Comparing Summarization Techniques For Generating News Summaries

Project Team 35
Team Member: Hrishikesh Pradhan

What is the task, and why is it important to users?
In recent years, many traditional news outlets have shifted to delivering news through their websites. This shift has intensified competition among these organizations, leading to practices such as SEO hacking and the inclusion of unnecessary fluff in articles.

The goal of this project is to develop and evaluate various Information Retrieval techniques for summarizing news articles. By providing concise and meaningful summaries, users will be able to quickly grasp the core information, helping them save time and avoid misleading or bloated content.

What do users' queries look like?
Normal google searches asking for news articles. These are simple English sentences.
Eg: ‘Latest development in Artificial Intelligence’, ‘New tariff rules’... etc

What kinds of results would be relevant to these queries? How many relevant results should there be per query?
A short paragraph summarizing the top news articles in about 150 to 200 words is a relevant result. The summary should include essential topics related to the user query. Ideally there will be only one result summary.

If relevant to your project, how should the results be organized (ranked list, clusters, summaries, etc.)?
Summaries.

What evaluation metrics would be appropriate for this task?
I am using ROUGE (Recall-Oriented Understudy for Gisting Evaluation) to evaluate generated summaries. ROUGE compares automatically produced summaries with human-written references, with scores ranging from 0 to 1—higher values indicate better similarity.

I focus on the following ROUGE metrics:

- ROUGE-1: Measures unigram (word-level) overlap.
- ROUGE-2: Measures bigram (two-word sequence) overlap.
- ROUGE-L: Based on the longest common subsequence, capturing sentence-level structural similarity.

Also use BertScore for evaluation.
BERTScore uses the power of BERT(Bidirectional Encoder Representations from Transformers), a state-of-the-art transformer-based model developed by Google, to understand the semantic meaning of words in a sentence. This leads to a more accurate representation of text similarity compared to traditional methods that rely on syntactic structures.

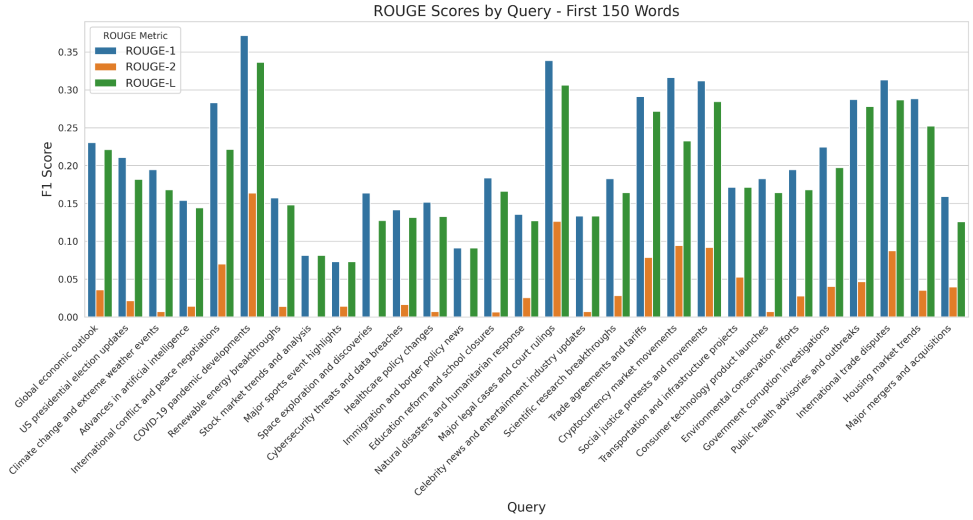
A description of your implementation and an analysis of its performance.

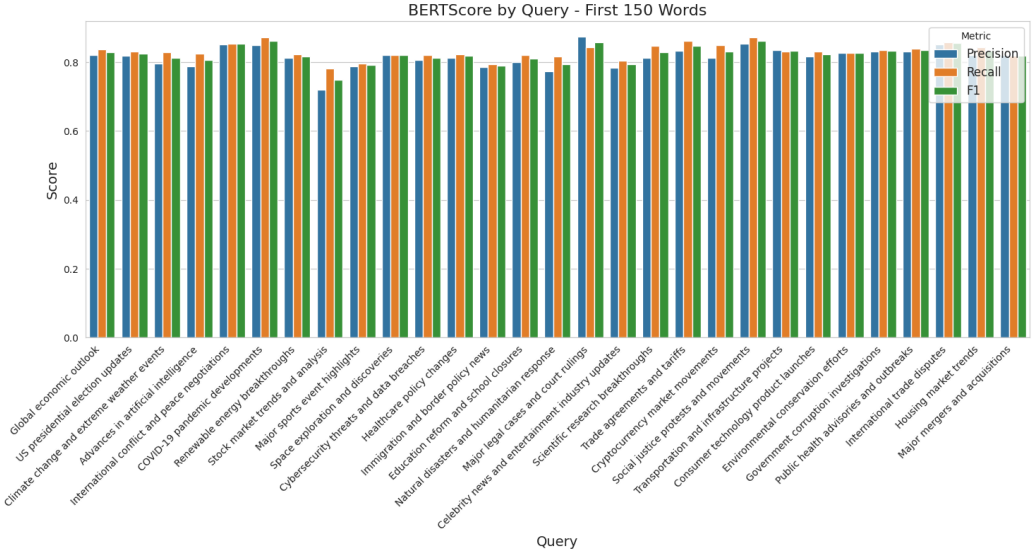
1. Write 30 user queries representing requests for news articles.
2. Use the DuckDuckGo search api library to fetch top 10 links(url to news websites hosting the news articles).
3. Store the results in json for further re-use.
4. Use newspaper4k library to fetch the article text and each of the 10 urls for each of the 30 articles.
5. Store the collected article text in json for further reuse.
6. Use Generative AI and self annotation to produce target summaries for each user query based on the corresponding collected articles.
7. Store the target summaries in json format for re-use.
8. Setup the rouge evaluation functions to compare the generated summaries(for each technique) and target summary.
9. Setup BertScore evaluation functions to compare the generated summaries(for each technique) and target summary.
10. Setup visualization functions that help in identifying trends.

Summarization Techniques Implemented and Evaluated

First 150 Words

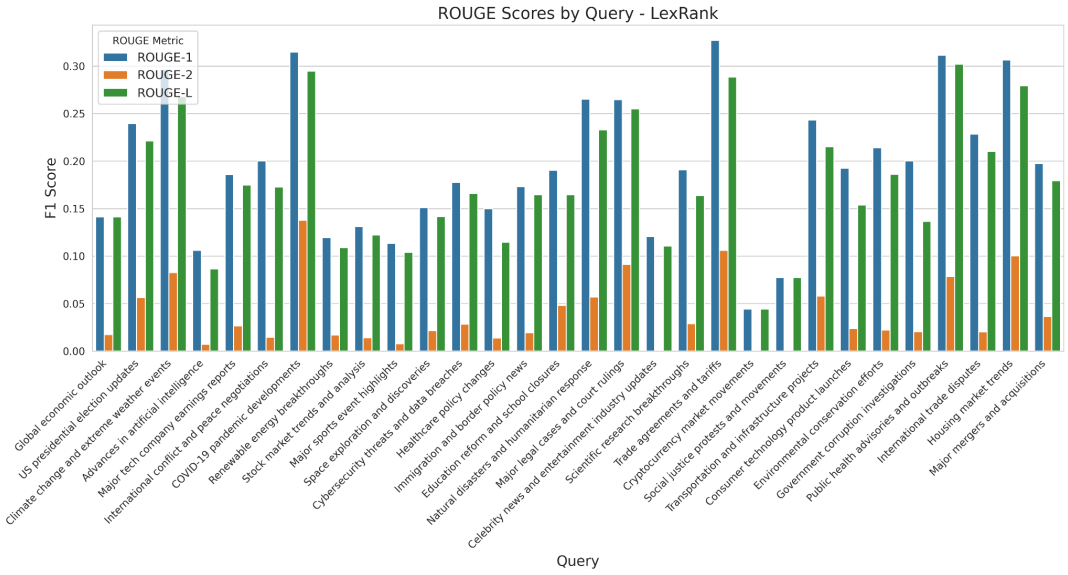
1. Simple Extraction: Takes the first 150 words from the source text with no advanced processing.
2. Position-Based: Based on the journalistic principle that important information appears first in news articles.
3. No Semantic Analysis: Doesn't consider the meaning or importance of sentences.
4. No Content Selection: Doesn't attempt to identify key information across the entire document.
5. Computationally Efficient: Very fast and requires minimal processing resources.

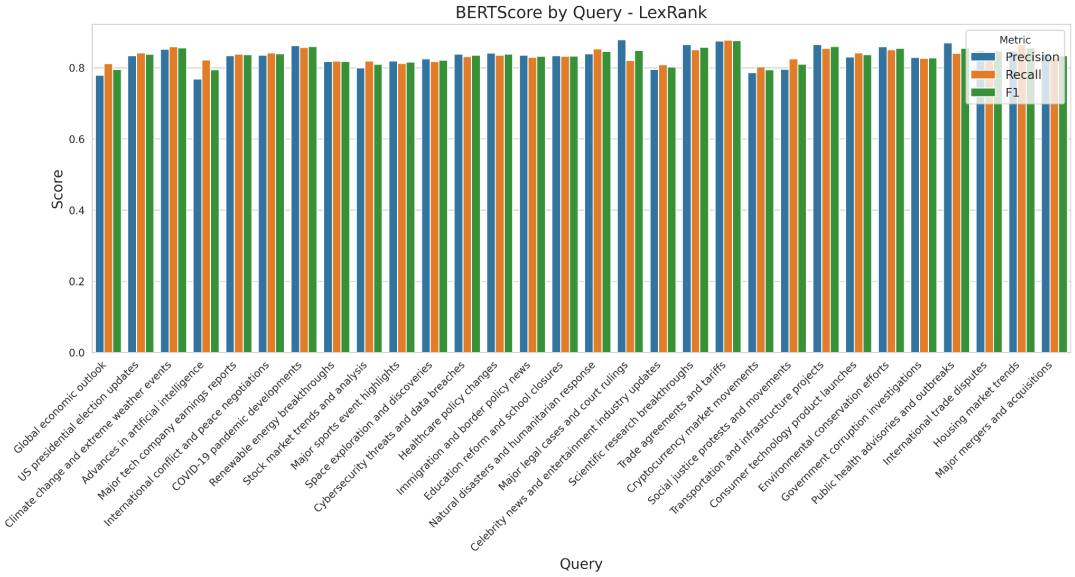




LexRank

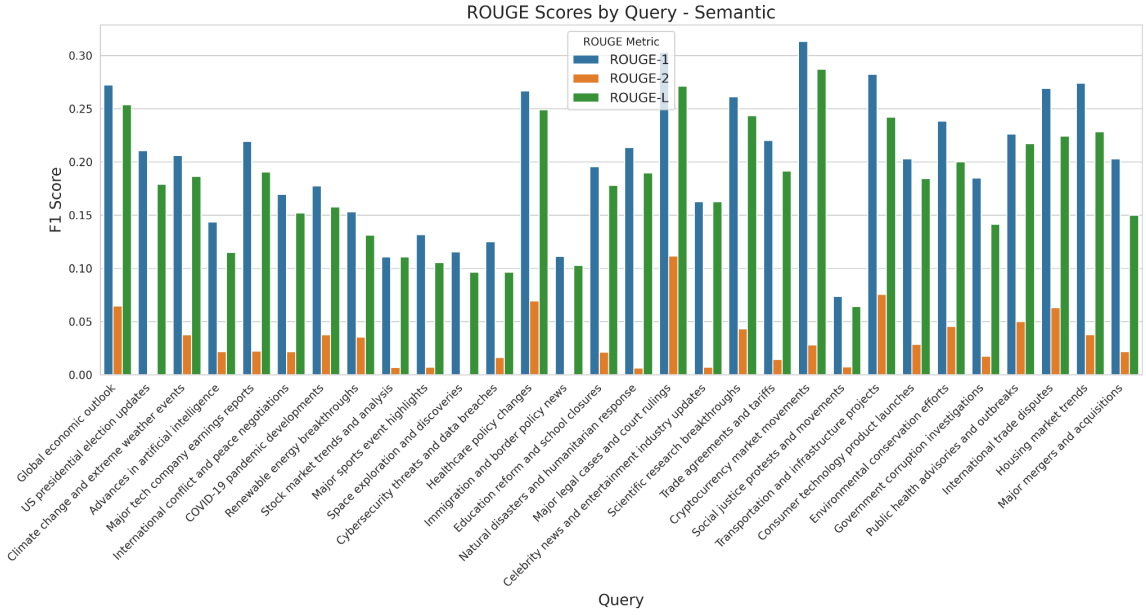
1. Graph-Based Algorithm: Represents sentences as nodes in a graph connected based on similarity.
2. Semantic Similarity: Uses cosine similarity of TF-IDF vectors to determine sentence relationships.
3. Centrality Measure: Applies PageRank-like algorithm to identify the most central (important) sentences.
4. Unsupervised Approach: Requires no training data or human annotations.
5. Content Diversity: Selects sentences that represent different aspects of the document.

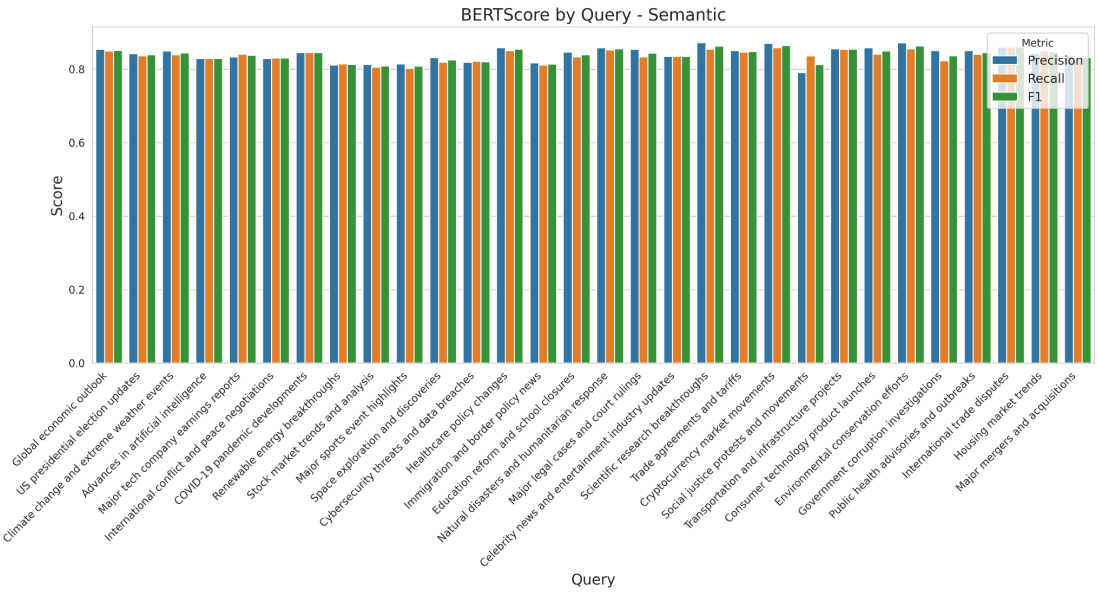




Semantic Summarization

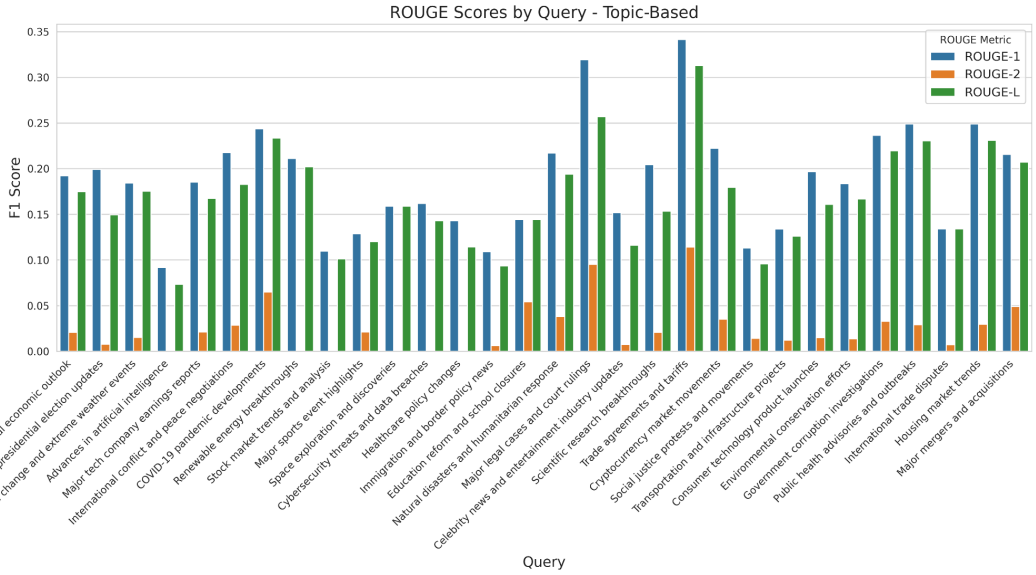
1. Meaning-Based: Focuses on capturing the semantic meaning of the text rather than just keywords.
2. Embedding-Based: Transforms sentences into semantic vector representations using language models.
3. Clustering Approach: Groups similar sentences based on their semantic embeddings.
4. Representative Selection: Selects sentences that best represent each semantic cluster.
5. Context-Aware: Better preserves contextual relationships between concepts than purely statistical methods.

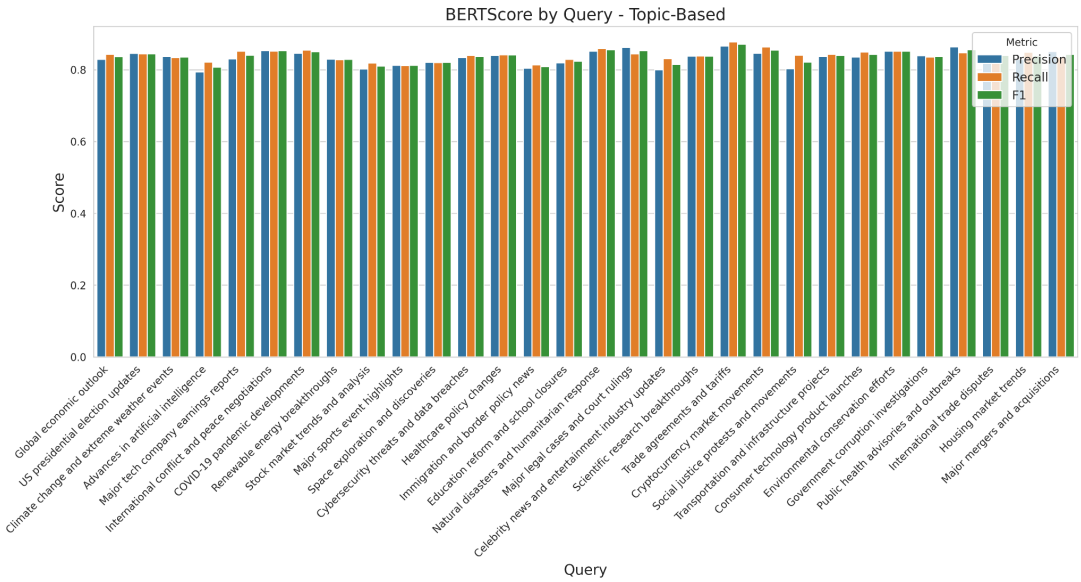




Topic-Based Summarization

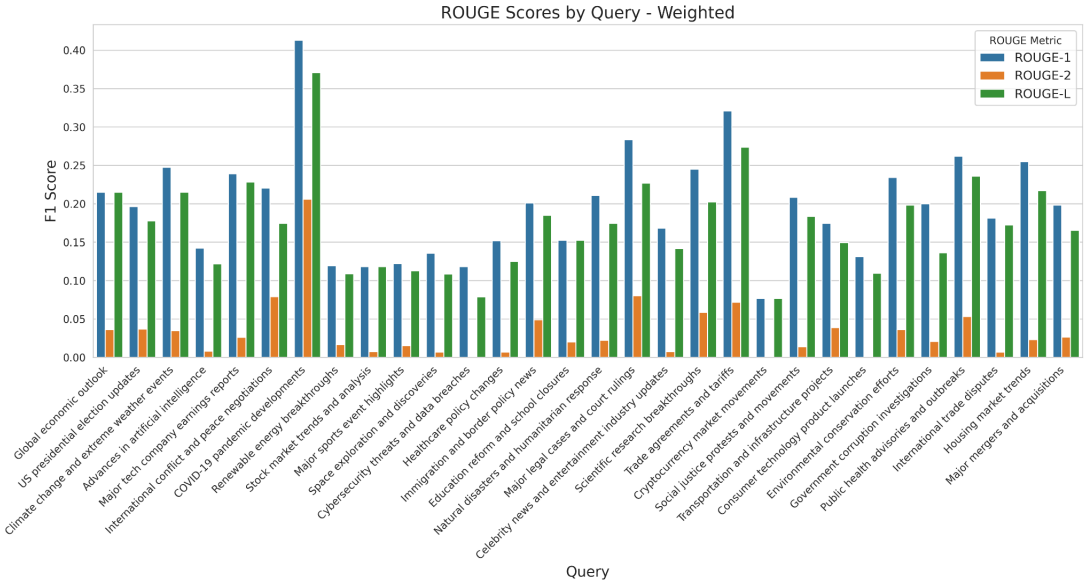
1. Topic Modeling: Uses techniques like LDA (Latent Dirichlet Allocation) to identify main topics in the text.
2. Multi-Document Awareness: Identifies common themes across multiple articles on the same subject.
3. Hierarchical Structure: Organizes content by topic importance and relevance.
4. Query Relevance: Prioritizes topics that align with the original query terms.
5. Coverage Optimization: Ensures the summary includes representation of all significant topics.

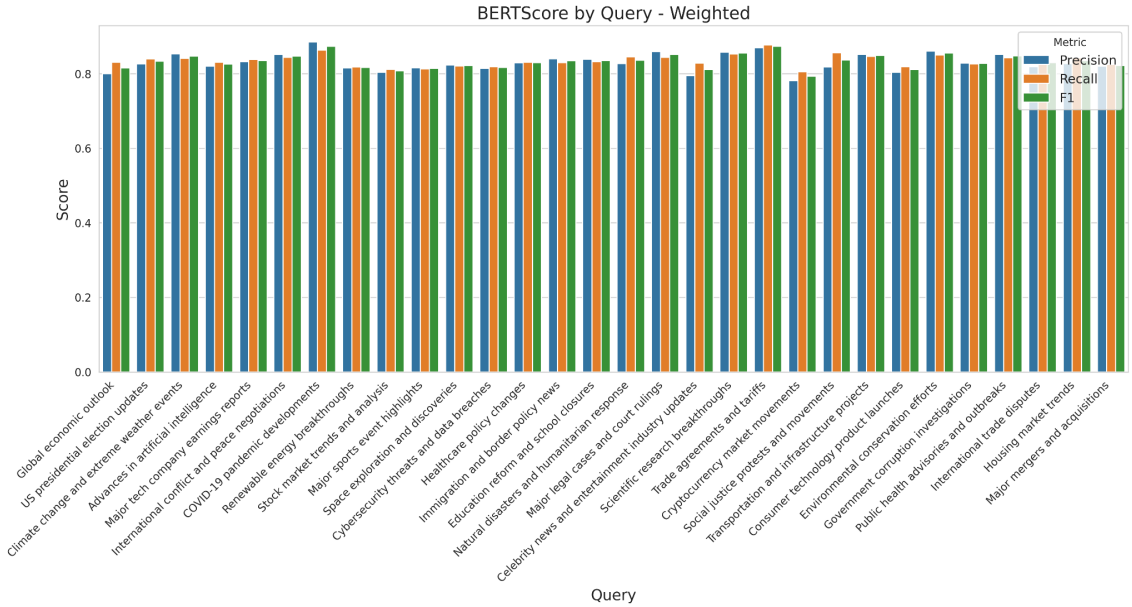




Weighted Summarization

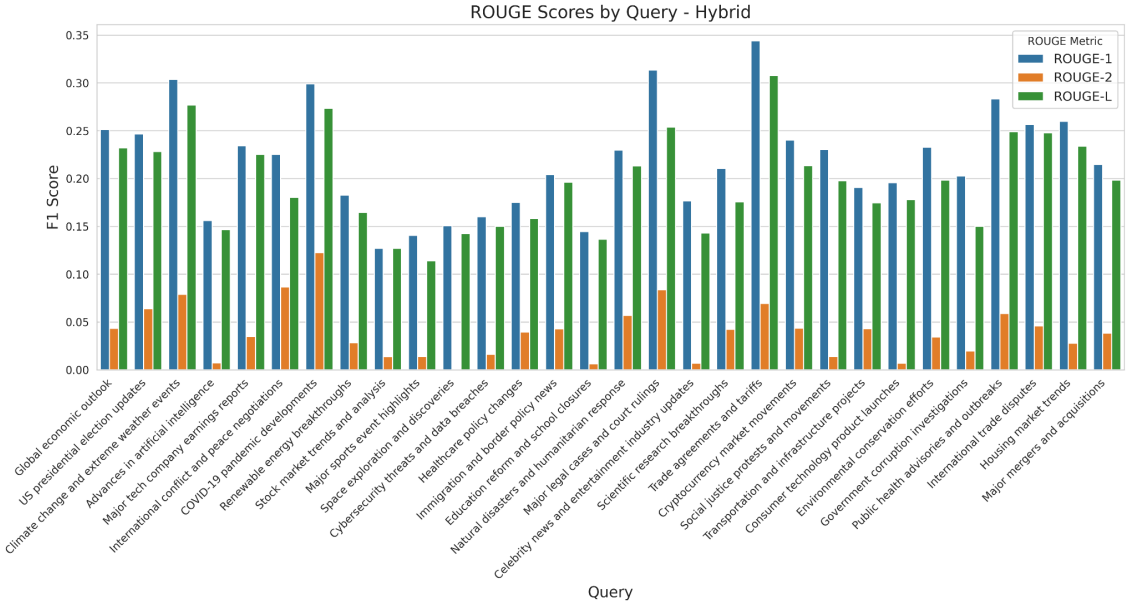
1. Multi-Factor Scoring: Assigns importance to sentences based on multiple weighted factors.
2. Customizable Weights: Allows adjustment of importance for different features (position, term frequency, etc.).
3. TF-IDF Application: Uses term frequency-inverse document frequency to identify key terms.
4. Position Bias: Incorporates the position of sentences in the document into the weighting scheme.
5. Threshold Selection: Selects sentences with scores above a certain threshold for inclusion in the summary.

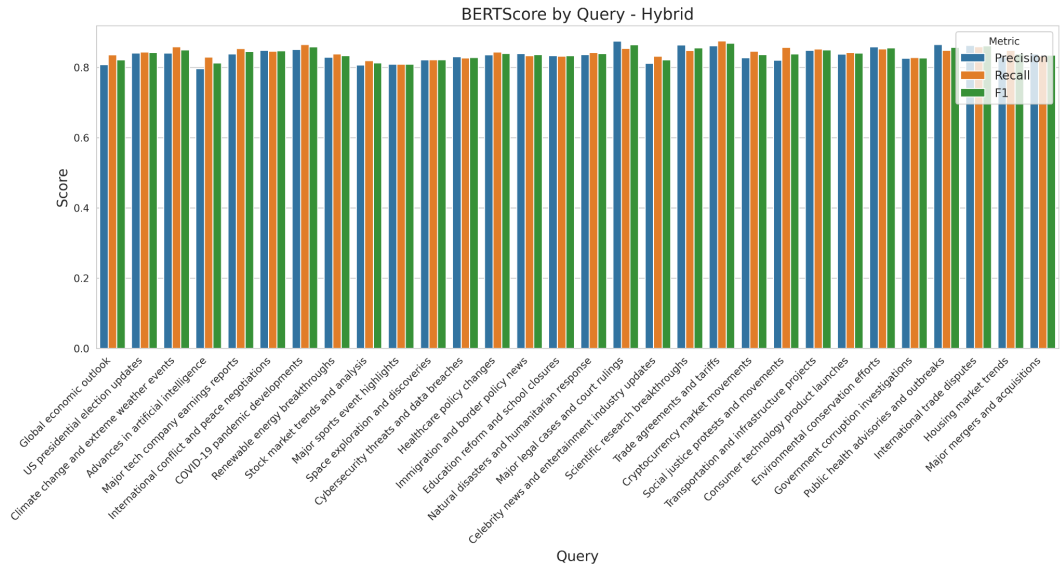




Hybrid Summarization

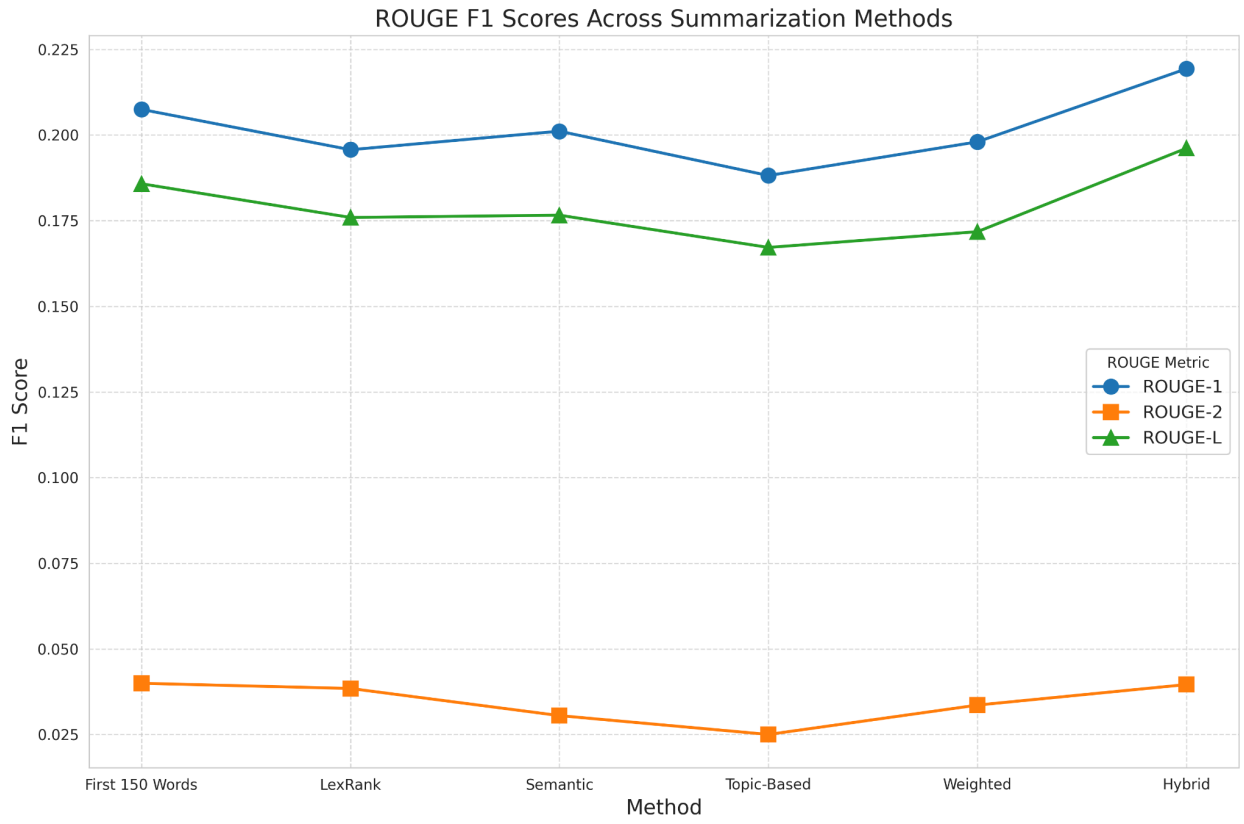
1. Combined Approach: Integrates multiple summarization techniques to leverage their complementary strengths.
2. Ensemble Method: May use voting or scoring from different algorithms to make final selections.
3. Balanced Processing: Considers both statistical importance (frequency) and semantic meaning of content.
4. Cross-Validation: Often employs verification across methods to confirm the importance of content.
5. Adaptability: Can emphasize different component methods depending on document type and content structure.





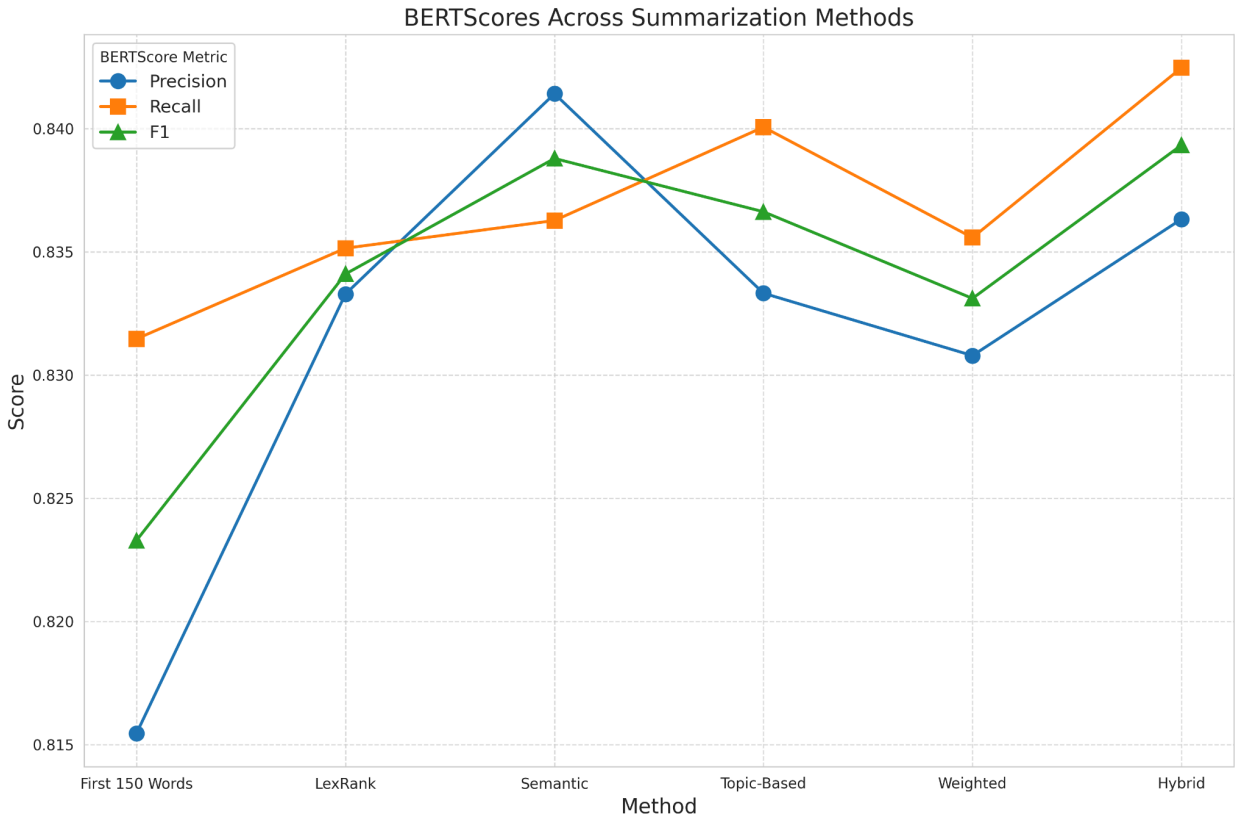
ROUGE Evaluation Summary Table

Technique	Average ROUGE-1 F1 Score	Average ROUGE-2 F1 Score	Average ROUGE-L F1 Score
First 150 Words	0.2075	0.0400	0.1858
LexRank	0.1957	0.0385	0.1759
Semantic	0.2011	0.0305	0.1766
Topic-Based	0.1882	0.0250	0.1672
Weighted	0.1980	0.0336	0.1718
Hybrid	0.2193	0.0396	0.1961



BertScore Evaluation Summary Table

Technique	Average Precision	Average Recall	Average F1 Score
First 150 Words	0.8155	0.8315	0.8233
LexRank	0.8333	0.8351	0.8341
Semantic	0.8414	0.8363	0.8388
Topic-Based	0.8333	0.8401	0.8366
Weighted	0.8308	0.8356	0.8331
Hybrid	0.8363	0.8425	0.8393



Analysis of Summarization Method Performance Results

The evaluation of six distinct text summarization techniques—First 150 Words, LexRank, Semantic, Topic-Based, Weighted, and Hybrid approaches—revealed surprisingly similar performance metrics across both ROUGE and BERTScore evaluations. This unexpected consistency suggests that for our particular news article dataset, the fundamental extractive nature of all methods led to comparable information selection, despite their algorithmic differences. The narrow range of scores may indicate that the journalistic writing style, which front loaders key information, creates a natural ceiling effect where even sophisticated methods struggle to meaningfully outperform simpler approaches. This pattern challenges the assumption that more complex summarization algorithms necessarily yield better results on well-structured news content.

A deeper inspection of query-specific performance reveals more nuanced distinctions between methods. While overall averages appear similar, some approaches demonstrated relative strengths on particular topics—with the Hybrid method showing marginally better consistency across diverse queries, and the Semantic approach capturing slightly more contextual meaning according to BERTScore metrics. These findings suggest that the effectiveness of summarization methods may be more dependent on content characteristics and query types than algorithmic sophistication alone. Future work should focus on creating more discriminative evaluation frameworks, potentially incorporating human judgments of summary quality, and exploring how these methods perform across different document genres where information distribution patterns vary more significantly from news articles.

Sample Results

Query: Global economic outlook
Narrative: The user wants to know the latest developments in global economics.

Target Summary: The global economic outlook for 2025 is marked by uncertainty, with trade tensions and tariffs weighing heavily on investor sentiment. Surveys indicate the most bearish outlook in decades, largely due to the Trump administration's trade policies and fears of stagflation—sluggish growth combined with persistent inflation. Mortgage refinance rates remain elevated, and experts do not expect a significant drop unless the Federal Reserve enacts multiple rate cuts. Meanwhile, China's economy is a bright spot, with companies like Royal Philips investing in healthcare innovation and AI, buoyed by government support and a large, skilled workforce. Despite ongoing geopolitical uncertainties, consumer confidence and healthcare demand in China are rising. Globally, markets remain volatile, with many investors wary of recession risks and inflationary pressures. The consensus is that, barring major policy shifts or resolution of trade disputes, global growth is likely to remain subdued in the near term.

First 150 word Summary: Mortgage refinance rates have been moderating in response to rising layoffs and worries about a potential economic downturn . Despite the recent dip in rates , experts don't expect another refinancing boom like in 2020 and 2021 , when mortgage rates dropped to historic lows . Until mortgage rates move below 6 % , which isn't guaranteed this year , refinancing activity is likely to remain limited . Most homeowners refinance to save money on their monthly payment with a lower interest rate . However , you might be considering refinancing for other reasons , such as modifying your loan term or type . Mortgage refinance rates change daily depending on multiple economic and political factors . Check out our weekly mortgage rate forecast for expert predictions on where rates are headed . Today 's average refinance rates 30-year fixed-rate 15-year fixed-rate 30-year fixed-rate jumbo 5/1

LexRank Summary: Today's average refinance rates 30-year fixed-rate 15-year fixed-rate 30-year fixed-rate jumbo 5/1 ARM 10-year fixed-rate 30-year fixed-rate refinance 15-year fixed-rate refinance 10-year fixed refinance Today's average mortgage rates on Apr. 30-year fixed-rate refinance The average 30-year fixed refinance rate right now is 6.82%, an increase of 8 basis points compared to one week ago. 15-year fixed-rate refinance For 15-year fixed refinances, the average rate is currently at 6.10%, an increase of 4 basis points compared to one week ago. Also, 15-year refinance rates are typically lower than 30-year refinance rates, which will help you save more in the long run. 10-year fixed-rate refinance The average 10-year fixed refinance rate right now is 6.02%, a decrease of 1 basis points compared to one week ago.

Semantic Summary: Despite the recent dip in rates, experts don't expect another refinancing boom like in 2020 and 2021, when mortgage rates dropped to historic lows. Until mortgage rates move below 6%, which isn't guaranteed this year, refinancing activity is likely to remain limited. Most homeowners refinance to save money on their monthly payment with a lower interest rate. Refinance rate news Late last year, mortgage rates climbed back into the 7%

range after the Federal Reserve said it would adopt a slower pace of interest rate cuts in 2025 due to limited progress on inflation. What to know about 2025 refinance rate expectations Experts say the Fed's rate cuts could help mortgage interest rates drop further by the end of 2025. Overall, refinance rates will not become significantly more affordable in the next year without multiple interest rate cuts and weaker economic data.

TopicBased Summary: Most homeowners refinance to save money on their monthly payment with a lower interest rate. About these rates: Bankrate's tool features rates from partner lenders that you can use when comparing multiple mortgage rates. (A basis point is equivalent to 0.01%.) 15-year fixed-rate refinance For 15-year fixed refinances, the average rate is currently at 6.10%, an increase of 4 basis points compared to one week ago. As shown below, this fell to a record low. Fears of higher inflation are also raising concerns about stagflation, a nightmare scenario for the economy that involves sluggish growth and stubbornly high prices. That's the most concentrated investors have been on a tail risk in 15 years, the bank added. \"We look at the specific consumer needs in China. At the same time, we use China as a base to provide services to the world,\" Jakobs said.

Weighted Summary: Today's average refinance rates 30-year fixed-rate 15-year fixed-rate 30-year fixed-rate jumbo 5/1 ARM 10-year fixed-rate 30-year fixed-rate refinance 15-year fixed-rate refinance 10-year fixed refinance Today's average mortgage rates on Apr. Refinance rate news Late last year, mortgage rates climbed back into the 7% range after the Federal Reserve said it would adopt a slower pace of interest rate cuts in 2025 due to limited progress on inflation. A 30-year fixed refinance will typically have lower monthly payments than a 15-year or 10-year refinance, but it will take you longer to pay off and typically cost you more in interest over the long term. A 10-year refinance typically has the lowest interest rate but the highest monthly payment of all refinance terms. It's around 10 percent of our total global revenue,\" Roy Jakobs, CEO of Royal Philips, said during a recent interview with China Daily.

Hybrid Summary: Most homeowners refinance to save money on their monthly payment with a lower interest rate. Today's average refinance rates 30-year fixed-rate 15-year fixed-rate 30-year fixed-rate jumbo 5/1 ARM 10-year fixed-rate 30-year fixed-rate refinance 15-year fixed-rate refinance 10-year fixed refinance\n\nToday's average mortgage rates on Apr. Refinance rate news Late last year, mortgage rates climbed back into the 7% range after the Federal Reserve said it would adopt a slower pace of interest rate cuts in 2025 due to limited progress on inflation. Overall, refinance rates will not become significantly more affordable in the next year without multiple interest rate cuts and weaker economic data. A 30-year fixed refinance will typically have lower monthly payments than a 15-year or 10-year refinance, but it will take you longer to pay off and typically cost you more in interest over the long term. \"The China market is very important for Philips.

Query: Natural disasters and humanitarian response
Narrative: The user wants to know the latest developments in disasters and humanitarian response.

Target Summary: In April 2025, natural disasters and humanitarian response efforts have been shaped by both innovation and crisis. In Myanmar, a catastrophic 7.7 magnitude earthquake struck amid ongoing civil conflict, compounding the suffering of already vulnerable communities. The military junta's aid blockade, restricted communications, and continued airstrikes have severely hampered relief operations, forcing civilians to rely on each other for rescue and survival. Despite calls for a ceasefire to facilitate humanitarian aid, skepticism remains high as the junta's actions mirror past patterns of obstructing international assistance and weaponizing disasters for political control. Meanwhile, technological advances are transforming disaster response elsewhere. The Philippines launched a pilot of the UN-led DISHA initiative, with Globe Telecom providing AI-driven data to map population movements and poverty in real time. This partnership aims to improve the speed and targeting of humanitarian aid, addressing challenges posed by manual assessments and limited resources.International military cooperation is also advancing, as seen in the India-U.S. Tiger Triumph 2025 exercise, which enhanced joint disaster relief capabilities through complex operations and integrated training.

First 150 word Summary: Kakinada (Andhra Pradesh) [India] , April 12 (ANI) : The Fourth Edition of Bilateral Tri-Service Humanitarian Assistance and Disaster Relief (HADR) Amphibious Exercise Tiger Triumph 2025 between India and the United States culminated with a Distinguished Visitors ' (DV) Day at Kakinada on April 11 . The DV Day was attended by the Flag Officer Commanding Tamil Nadu & Puducherry Naval Area (FOTNA) , US Consular General , Commander US Navy Strike Group Five and Dy GOC 54 Infantry Division , along with other senior dignitaries . The Distinguished Visitors ' Day witnessed seamless execution of complex operations at/off Kakinada , including Standoffs and Hard Beaching , Slithering Operations by Special Operations Forces from SC and Mi17V5 Helicopters , the participation of C130 Aircraft , and integrated Air Operations by the Indian Navy , Indian Army , Indian

LexRank Summary: Many local organizations still rely on manual assessments for aid distribution, which can delay response efforts. Foreign governments and international organizations must categorically reject the military junta's demand that it control aid distribution and establish direct funding mechanisms with local actors, such as community-based organizations and civil society organizations that are directly supporting quake-affected areas. MANILA, Philippines, April 2, 2025 /PRNewswire/ -- In a landmark pro bono agreement, Globe, a top telco and digital solutions provider in the Philippines, has become the first mobile network operator (MNO) to partner with the UN Global Pulse's Data Insights for Social and Humanitarian Action (DISHA) coalition, harnessing AI-driven data for disaster preparedness and response. We're excited to welcome Globe Telecom to the DISHA Partner community and look forward to working side-by-side with Globe and humanitarian agencies operating in the Philippines to validate and scale the DISHA Socio-Economic Mapping solution in the country.

Semantic Summary: Globe has become the first mobile network operator to join the UN Global Pulse 2019s Data Insights for Social and Humanitarian Action (DISHA) coalition. As the international community rallies to provide aid, organizations must ensure that this support directly reaches affected civilians on the ground. The Myanmar military junta\u2019s current prevention of international humanitarian aid for earthquake victims in Sagaing mirrors their deadly playbook from 2008 Cyclone Nargis, when they blocked critical relief efforts while thousands perished. For years, even before this disaster, these local actors have successfully delivered aid to communities most in need, even in areas considered inaccessible by the junta. These tools help humanitarian organizations better respond to disasters, reaching those in most need. \"Our partnership with DISHA provides a crucial data-for-good solution that will enable disaster management agencies and partner organizations to deliver life-saving aid to those who need it the most.

TopicBased Summary: The Sea Phase, undertaken off the coast of Kakinada from April 8 to 11, featured complex maritime operations, aircraft cross-deck landings, troops landing at Kakinada, and NDRF activities. Exercise Tiger Triumph 2025 marks a significant milestone in reinforcing the shared commitment of India and the United States towards regional stability, disaster response cooperation, and enhancing interoperability between their armed forces. Globe has become the first mobile network operator to join the UN Global Pulse’s Data Insights for Social and Humanitarian Action (DISHA) coalition. Many local organizations still rely on manual assessments for aid distribution, which can delay response efforts. From March 28 until the present, DVB News reports that the junta has launched over 40 air and artillery strikes across Myanmar, resulting in at least 12 fatalities. Additionally, world leaders must demand the immediate cessation of airstrikes and military operations, especially in earthquake-affected regions.

Weighted Summary: Globe has become the first mobile network operator to join the UN Global Pulse’s Data Insights for Social and Humanitarian Action (DISHA) coalition. Many relief organizations and local communities remain wary with the understanding that while airstrikes might temporarily cease in the earthquake ravaged central regions of Sagaing and Mandalay, military operations could continue unabated in other parts of the country. MANILA, Philippines, April 2, 2025 /PRNewswire/ -- In a landmark pro bono agreement, Globe, a top telco and digital solutions provider in the Philippines, has become the first mobile network operator (MNO) to partner with the UN Global Pulse's Data Insights for Social and Humanitarian Action (DISHA) coalition, harnessing AI-driven data for disaster preparedness and response.

Hybrid Summary: Many local organizations still rely on manual assessments for aid distribution, which can delay response efforts. These tools enhance humanitarian response efforts by enabling targeted aid distribution. MANILA, Philippines, April 2, 2025 /PRNewswire/ -- In a landmark pro bono agreement, Globe, a top telco and digital solutions provider in the Philippines, has become the first mobile network operator (MNO) to partner with the UN Global Pulse's Data Insights for Social and Humanitarian Action (DISHA) coalition, harnessing AI-driven data for disaster preparedness and response. \"We're excited to welcome Globe Telecom to the DISHA Partner community and look forward to working side-by-side with Globe

and humanitarian agencies operating in the Philippines to validate and scale the DISHA Socio-Economic Mapping solution in the country. \"Our partnership with DISHA provides a crucial data-for-good solution that will enable disaster management agencies and partner organizations to deliver life-saving aid to those who need it the most.

What milestones in your grade contract did you complete?

I have completed all milestones stated in my grade contract.

External Links

Github

Data collection and method implementation Jupyter Notebook

Evaluation Notebook