Graph Neural Network and NER-Based Text Summarization

Imaad Zaffar Khan
NetId: izkhan2
izkhan@illinois.edu

Amaan Aijaz Sheikh
NetId: asheikh4
asheikh4@illinois.edu

Utkarsh Sinha
NetId: usinha3
usinha3@illinois.edu

Abstract—**With the abundance of data and information in today's time, it is nearly impossible for man, or, even machine, to go through all of the data line by line. What one usually does is to try to skim through the lines and retain the absolutely important information, that in a more formal term is called summarization. Text summarization is an important task that aims to compress lengthy documents or articles into shorter, coherent representations while preserving the core information and meaning. This project introduces an innovative approach to text summarization, leveraging the capabilities of Graph Neural Networks (GNNs) and Named Entity Recognition (NER) systems. GNNs, with their exceptional ability to capture and process the relational data inherent in textual information, are adept at understanding the complex structures within large documents. Meanwhile, NER systems contribute by identifying and emphasizing key entities, ensuring that the summarization process maintains a focus on the most critical aspects of the text. By integrating these two technologies, our method aims to enhances the efficiency of summarization and also tries to ensures a high degree relevance in the condensed content. This project, therefore, offers a promising direction for handling the ever-increasing volume of textual data in an information-saturated world.**

Index Terms—**Graph Neural Networks, Named Entity Recognition, Text Summarization**

I. Introduction

A. Background

1) Introduction to Text Summarization: Text summarization, in its essence, is the process of extracting the most important information from a text to produce a concise version that conveys the core meaning while reducing size of the text. This technique is divided into two primary categories: extractive summarization and abstractive summarization. Extractive summarization focuses on selecting and compiling key segments directly from the source text, whereas abstractive summarization involves generating new text that conveys the main ideas in a condensed form. Our focus lies on the former, exploring the nuances of extractive text summarization and its advancements.

2) Relevance in Today's World: In the current digital era, we are overwhelmed with textual information. From online news articles and scientific journals to social media feeds and corporate reports, the sheer volume of text available poses a significant challenge for individuals seeking to grasp information efficiently. Text summarization emerges as a potential and important solution in this context, enabling quick and effective access to essential content. Its applications span various domains where timely and accurate information extraction is quite important.

3) Transition to GNN and NER in Text Summarization: While there have been various advancements in the field of text summarization, the incorporation of Graph Neural Networks (GNN) and Named Entity Recognition (NER) stands out. These technologies address the unique challenges posed by the volume and complexity of modern data. GNN, with its capacity to model relational data, and NER, effective in identifying and classifying named entities in text, together improve the extractive summarization process. This integration can not only improves the accuracy of summarization but also ensure relevance and context-awareness, a critical aspect when dealing with diverse and extensive datasets.

B. Problem Statement

In the current times, we are confronted with an ever-growing size of textual data of great variety. This information overload not only challenges our capacity for comprehension but also strains traditional text summarization methods. These conventional approaches often fall short in addressing the complexity and scale of the data, leading to inefficiencies and missed insights.

Amidst this, Large Language Models (LLMs) have emerged as a powerful tool in text processing, showcasing an impressive ability to understand and generate human-like text. However, their implementation has the downside of significant computational costs and resource demands, making them less feasible for continuous use in various resource-constrained environments. This situation presents a critical need for more accessible and efficient text summarization solutions.

In this context, the integration of Graph Neural Networks (GNN) and Named Entity Recognition (NER) holds immense potential for revolutionizing extractive text summarization. GNNs excel in modeling complex data relationships, while NER systems are proficient in identifying and categorizing key entities within texts. Individually, they have shown promise in various text processing applications, but their combined use in extractive text summarization remains an uncharted territory.

This convergence presents a unique opportunity to address the limitations of current summarization methods, especially in managing large volumes of data efficiently and effectively. The

core challenge lies in investigating the feasibility and effectiveness of combining GNN and NER in a unified framework. Such an integration aims not only to enhance the capacity for handling large-scale data but also to provide a resource-efficient alternative to the more demanding LLMs.

C. Objectives

The primary objectives of this project are outlined as follows:

1) **Explore the Integration of GNN and NER for Extractive Text Summarization:** Implement the integration and test if in reality, the theoretical idea of the integration is practically effective once poised with the task of summariaztion.
2) **Create a Resource-Efficient Summarization Tool:** Design a text summarization system that operates with significantly lower computational resources compared to Large Language Models (LLMs), making advanced text summarization more accessible in various environments.
3) **Evaluate the Performance of the GNN-NER Framework:** Conduct thorough testing and evaluation of the proposed GNN-NER framework across diverse real-world scenarios, assessing accuracy and efficiency.
4) **Enhance Context-Awareness in Summarization:** Focus on improving the context-awareness in the summarization process, ensuring the summarized content retains the essential context and meaning of the original text with the help of proposed NER integration.

II. RELATED WORK

The field of extractive text summarization has seen significant advancements in recent years, with researchers exploring various methodologies to enhance the efficiency and accuracy of summarization processes.

1) **Graph-Based Approaches**:
   [1]Kadriua and Obradovica (2021) explored an extractive text summarization method using graphs, which offers insights into the structural analysis of texts for summarization purposes. Further advancements in graph-based methods can be seen in other works, where the focus is on integrating graph structures with neural network models to capture contextual relationships in text.
2) **Named Entity Recognition in Summarization**:
   [2]Alshibly et al. (2023) investigated text summarization of news articles using Named Entity Recognition (NER) via the SpaCy library, underscoring the importance of identifying key entities in text for effective summarization.
   This is complemented by studies which delve into the integration of NER with machine learning algorithms to enhance the precision of extractive summarization.
3) **Language-Independent Summarization Techniques**:
   [3]Hernández-Castañeda et al. (2022) discussed a language-independent approach to extractive automatic text summarization, focusing on automatic keyword extraction, which is crucial for understanding multi-lingual and cross-lingual summarization techniques.
4) **Algorithmic Innovations in Summarization**:
   [4]P.Verma, A.Verma, and Pal (2022) presented an approach using fuzzy evolutionary and clustering algorithms for extractive text summarization, highlighting the role of advanced computational methods in this domain. The evolution of algorithmic approaches can further be seen in various other studies, which explore the use of deep learning techniques for enhancing the adaptability and accuracy of summarization tools.

Each of these studies contributes to a refined understanding of extractive text summarization. However, there remains a gap in exploring the combined potential of GNN and NER for this purpose. While individual studies have shown the efficacy of GNN in understanding complex relational data and NER in accurately identifying key entities, their integration in the context of extractive text summarization remains under-explored. This gap presents an opportunity for novel research, potentially leading to efficient summarization tools capable of handling the vast and varied nature of today's textual data.

III. METHODOLOGY

This section outlines the methodology adopted for integrating Graph Neural Networks (GNN) and Named Entity Recognition (NER) in extractive text summarization. The methodology is structured into several key phases:

A. Proposed Architecture

The proposed architecture for our approach can be seen below:
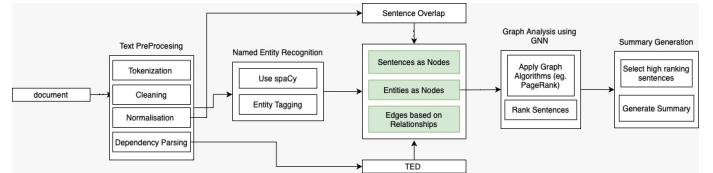

Figure 1. Architecture of the GNN and NER based text summarization system

The architecture diagram (Figure 1) illustrates the systematic flow of the proposed integration. The process is initiated with a document that undergoes a series of text pre-processing steps. These steps include tokenization, where the text is split into individual tokens, cleaning, to remove any irrelevant characters or spaces, normalization, which standardizes the text, and dependency parsing to analyze the grammatical structure of the text.

Following pre-processing, the text is passed through the NER phase, where spaCy is used for entity tagging, identifying named entities within the text. This phase is critical for recognizing important nouns and noun phrases that are essential to understanding the context of the document.

The extracted sentences and entities are then constructed into a graph, with sentences as nodes, entities also as nodes,

and edges based on the relationships derived from the text and entity analysis. This graph is analyzed using GNN, specifically employing algorithms to evaluate the importance of each node within the graph structure.

In the final phase of Summary Generation, sentences are selected based on their ranking from the graph analysis. High-ranking sentences are deemed to carry significant informational weight and are chosen to be a part of the final summary. The summary is generated by strategically combining these sentences to reflect the main ideas of the original document while maintaining coherence and conciseness.

This architecture is designed to optimize the summarization process by leveraging the strengths of both GNN for capturing relational information and NER for emphasizing key entities, resulting in summaries that are both informative and representative of the source document.

B. *Data Collection and Preprocessing*

The project utilizes the CNN Daily Mail dataset, renowned for its diverse compilation of news stories, ideal for tasks in natural language processing, including summarization. The selection criteria for textual data included relevance to current events, diversity in topics, and rich entity representation to facilitate effective NER.

- **Data Collection:** The dataset is accessed through established repositories, ensuring that the data is reliable and can serve its purpose. A brief table on the statistics of the CNN dataset can be seen below:

Table I
OVERVIEW OF THE CNNDAILYMAIL DATASET

Statistic	Value
Number of Articles	~93,000(CNN)+~220,000(Daily Mail)
Number of Summaries	Same as number of articles
Average Length of Articles	~800 (CNN), ~1000 (Daily Mail)
Average Length of Summaries	~55 (bullet points)
Vocabulary Size	~350,000 to 400,000 unique tokens
Min Length of Articles	Generally around 100
Max Length of Articles	Varies, some exceed 2,000 words
Min Length of Summaries	As low as 20
Max Length of Summaries	Can exceed 100 words

The CNN Daily Mail dataset, can be seen from the statistics above, is a good choice for our project on extractive text summarization due to its substantial volume and diversity of length of articles as well as the diversity in the length of summaries provided. With over 300,000 combined articles from CNN and Daily Mail, the dataset offers a rich resource in terms of diversity and volume of data. The varying lengths of articles and summaries ensure that our system can adapt to a range of contexts and detail levels, which is an essential aspect for various use-cases. Moreover, the extensive vocabulary present in the dataset is advantageous for developing a robust model capable of understanding and summarizing a wide array of topics. This variability and complexity make the CNN Daily Mail dataset an ideal candidate to serve its purpose for our research.

- **Pre-Processing:**
Data pre-processing forms the foundation of our text summarization model, as the subsequent graph representation relies heavily on the quality and richness of the input. The process comprises several stages, to refine the data for the construction of text-rich nodes and their edges/relationships within the graph. The steps are as follows:

  - **Tokenization:** The text is broken down into tokens. This step is crucial as it determines the granularity of the nodes in our graph. Each token potentially represents a node implying it to require careful segmentation.
  - **Cleaning (Punctuation/Special Characters):** We remove punctuation and special characters that do not contribute to the semantic relationships in the graph. This step helps in reducing noise and focusing on the text that contributes to the meaning of the content.
  - **Lemmatization:** Lemmatization is applied to reduce words to their base or dictionary form. We aim to capture the essence of the text in the graph, this step helps in generalizing the connections between nodes by reducing the words to their base form.
  - **Lower Casing:** Converting all text to lower case ensures that the same words are not treated as different nodes due to case differences.
  - **Stop-Word Removal:** Common stop-words are filtered out as they usually do not carry significant meaning and are not useful as nodes in the graph. Their removal increases the text-richness of the remaining nodes, which bear more semantic weight.
  - **Dependency Parsing:** Perhaps the most critical step, dependency parsing, is employed to understand the grammatical structure of the text, which informs the edges in our graph. It elucidates the syntactic relationships between tokens, allowing for a richer and more meaningful representation of the text as a graph.

Each of these pre-processing steps contributes to the construction of a graph where the nodes (words or entities) and their relationships (edges) are indicative of the content's semantic structure. Given that our approach heavily relies on the accurate representation of textual relationships through graphs, the pre-processing stage is not merely a prerequisite but a determinant of the summarization quality. The more text-rich and semantically accurate our graph, the better our model can identify and extract the most important/appropriate information for summarization. This attention to pre-processing detail ensures that our GNN can operate on data that closely mimics the interconnected nature of human language and thought.

C. Implementation of Named Entity Recognition (NER)

Named Entity Recognition (NER) is a critical stage in our text summarization process, going beyond basic pre-processing by enriching the text data with semantic information. This step acts as an additional input generation step. Output from the NER is given into the Graph Construction part of the pipeline. We utilize the spaCy library for its state-of-the-art NER capabilities.
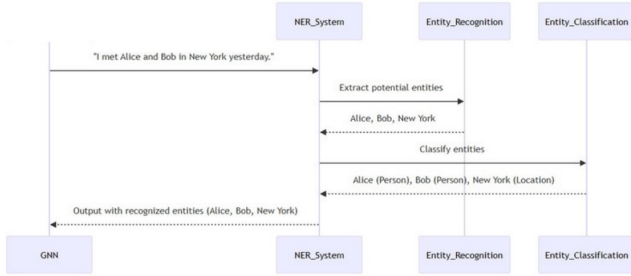
An overview of this part of our pipeline can be seen below:

Figure 2. Architecture of NER Layer in Proposed Approach

Utilizing Figure 2, let us further look at the importance this layer:

– **NER Model Selection:** We made use of spaCy's advanced pre-trained NER model, which identifies a wide range of entities with high precision. The model's robustness is derived from its training on a diverse and extensive corpus, enabling it to recognize and tag entities accurately across various domains [2]. Our selection of spaCy's model is based on its *exceptional performance in benchmarks* and its ease of integration into our proposed pipeline.

– **Entity Identification:** The NER model tags text with entity types such as persons, organizations, locations, and others. This step is essential, as entities often carry significant weight in the meaning of a sentence and play a crucial role in summarization. By identifying these entities, we are able to preserve critical information that may otherwise be lost in a purely extractive approach.

– **Graph Construction with Entities:** In the subsequent stage of our pipeline, the recognized entities become nodes within our graph. This is not a simple transformation of text into graph form; rather, it is a careful transformation of the text into a structured representation that highlights the most meaningful components. The entities serve as pivotal points in the graph, often becoming central nodes that connect with various parts of the text, reflecting their importance in the narrative structure.

– **Semantic Enrichment:** The classification of entities facilitates a deeper understanding of the text by the GNN. It allows the GNN to not just process the text as a sequence of words but as a semantically rich network where the relationships between entities and other textual elements can be analyzed in a way that is closer to human cognition.

The implementation of NER is a decisive step that, in theory, should significantly influence the performance of our GNN model. By accurately identifying and classifying entities, we create a detailed semantic map of the text, which is essential for generating summaries that are both informative and reflective of the text's core themes.

D. Incorporation of Graph Neural Networks (GNN)

Graph Neural Networks (GNNs) are leveraged to capture the complex relational information among entities and contextual words in the text, effectively modeling the data as a graph. The process is visualized in Figure 3, which illustrates the transition from isolated nodes to a fully connected graph.
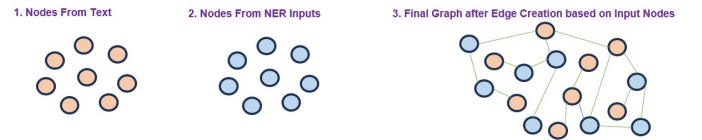

Figure 3. Progression of graph construction in GNN

Based on Figure 3, lets expand more below:

– **GNN Model Selection:** Our selected GNN architecture is specifically designed to process text-based graph structures. It is designed to understand and make use of the relational dynamics between words and entities, which are essential for identifying crucial information in the text summarization process. The architecture is chosen for its ability to handle the heterogeneity of the graph's nodes and the complexity of their interrelations.

– **Graph Construction:** The graph is constructed in a two-step process. Initially, sentences and identified entities are treated as separate nodes (as shown in parts 1 and 2 of Figure 3). These nodes are then interconnected through edges that represent their relationships, which are determined by co-occurrence within sentences, syntactical dependencies, and semantic relatedness (part 3 of Figure 3). This results in a rich graph where the nodes are text elements and the edges represent the contextual and relational nuances of the original document.

– **Edge Formation:** Edges are not uniformly created; instead, their formation is informed by the depth of the relationship between nodes. Factors such as the frequency of co-occurrence and the strength of dependency links play a role in defining the edge weights, which are crucial for the GNN as it processes the graph to identify which nodes (text segments) are pivotal for the summary.

The incorporation of GNNs enables the system to move beyond linear text analysis, instead treating the text as

a complex network, somewhat similar to human cognitive processes that interpret information. This networked approach allows for a more detailed understanding and generation of summaries, which are both comprehensive and contextually coherent.

E. Graph Analysis Using GNN

After graph construction, the next step is the graph analysis. This phase involves the application of graph-based algorithms to determine the centrality and importance of each node (sentence and entity) within the graph.

- **Algorithm Selection:** For graph analysis, algorithms such as PageRank or other centrality measures are employed (discussed in detail in the following subsection). These algorithms evaluate the importance of a node based on its connections, with the understanding that nodes with more connections or nodes connected to other highly connected nodes signify greater importance in the text.
- **Node Ranking:** Each node is scored and ranked according to the chosen algorithm. This ranking reflects the node's significance in the overall graph, which correlates to the importance of the corresponding sentence or entity in the context of the summary.

The outcome of this analysis is a list of sentences and entities ranked by their importance and relevance to the core content of the document, setting the stage for the final summarization.

F. Graph Analysis Algorithms

Graph analysis algorithms play an important role in interpreting the constructed graph and determining the importance of nodes for the summarization task. We have utilized and compared a variety of algorithms, each contributing a unique interpretation on node importance and network structure.

- **PageRank:** PageRank is used to measure the importance of nodes in the graph. Nodes that are connected to many nodes, or to nodes that are themselves important, receive a higher score.
- **HITS (Hyperlink-Induced Topic Search):** HITS, focuses on identifying 'hubs' and 'authorities' within the graph [2]. A good hub represents a node that points to many authorities.A good authority is a node that is pointed at by many hubs.
- **Closeness Centrality:** This metric identifies nodes that are able to spread information efficiently through the graph. Nodes with a high closeness centrality can more effectively interact with all other nodes in the network.
- **Degree Centrality:** This is a measure of the number of connections a node has. In the context of summarization, nodes with high degree centrality have a greater likelihood to contain words or phrases common across the document, hence, pointing towards potential keywords or topics.

- **Betweenness Centrality:** This algorithm identifies nodes that serve as link between different parts of the graph. Nodes with high betweenness centrality can indicate sentences that relate various key concepts in the text.
- **Clusters:** Identifying clusters within the graph helps in understanding the modular structure of the text. Clusters can represent themes or topics within the document, aiding in the generation of a summary that covers a diverse range of points.

By utilizing and evaluating all the above mentioned algorithms, our goal is to try to select the algorithm that provided the best possible results. Making use of these algorithms gives us the opportunity to in-depth compare our work, against previous works like [1], [2] and [3], and actually visually verify if our approach of integrating the NER to our graph is fruitful or not.

G. Summary Generation

The final phase of our methodology is the generation of the summary. This phase capitalizes on the insights provided by the graph analysis algorithms, mentioned in the previous sub-section, to determine which parts of the text are most informative and are to be included in the summary.

- **Sentence Selection:** Utilizing the rankings from graph analysis, sentences that are central to the text's meaning are selected. This selection ensures that the most relevant and important parts of the original document are preserved.
- **Redundancy Elimination:** To avoid repetitiveness, the algorithm identifies and removes redundant information. This ensures that the summary is not only concise but also non-repetitive.

The output is a summary that represents the key points and themes of the original text while being short and clear. This process mimics the human ability to identify and interpret the most useful information from a large text, providing users with a concise and informative version of the content.

H. Evaluation and Testing

Following the generation of summaries,evaluation and testing are conducted to ensure the effectiveness of the summarization pipeline.

- **Evaluation Metrics:** We utilize standard evaluation metrics such as ROUGE to quantify the quality of the summaries by comparing them to reference summaries. These metrics assess the overlap of n-grams, word sequences, and word pairs between the generated and reference summaries.
- **User Studies:** Beyond automated metrics, user studies were conducted to gather qualitative feedback on the summaries' readability, informativeness, and overall utility.

– **Iterative Refinement:** Based on the evaluations, the summarization system undergoes iterative refinement to address any identified shortcomings, ensuring continuous improvement in performance.

In the next section, we look at the various challenges and pitfalls of evaluation techniques in the field of text summarization, followed by the overall results obtained by our proposed integration of NER in GNN.

IV. RESULTS

A. *Evaluation Challenges in Text Summarization*

Evaluation of text summarization systems presents unique challenges, mainly due to the subjective nature of what constitutes a *good summary*. The lack of universally accepted evaluation metrics further complicates the assessment process. Traditional metrics tend to focus on surface-level text features which may not fully capture the semantic quality of a generated summary (an example of traditional metrics is word overlap).

– **Semantic Richness:** Summaries need to be evaluated not just by the words they contain, but also by the semantic richness and the delivery of the main important ideas from the inputted text.
– **Diversity of Expression:** The same concept can be expressed in numerous ways, making word-to-word comparison not a sufficient measure for evaluation. A good summary might use different wording than the source text or reference summaries, leading to low scores from automatic metrics.

These challenges highlight the need for more sophisticated evaluation approaches that can appreciate the nuanced task of summarization beyond mere lexical matching.

B. *Limitations of Automatic Evaluation Metrics*

Automatic evaluation metrics are often employed for their efficiency and objectivity. However, they inherently possess limitations, particularly when assessing the quality of text summaries. A more detailed view of the same can be seen below:

– **Word-to-Word Comparison:** Metrics like precision and recall, while useful, may not fully recognize the quality of a summary if they only rely on direct word matches. Summaries containing synonymous expressions or paraphrases might be undervalued.
– **Potential for Low Scores:** Given the diversity of valid summaries for any given text, strict word-to-word comparison can result in lower scores for summaries that are, in reality, accurate and serve the required purpose, simply because they do not use the exact wording of the reference summaries.

Such limitations indicate the need for a balanced approach to evaluation, combining both automatic metrics and human judgment to capture the effectiveness of summarization systems more holistically.

C. *ROUGE*

The *Recall-Oriented Understudy for Gisting Evaluation* (ROUGE) metric is widely used for evaluating automatic summarization and machine translation [3]. We have a look below at the ROUGE metrics that we use:

– **ROUGE-N:** This variant of ROUGE measures the overlap of n-grams between the generated summary and reference summaries. It is a valuable tool for assessing the presence of key phrases.

Despite its widespread adoption, ROUGE has its own limitations. Therefore, while ROUGE scores can provide valuable insights, they should be interpreted in the context of these known constraints and supplemented with qualitative analyses to ensure a comprehensive evaluation of summarization performance.

In this section, we present the results of our text summarization system evaluated using the ROUGE metric, which is standard in the field of automatic text summarization. The results are presented in terms of both F1 scores, which combine precision and recall, and recall values alone, offering insight into the system's ability to capture the content of the reference summaries.

D. *Findings*

The following table presents the F1 scores for our approach across different ROUGE metrics.

Table II
ROUGE F1 SCORES FOR PROPOSED APPROACH

Graph Algorithm	ROUGE 1	ROUGE 2
PageRank	0.29	0.11
HITS	0.30	0.11
Closeness	**0.30**	**0.12**
Betweenness	0.29	0.10
Degree	0.30	0.11
Clusters	0.28	0.09

These scores reflect the system's effectiveness in producing summaries that are similar to the reference summaries, with a higher score indicating a more accurate summarization.

While F1 scores provide a harmonic mean of precision and recall, focusing on recall alone allows us to understand how well the system captures the entirety of the reference content.

Table III
ROUGE RECALL SCORES FOR PROPOSED APPROACH

Graph Algorithm	ROUGE 1	ROUGE 2
PageRank	0.49	0.20
HITS	0.42	0.17
Closeness	0.46	0.18
Betweenness	0.47	0.19
Degree	0.48	0.20
Clusters	0.42	0.15

The recall scores are especially important in scenarios where the inclusion of all relevant information is critical,

such as in comprehensive document summaries where missing details could lead to misinterpretation of the content.

E. *Discussion*

The obtained ROUGE scores suggest that our summarization system performs well in capturing the gist of the source documents. Again, it is to be noted that this statement is made on the grounds of the limitations of the overall evaluation metrics in the field of text summarization. Specifically, the F1 scores indicate a strong balance between precision and recall, while the recall scores suggest that the approach is quite effective at ensuring that the summaries are comprehensive. These findings are promising, showing that the novel approach can reliably generate summaries.

V. COMPARATIVE ANALYSIS

A. *Comparison Against Previous Works*

A critical part of evaluating the effectiveness of our proposed summarization system is to compare its performance with that of existing approaches. We have selected several notable studies for comparison, with their results reported using the same ROUGE metrics on the same CNN/Dailymail Datset for consistency.

Below is a table that compares the best ROUGE F1 scores of our system with those of previous works of [1] and [2]. It is important to note that differences in test sets, pre-processing, and other experimental settings can affect comparability and hence provide an approxiamte comparison at the best.

Table IV
COMPARISON OF ROUGE F1 SCORES WITH PREVIOUS WORKS

System	ROUGE-1	ROUGE-2
Previous Work [1]	0.28	0.11
Previous Work [2]	0.21	-
Our Approach	**0.30**	**0.12**

As illustrated in Table IV, our integrated approach demonstrates a better performance in comparison to the established benchmarks. This suggests that *the integration of GNN and NER can indeed contribute to the improvement of extractive text summarization.*

Similarly, we compare the recall scores to assess the comprehensiveness of the summaries generated.

Table V
COMPARISON OF ROUGE RECALL SCORES WITH PREVIOUS WORKS

System	ROUGE-1	ROUGE-2
Previous Work [1]	0.48	0.17
Our System	0.46	0.18

The recall scores shown in Table V highlight the system's ability to include relevant content in the summaries. Our

system's performance in this area is indicative of its potential for practical applications where content coverage is crucial.

B. *Comparison Against LLM's*

Our approach, which integrates Graph Neural Networks (GNN) and Named Entity Recognition (NER), offers a different approach to that of LLM's, which we believe can be more cost effective. Here, we compare its performance against these LLMs to highlight the advantages and potential trade-offs.

The following table compares the summarization performance of our system against that of selected LLMs [5]. The comparison is based on the ROUGE metrics, considering both the F1 scores and recall.

Table VI
PERFORMANCE COMPARISON WITH LLMS

Model	ROUGE-1 F1	ROUGE-2 F1
BERTSUM+Transformer	0.43	0.20
Our Approach	0.30	0.12

This comparison indicates where our GNN-NER system stands in comparison to the current LLMs. Particularly, it showcases the system's efficiency in generating summaries without the extensive computational resources typically required by LLMs.

One of the key advantages of our approach over LLMs is resource efficiency. LLMs, while powerful, require significant computational power, which can be a limiting factor in certain applications.

C. *Case Study for effectiveness of Novel Approach*

This case study aims to demonstrate the practical outcomes of different summarization methods. We compare the results of using Graph Neural Networks (GNNs) alone and in combination with a Named Entity Recognition (NER) based approach, and a Large Language Model (LLM-GPT) to provide a view on the summaries generated.

1) Method Description: The Various methods used:

- **Integrated GNN-NER Approach:** Our proposed approach following the pipeline showcased in the above sections.
- **GNN-only Approach:** The process flow is similar to that of our proposed approach, apart from the additional input from NER block.
- **LLM-based Approach:** Using GPT to generate an extractive summary.

2) Input Text: "Cristiano Ronaldo and Lionel Messi will go head-to-head once more in the race to be this season's top scorer in the Champions League —xe2—x80—x93 although Luiz Adriano threatens to spoil the party. Both Barcelona and Real Madrid booked their spots in the semi-finals this week with victories over Paris Saint-Germain and Atletico Madrid respectively. The planet's

best footballers have scored eight times in Europe this season. But Shakhtar Donetsk—xe2—x80—x99s Adriano, courted by Arsenal and Liverpool, has netted on nine occasions this term. Cristiano Ronaldo, in action against Atletico Madrid on Wednesday evening, has scored eight goals in Europe. Lionel Messi also has eight goals in the Champions League this term; one fewer than Luiz Adriano. Ronaldo and Messi will both play at least two more times after Real Madrid and Barcelona reached the last four. Adriano, who moved to Donetsk in 2007, scored five against BATE Borsiov in the group stages. His performance that night made history, with the 27-year-old becoming only the second player to score five times in a Champions League game. The other was Messi for Barcelona against Bayer Leverkusen in 2012. He also scored the third quickest hat-trick in the competition's history (12 minutes) as the Ukrainian side, knocked out by Bayern Munich in the round of 16, racked up the biggest-ever half-time lead (6-0) in Europe's premier tournament. —xe2—x80—x98I am in a good moment of my career and we'll do what will be best for me and for the club,—xe2—x80—x99 said Adriano last month when quizzed over his future. Adriano, who netted five times against BATE Borisov in the group, has scored more goals than any other player in the Champions League... he is out of contract in December and could move to the Premier League. —xe2—x80—x98With my contract set to expire and many good performances, it'll be difficult to stay in Ukraine. —xe2—x80—x99 Arsenal have sent scouts to watch Adriano in recent months, while Liverpool are also keen on the Brazilian. His contract with Shakhtar Donetsk runs out at the end of the year. Ronaldo and Messi however, remain in pole-position to top the scoring charts with Barcelona and Real Madrid both in the hat for the two-legged semi-finals to be played next month. Of the teams still in the pot, Neymar and Luis Suarez of Barcelona, Real Madrid's Karim Benzema and former Manchester United and City striker Carlos Tevez, now plying his trade for Juventus, each have six goals. The draw for the last four will take place on Friday."

3) Summarization Results: The summaries obtained from each method can be seen in the figures below:

Figure 4. GNN+ NER Approach Generated Summary

Figure 5. GNN Approach Generated Summary

Figure 6. LLM(GPT) Approach Generated Summary

D. Analysis

This case study presents a comparative analysis of summaries generated using three different approaches. The comparison is based on the ability of the approaches for content coverage, coherence, conciseness, and focus on the main themes.

Table VII
COMPARATIVE ANALYSIS OF SUMMARIZATION APPROACHES

Aspect	GNN+NER	GNN Only	GPT
Content Coverage	Covers relevant details	Focused on Adriano, No Clubs Mentioned.	Good Coverage.
Coherence	Moderate Coherency	Less coherent	Good Coherency
Conciseness	Moderately concise	Less efficient organization of information.	Balanced
Focus on Main Themes	Multiple themes covered	Misses broader context	Successfully captures the central themes

The table reveals that while GNN-based methods are effective in extracting key information, they may not always provide a coherent narrative. The GPT approach, in contrast, offers a more narrative and contextually rich summary, highlighting its strength in generating coherent and comprehensive summaries. However, based on human analysis, the ultimate question still lies on whether the computational expense is still worth it as the GNN+NER based approach generates a great quality summary nonetheless.

VI. CONCLUSION

A. Summary

In this study, we presented a novel approach to extractive text summarization by integrating Graph Neural Networks (GNN) with Named Entity Recognition (NER). Our methodology leveraged the GNN's ability to understand complex relationships within text and NER's precision in identifying relevant entities. Through a series of data pre-processing steps and the use of graph analysis algorithms, we constructed a model that understands and summarizes large volumes of textual data with a degree of accuracy and semantic coherence. The approaches's performance, evaluated using the standard ROUGE metric, showcases its promising potential in generating concise summaries that retain the essential information of the original documents while its comparison with LLM's further validates its immense potential .

B. Future Work

The field of text summarization, particularly when combined with GNN and NER, is an uncharted territory with opportunities for further research. Future work may focus on several areas, including:

- **Model Enhancement:** Refining the GNN architecture and exploring additional NER systems to improve the accuracy and richness of the summarization.
- **Evaluation Metrics:** Developing more sophisticated evaluation metrics that can capture the quality of summaries beyond word overlap, such as semantic fidelity and coherence.
- **Scalability:** Enhancing the model's scalability to handle datasets larger than CNN/Daily Mail, including real-time news feeds and extensive academic corpora.

REFERENCES

[1] K. Kadriua and M. Obradovica, "Extractive Approach For Text Summarization Using Graphs," *Arxiv*, vol. 2106.10955v1, June 2021. [Online]. Available: https://arxiv.org/abs/2106.10955

[2] I. Alshibly, S. Al-Shorfat, M. Otair, M. Shehab, O. Tarawneh, and M. Sh. Daoud, "Text Summarization Of News Articles Based On Named Entity Recognition Using Spacy Library," *Research Square*, March 2023.

[3] A. Hernández-Castañeda, R. A. García-Hernández, Y. Ledeneva, and C. E. Millán-Hernández, "Language-independent Extractive Automatic Text Summarization Based On Automatic Keyword Extraction," *Computer Speech and Language*, vol. 71, January 2022, 101267.

[4] P. Verma, A. Verma, and S. Pal, "An Approach For Extractive Text Summarization Using Fuzzy Evolutionary And Clustering Algorithms," *Applied Soft Computing*, vol. 120, May 2022, 108670.

[5] Yang Liu, "Fine-tune BERT for Extractive Summarization," *Arxiv*, vol. 2106.10955v1, September 2021. [Online]. Available: https://arxiv.org/abs/1903.10318

[6] A. Oussous, F.-Z. Benjelloun, A. A. Lahcen, S. Belfkih, "Big data technologies: A survey," *Journal of King Saud University-Computer and Information Sciences*, vol. 30, no. 4, pp. 431–448, 2018.

[7] R. Paulus, C. Xiong, R. Socher, "A deep reinforced model for abstractive summarization," *arXiv preprint arXiv:1705.04304*.

[8] H. Jin, T. Wang, X. Wan, "Multi-granularity interaction network for extractive and abstractive multi-document summarization," in *Proceedings of the 58th annual meeting of the association for computational linguistics*, 2020, pp. 6244–6254.

[9] R. K. Singh, S. Khetarpaul, R. Gorantla, S. G. Allada, "Sheg: summarization and headline generation of news articles using deep learning," *Neural Computing and Applications*, vol. 33, pp. 3251–3265, 2021.

[10] G. Erkan, D. R. Radev, "Lexrank: Graph-based lexical centrality as salience in text summarization," *Journal of Artificial Intelligence Research*, vol. 22, pp. 457–479, 2004.

[11] B. Manly, L. McDonald, D. L. Thomas, T. L. McDonald, W. P. Erickson, *Resource selection by animals: statistical design and analysis for field studies*, Springer Science and Business Media, 2007.

[12] R. Nallapati, F. Zhai, B. Zhou, "Summarunner: A recurrent neural network based sequence model for extractive summarization of documents," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, 2017.

[13] R. Socher, J. Pennington, E. H. Huang, A. Y. Ng, C. D. Manning, "Semi-supervised recursive autoencoders for predicting sentiment distributions," in *Proceedings of the 2011 conference on empirical methods in natural language processing*, 2011, pp. 151–161.

[14] X.-C. Yin, W.-Y. Pei, J. Zhang, H.-W. Hao, "Multi-orientation scene text detection with adaptive clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1930–1937, 2015.

[15] A. Fan, D. Grangier, M. Auli, "Controllable abstractive summarization," *arXiv preprint arXiv:1711.05217*.

[16] R. Nallapati, B. Zhou, C. Gulcehre, B. Xiang, et al., "Abstractive text summarization using sequence-to-sequence rnns and beyond," *arXiv preprint arXiv:1602.06023*.

[17] A. M. Rush, S. Chopra, J. Weston, "A neural attention model for abstractive sentence summarization," *arXiv preprint arXiv:1509.00685*.

[18] J. Gerlings, A. Shollo, I. Constantiou, "Reviewing the need for explainable artificial intelligence (xai)," *arXiv preprint arXiv:2012.01007*.

[19] S. R. Senthamizh, K. Arutchelvan, "Automatic text summarization using document clustering named entity recognition," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 9.

[20] B. Bhardwaj, S. I. Ahmed, J. Jaiharie, R. S. Dadhich, M. Ganesan, "Web scraping using summarization and named entity recognition (ner)," in *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, IEEE, 2021, pp. 261–265.

[21] S. Naseer, M. M. Ghafoor, S. bin Khalid Alvi, A. Kiran, S. U. Rahmand, G. Murtazae, G. Murtaza, "Named entity recognition (ner) in nlp techniques, tools accuracy and performance.," *Pakistan Journal of Multidisciplinary Research*, vol. 2, no. 2, pp. 293–308, 2021.

[22] B. Song, F. Li, Y. Liu, X. Zeng, "Deep learning methods for biomedical named entity recognition: a survey and qualitative comparison," *Briefings in Bioinformatics*, vol. 22, no. 6, bbab282, 2021.

[23] S. Huynh, K. Le, N. Dang, B. Le, D. Huynh, B. T. Nguyen, T. T. Nguyen, N. Y. Ho, "Named entity recognition for vietnamese real estate advertisements," in *2021 8th NAFOSTED Conference on Information and Computer Science (NICS)*, IEEE, 2021, pp. 23–28.

[24] J. Zenkert, A. Klahold, M. Fathi, "Towards extractive text summarization using multidimensional knowledge representation," in *2018 IEEE International Conference on Electro/Information Technology (EIT)*, IEEE, 2018, pp. 0826–0831.

[25] A. Alambo, T. Banerjee, K. Thirunarayan, M. Cajita, "Improving the factual accuracy of abstractive clinical text summarization using multi-objective optimization," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, IEEE, 2022, pp. 1615–1618.

[26] Z. A. Guven, M. O. Unalir, "Improving the bert model with proposed named entity recognition method for question answering," in *2021 6th International Conference on Computer Science and Engineering (UBMK)*, IEEE, 2021, pp. 204–208.

[27] S. T. Chavali, C. T. Kandavalli, T. Sugash, D. Gupta, "A study on named entity recognition with different word embeddings on gmb dataset using deep learning pipelines," in *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, 2022, pp. 1–5.

[28] Y. P. Kilic, D. Dinc, P. Karagoz, "Named entity recognition on morphologically rich language: Exploring the performance of bert with varying training levels," in *2020 IEEE International Conference on Big Data (Big Data)*, IEEE, 2020, pp. 4613–4619.

[29] P. N. Achyutha, S. Chaudhury, S. C. Bose, R. Kler, J. Surve, K. Kaliyaperumal, "User classification and stock market-based recommendation eng..."