BERT pipeline for question-answering tasks.
    2. **Define answer_question():**
        - Accept Question and Context as input.
        - Use the BERT pipeline to predict causal relationships as answers.

6.4.3 Answer Extraction

- **Objective:** Extract predicted answers for each row in the input dataset.
- **Steps:**
    1. **Iterate Through Rows:** For each row in the input dataset:
        - Pass the Question and Context to answer_question().
        - Store the predicted answer.

6.4.4 Output Generation

- **Objective:** Save the processed data with predictions to a CSV file.
- **Steps:**
    1. **Add Predicted Answers:** Append a new column, Predicted Answer, to the dataset.
    2. **Save Output File:** Write the enriched dataset to a new CSV file.
    3. **Provide File Path:** Display or return the location of the output file.

6.5 Conclusion

This chapter presented a detailed account of the design and implementation of the Financial Causality Detection system. By integrating advanced transformer-based NLP models with principles of causal inference, the system delivers precise and meaningful extraction of causal relationships from financial texts. This capability empowers stakeholders to make well-informed decisions, ultimately fostering greater financial transparency and stability.

CHAPTER-7

TIMELINE FOR EXECUTION OF PROJECT (GANTT CHART)

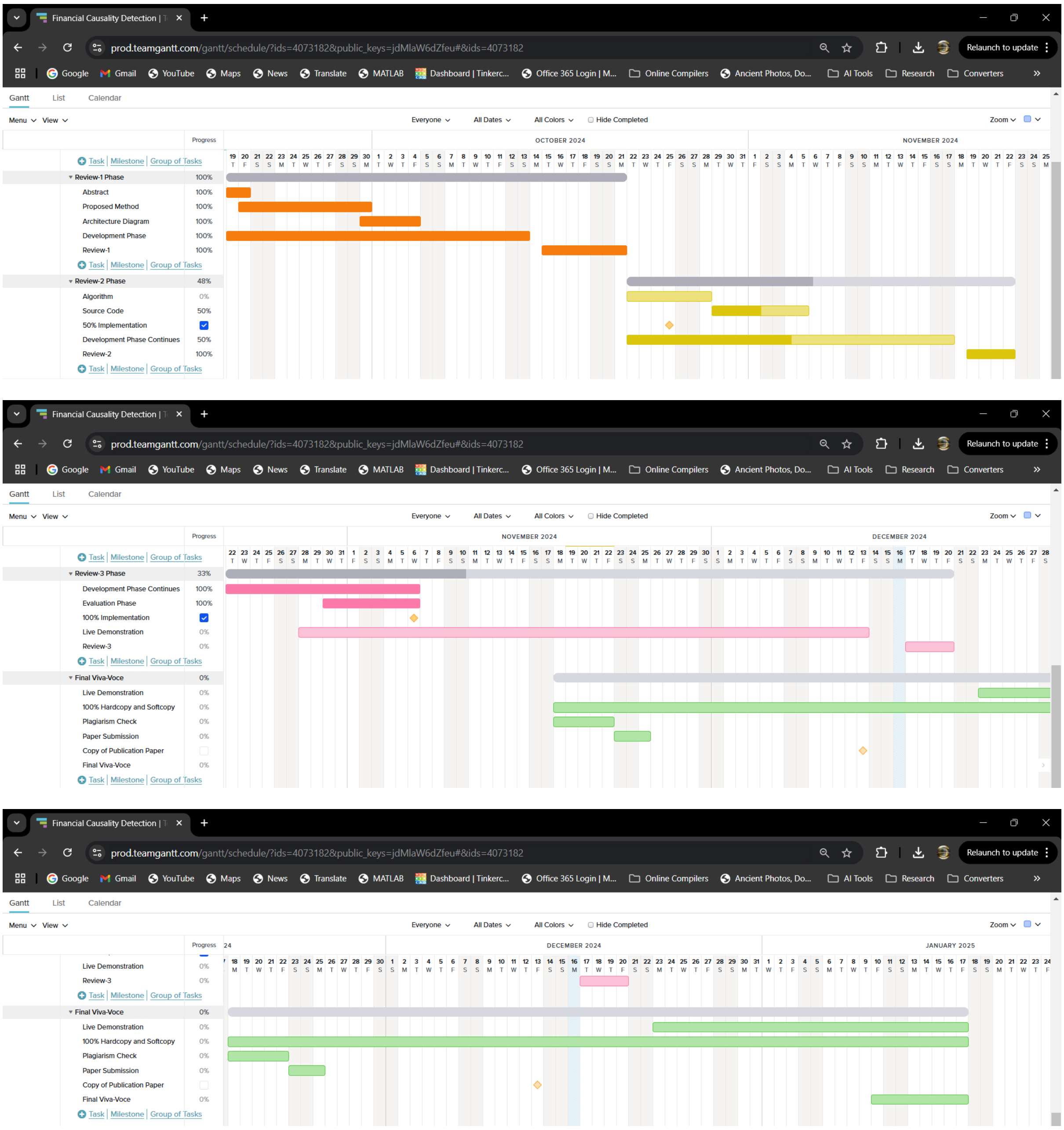


Figure 2 Timeline of the Project

For more details refer to the website:

https://prod.teamgantt.com/gantt/schedule/?ids=4073182&public_keys=jdMlaW6dZfeu#&ids=4073182

CHAPTER-8
OUTCOMES

8.1 CSV File Generated with the Predicted Answers to the Questions:

The output of the model was saved in a structured CSV format, with each row capturing the input and output details. This provides a clear record of the predictions for further analysis.

8.2 Each Row Contains the Id, Text, Question, and the Predicted Answer:

Each row in the CSV file includes four key fields: the unique identifier (**Id**) for each question, the **Text** providing the context, the **Question**, and the corresponding **Predicted Answer** generated by the model.

8.3 Causal Relationships are Extracted from Financial Texts in Spanish and English:

The models effectively identified and extracted causal relationships embedded within financial texts in both English and Spanish.

8.4 Model was Evaluated Using the Metrics SAS and EM:

The evaluation employed **Exact Match (EM)** for assessing the precision of verbatim answers and **Semantic Answer Similarity (SAS)** for capturing semantic alignment between the predictions and ground truth.

8.5 The project contributes to 'Decent Work and Economic Growth' (SDG 8):

By improving the analysis and understanding of financial documents, it aids in identifying the factors driving economic changes. By detecting causal relationships in financial data, the project helps businesses and financial institutions make informed decisions, fostering sustainable economic growth and promoting transparency, leading to better work conditions and economic stability.

Summary of Language Models

Aspect	English Model	Spanish Model
Model Name	bert-large-uncased-whole-word-masking-finetuned-squad	deepset/xlm-roberta-large-squad2
Model Type	BERT-based QA model	XLM-RoBERTa
Tokenizer Used	**Same as model**: bert-large-uncased-whole-word-masking-finetuned-squad	**Same as model**: deepset/xlm-roberta-large-squad2
Pipeline Type	question-answering	question-answering
Primary Dataset Used for Fine Tuning	SQuAD (Stanford Question Answering Dataset)	SQuAD 2.0 (multilingual version)

Table 2: Summary of our Selected Models

CHAPTER-9
RESULTS AND DISCUSSIONS

9.1 Summary of Results

The evaluation phase involved iterative submissions designed to enhance the pipeline's performance for English and Spanish tasks, with each iteration addressing specific challenges and refining the approach. This iterative process focused on improving the detection of causal relationships in financial texts, utilizing advanced question-answering models and multilingual translation tools.

Key Highlights

i. **Performance on English Task**

- The deepset/roberta-base-squad2 model, fine-tuned on the SQuAD2.0 (Rajpurkar, Jia and Liang 2018) dataset, delivered outstanding results, excelling in capturing both explicit and implicit causal relationships.
- It achieved a Semantic Answer Similarity (SAS) score of 82.41%, demonstrating its ability to align predicted answers with the ground truth effectively.

ii. **Performance on Spanish Task**

- The deepset/xlm-roberta-large-squad2 model translation proved highly effective for multilingual causality detection.
- This approach yielded impressive results, achieving an SAS (Risch, et al. 2021) score of 75.20%, making it a robust solution for cross-lingual question answering.

9.2 Error Analysis

While the system demonstrated strong performance overall, certain limitations were observed:

i. **Ambiguous Causality**: Sentences containing overlapping or multiple causal relationships introduced ambiguity, resulting in inconsistent model predictions.

ii. **Translation Noise**: The process of translating Spanish text to English sometimes introduced errors, leading to inaccuracies in the extracted causal relationships.

iii. **Implicit Causality**: The system faced challenges with detecting highly nuanced or context-dependent causal relationships that required a deeper understanding of linguistic subtleties in both Spanish and English.

9.3 Explainability Insights

i. High attributions were given to explicit causal keywords like "*caused by*" and "*led to*".
ii. Implicit causality often relied on contextual embeddings, showcasing the model's deeper understanding of language semantics.

9.4 Practical Applications

The refined pipeline offers a wide range of practical applications:

i. **Risk Assessment**: Identifying causal relationships to predict market trends and mitigate financial risks.

ii. **Policy Impact Analysis**: Understanding the effects of economic policies or global events on markets.

iii. **Cross-Lingual Utility**: Ensuring consistent causality detection across English and Spanish datasets supports global financial institutions.

CHAPTER-10
CONCLUSION

10.1 Summary of Achievements

This project successfully developed and evaluated a hybrid pipeline for financial causality detection using state-of-the-art language models and machine translation tools. The system demonstrated:

i. High accuracy in detecting explicit and implicit causal links.
ii. Effective multilingual processing using XLM-RoBERTa (Conneau, et al. 2020) and BERT (Devlin, et al. 2019) models.

10.2 Broader Implications

The developed pipeline is a step forward in leveraging AI for financial decision-making. By providing accurate and interpretable causal insights, it bridges the gap between raw data and actionable strategies for stakeholders, such as financial analysts and policymakers.

10.3 Challenges and Limitations

i. **High Computational Requirements**: Deep learning models, especially for multilingual tasks, require significant computational resources.
ii. **Translation-Dependent Performance**: The reliance on machine translation for Spanish texts introduced minor inaccuracies.

10.4 Future Directions

i. **Real-Time Processing**: Implementing real-time causality detection for high-frequency financial data.
ii. **Domain-Specific Fine-Tuning**: Fine-tuning models on larger, domain-specific datasets for enhanced accuracy.
iii. **Interactive Dashboards**: Building visualization tools for stakeholders to explore and interpret causal relationships dynamically.

REFERENCES

Conneau, Alexis, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. "Unsupervised Cross-lingual Representation Learning at Scale." Edited by Joyce Chai, Natalie Schluter, and Joel Tetreault Dan Jurafsky. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics.* Online: Association for Computational Linguistics. 8440-8451. doi:10.18653/v1/2020.acl-main.747.

Day, Min-Yuh, and Chia-Chou Lee. 2016. "Deep learning for financial sentiment analysis on finance news providers." 1127-1134.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers).* Minneapolis, Minnesota: Association for Computational Linguistics. 4171-4186. doi:10.18653/v1/N19-1423.

Feder, Amir, Katherine A Keith, Emaad Manzoor, Reid Pryzant, Dhanya Sridhar, Zach Wood-Doughty, Jacob Eisenstein, et al. 2022. "Causal Inference in Natural Language Processing: Estimation, Prediction, Interpretation and Beyond." *Transactions of the Association for Computational Linguistics.* Cambridge, MA: MIT Press. 1138–1158.

Hong, Yongmiao and Liu, Yanhui and Wang, Shouyang. 2009. "Granger causality in risk and detection of extreme risk spillover between financial markets." *Journal of Econometrics* (Elsevier) 150: 271-287.

Khetan, Vivek, Roshni Ramnani, Mayuresh Anand, Shubhashis Sengupta, and E, Andrew Fano. 2020. "Causal BERT: Language models for causality detection between events expressed in text." arXiv preprint arXiv:2012.05453.

Kumar, Satyam, Yelleti Vivek, Vadlamani Ravi, and Indranil Bose. 2023. "Causal Inference for Banking Finance and Insurance A Survey."

Liu, Yinhan. 2019. "Roberta: A robustly optimized bert pretraining approach." *arXiv preprint* 364. https://arxiv.org/abs/1907.11692.

Man, Hieu, Minh Nguyen, and Thien Nguyen. 2022. "Event Causality Identification via Generation of Important Context Words." *Proceedings of the 11th Joint Conference on Lexical and Computational Semantics.* Seattle, Washington: Association for Computational Linguistics. 323–330.

Mariko, Dominique, Abi, Hanna Akl, Estelle Labidurie, Stephane Durfort, De, Hugues Mazancourt, and Mahmoud El-Haj. 2020. "Financial document causality detection shared task (fincausal 2020)." *The 28th International Conference on Computational Linguistics (COLING'2020).* Barcelona, Spain: arXiv preprint arXiv:2012.02505.

Qu, Haizhou, and Dimitar Kazakov. 2019. "Detecting Causal Links between Financial news and stocks." *2019 IEEE Conference on Computational Intelligence for Financial Engineering \& Economics (CIFEr).* IEEE. 1-8.

Rajpur, Pranav, Jian Zhang, Konstantin Lopyrev, and Percy Liang. 2016. "SQuAD: 100,000+ Questions for Machine Comprehension Text." Edited by Jian Su, Kevin Duh and Xavier Carreras. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing.* Austin, Texas: Association for Computational Linguistics. 2383-2392. doi:10.18653/v1/D16-1264.

Rajpurkar, Pranav, Robin Jia, and Percy Liang. 2018. "Know What You Don't Know: Unanswerable Questions for SQuAD." Edited by Iryna Gurevych and Yusuke Miyao. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers).* Melbourne, Australia: Association for Computational Linguistics. 784-789. doi:10.18653/v1/P18-2124.

Risch, Julian, Timo Möller, Julian Gutsch, and Malte Pietsch. 2021. "Semantic Answer Similarity for Evaluating Question Answering Models." *Proceedings of the 3rd Workshop on Machine Reading for Question Answering.* Online: Association for Computational Linguistics. 149-157. https://aclanthology.org/2021.mrqa-1.15/.

Sakaji, Hiroki, and Kiyoshi Izumi. 2023. "Financial Causality Extraction Based on Universal Dependencies and Clue Expressions." Springer. 839-857.

Scaramozzino, Roberta, Paola Cerchiello, and Tomaso Aste. 2021. "Information theoretic causality detection between financial and sentiment data." MDPI. 621.

Wolf, Thomas, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, et al. 2020. "Transformers: State-of-the-Art Natural Language Processing." Edited by Qun Liu and David Schlangen. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations.* Online: Association for Computational Linguistics. 38-45. doi:10.18653/v1/2020.emnlp-demos.6.

Yang, Jie, Soyeon, Caren Han, and Josiah Poon. 2021. "A Survey on Extraction of Causal Relations from Natural Language Text."

Zaremba, Anna, and Tomaso Aste. 2014. "Measures of causality in complex datasets with application to financial data." Molecular Diversity Preservation International (MDPI). 2309-2349.

APPENDIX-A
PSUEDOCODE

Generate Predictions Using QA Models

1. **Initialize QA Model**:
   - Load either **XLM-RoBERTa** (multilingual) or **BERT** (English) based on the use case.
2. **Define answer_question**:
   - Input: question, context.
   - Use the chosen model to predict the answer.
3. **Process Input CSV**:
   - Open input CSV and read rows.
   - Add "Answer" to the header.
   - For each valid row:
     - Extract id, context, and question.
     - Predict the answer and append it to results.
   - Skip invalid rows and log them.
4. **Write Output CSV**:
   - Save the results (header + answers) to an output file.
5. **Run**:
   - Call the process with specified input/output file paths.
   - Print completion message.

APPENDIX-B

SCREENSHOTS

English Example:

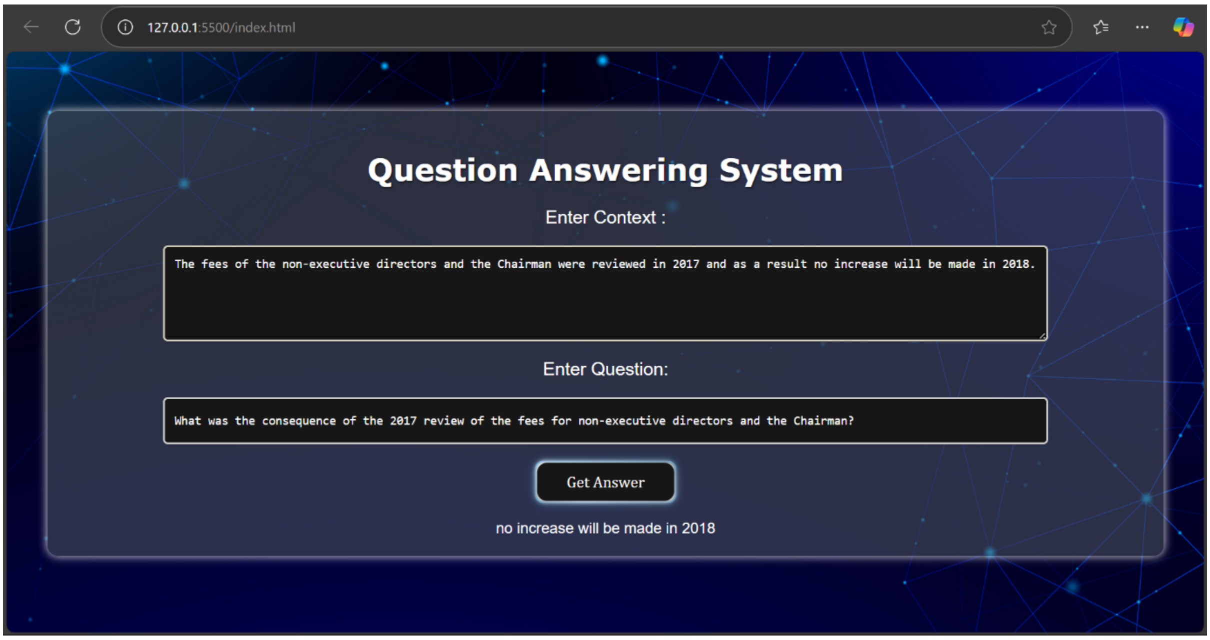


Figure 3 English Example

Spanish Example:

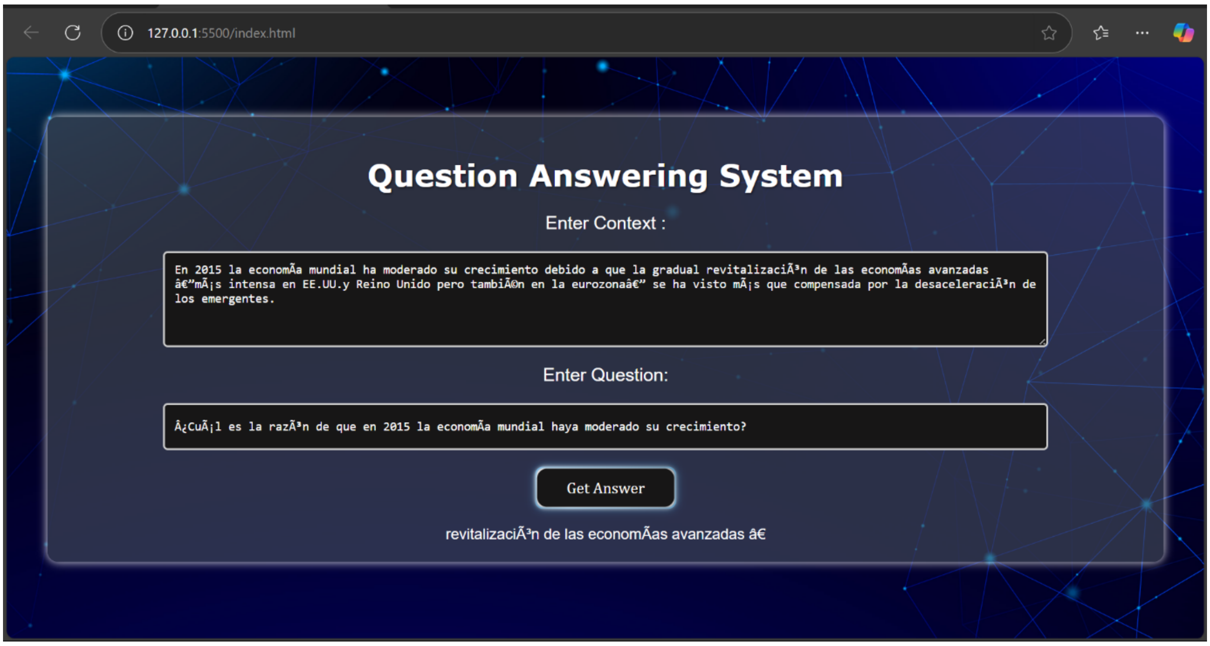


Figure 4 Spanish Example

APPENDIX-C

ENCLOSURES

1. Journal publication/Conference Paper Presented Certificates of all students.

Your FinNLP-FNP-LLMFinLegal 2025 Submission (Number 36)

FinCaus...pdf.pdf

Hide message history

From: Antonio Moreno-Sandoval <noreply@softconf.com>
Sent: Thursday, December 5, 2024 4:45:49 pm
To: Dr. Sandeep Albert Mathias -Asst. Selection Grade-SCSE <sandeepalbert@presidencyuniversity.in>
Cc: pmf@inf.uc3m.es <pmf@inf.uc3m.es>
Subject: Your FinNLP-FNP-LLMFinLegal 2025 Submission (Number 36)

Dear Sandeep Mathias:

On behalf of the FinNLP-FNP-LLMFinLegal 2025 Program Committee, I am delighted to inform you that the following submission has been accepted to appear at the conference:

PresiUniv at FinCausal 2025 Shared Task: Applying Fine-tuned Language Models to Explain Financial Cause and Effect with Zero-shot Learning

```
============================================================================
                              REVIEWER #1
============================================================================


----------------------------------------------------------------------------
Reviewer's Scores
----------------------------------------------------------------------------
                         Appropriateness (1-5): 5
                                 Clarity (1-5): 4
            Originality / Innovativeness (1-5): 3
                 Soundness / Correctness (1-5): 4
                   Meaningful Comparison (1-5): 4
                            Thoroughness (1-5): 3
              Impact of Ideas or Results (1-5): 3
                          Recommendation (1-5): 4
                     Reviewer Confidence (1-5): 5
```

```
================================================================================
                              REVIEWER #2
================================================================================

--------------------------------------------------------------------------------
Reviewer's Scores
--------------------------------------------------------------------------------
                   Appropriateness (1-5): 5
                           Clarity (1-5): 4
     Originality / Innovativeness (1-5): 3
          Soundness / Correctness (1-5): 5
            Meaningful Comparison (1-5): 4
                      Thoroughness (1-5): 3
       Impact of Ideas or Results (1-5): 3
                    Recommendation (1-5): 4
               Reviewer Confidence (1-5): 4
```

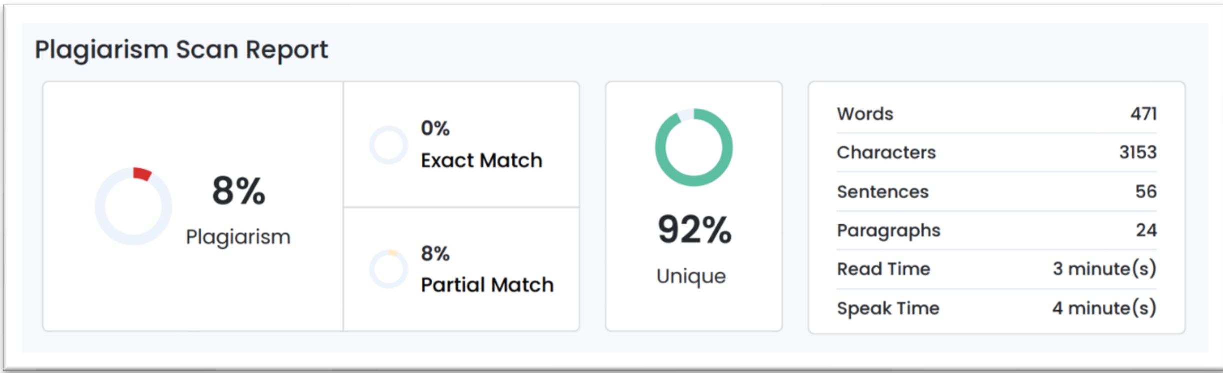
```
================================================================================
                              REVIEWER #3
================================================================================

--------------------------------------------------------------------------------
Reviewer's Scores
--------------------------------------------------------------------------------
                   Appropriateness (1-5): 5
                           Clarity (1-5): 5
     Originality / Innovativeness (1-5): 3
          Soundness / Correctness (1-5): 5
            Meaningful Comparison (1-5): 4
                      Thoroughness (1-5): 4
       Impact of Ideas or Results (1-5): 3
                    Recommendation (1-5): 4
               Reviewer Confidence (1-5): 5
```

2. Include certificate(s) of any Achievement/Award won in any project-related event.

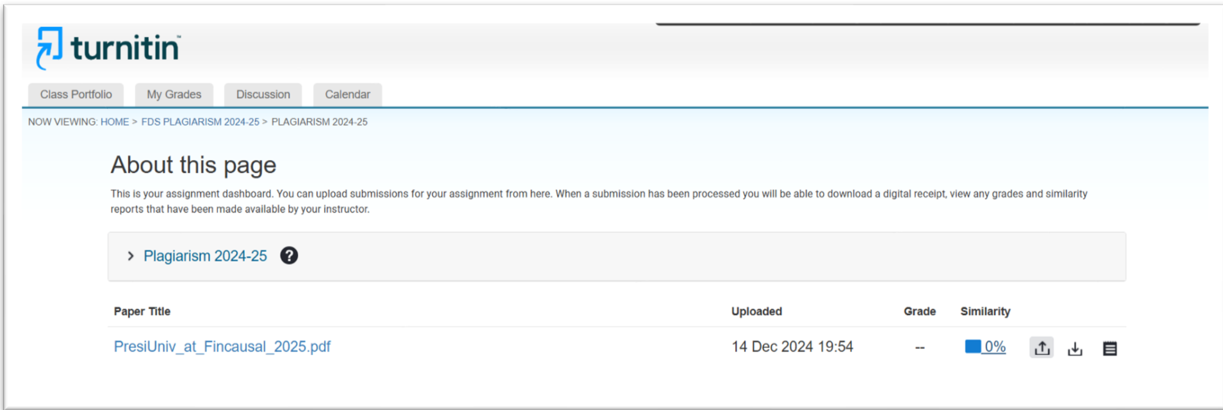
The Joint Workshop of the 9th Financial Technology and Natural Language Processing (FinNLP), the 6th Financial Narrative Processing (FNP), and the 1st Workshop on Large Language Models for Finance and Legal (LLMFinLegal) is scheduled on **19th and 20th January, 2025.**

3. Similarity Index / Plagiarism Check report clearly showing the Percentage (%). No need for a page-wise explanation.

Report:



Paper:



4. Details of mapping the project with the Sustainable Development Goals (SDGs).

Contribution to Sustainable Development Goal (SDG) 8: Decent Work and Economic Growth

The project aligns with SDG 8, "Decent Work and Economic Growth," by enhancing the analysis and comprehension of financial documents. By employing advanced techniques to detect causal relationships in financial data, the project provides valuable insights into the underlying factors driving economic changes. These insights enable businesses and financial institutions to identify trends, risks, and opportunities, leading to more strategic decision-making.

Improved decision-making fosters sustainable economic growth by promoting efficiency, reducing financial uncertainties, and encouraging investments in areas with the highest impact. Moreover, the project contributes to greater transparency in financial systems, which is crucial for building trust among stakeholders and ensuring ethical practices. By helping organizations understand the root causes of economic challenges and successes, the project supports the creation of stable and equitable work environments.

Ultimately, this fosters economic stability, promotes innovation, and improves job quality, thereby contributing to a resilient and inclusive economy that benefits both individuals and organizations.