# Executive summary Phase 3

## Triage System Results and Analysis

### Introduction

In this project phase, we developed and implemented a triage system to prioritize and categorize comments within collaborative documents. We integrated these results using advanced NLP techniques, including Hierarchical Capsule Networks, Zero-shot Learning, and N-shot Learning, to create a comprehensive triage system.

### Triage System Overview

### Objectives

**Prioritization:** Assign priority levels (High, Medium, Low) to comments based on their content and relevance.

**Categorization:** Classify comments into specific categories such as "Legal/Contractual Issues," "Operational Details," "Research/Technical Discussions," "General Updates," and "Data Analysis."

**Integrated Labeling:** Combine insights from Zero-shot, N-shot, and LDA models to provide a unified triage label.

### Methodology

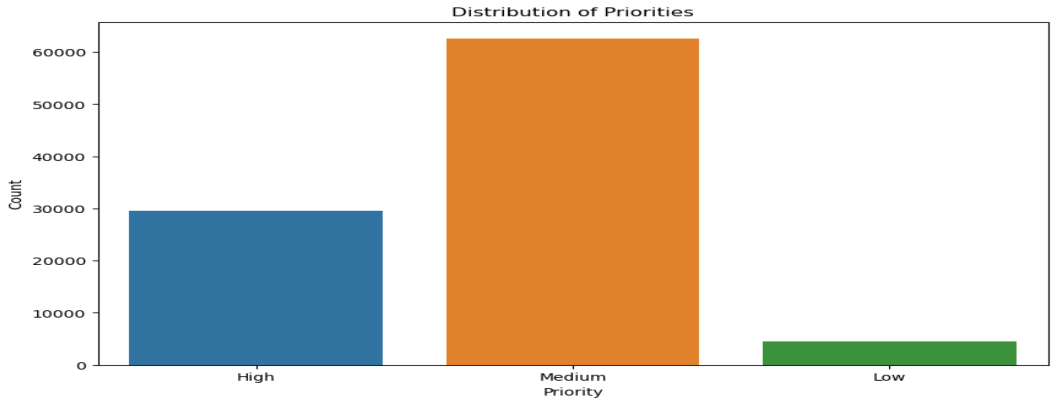
**Data Integration:** We combined results from Zero-shot, N-shot, and LDA models into a unified dataset.

**Priority Assignment:** Priority levels were determined based on the content and predicted labels.

**Category Assignment:** Categories were assigned based on LDA model results and refined by additional classification layers.

### Results and Visualizations

### Priority Distribution



**Visualization:** Bar chart showing the distribution of comments across different priority levels (High, Medium, Low).

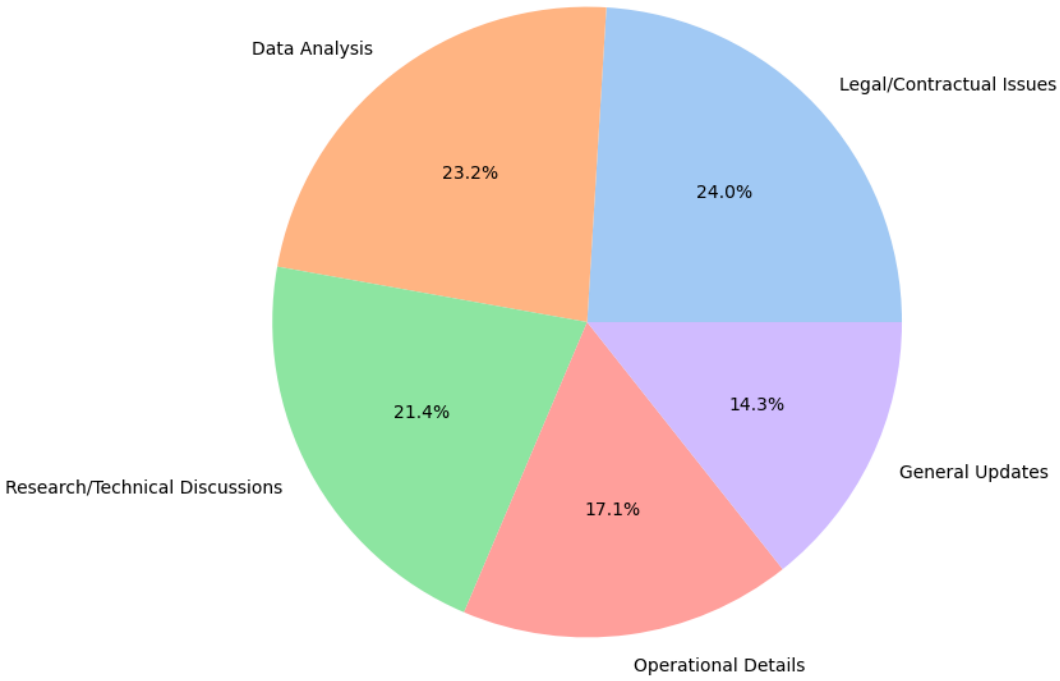
**Findings:**

- The majority of the comments were categorized as medium priority.

- A smaller proportion of comments were categorized as High and Low priority.

## Category Distribution

**Visualization:** Pie chart representing the distribution of comments across different categories.

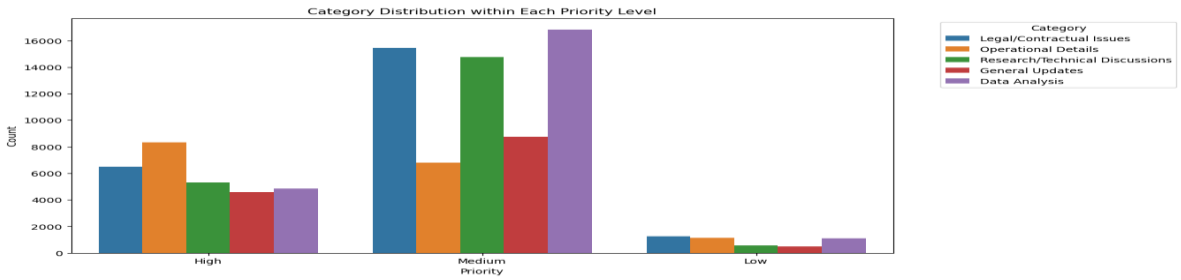



**Findings:**

- The comments were distributed across five main categories: Legal/Contractual Issues, Operational Details, Research/Technical Discussions, General Updates, and Data Analysis.

- Most comments fell under "Legal/Contractual Issues" and "Data Analysis."

## Category Distribution within Each Priority Level

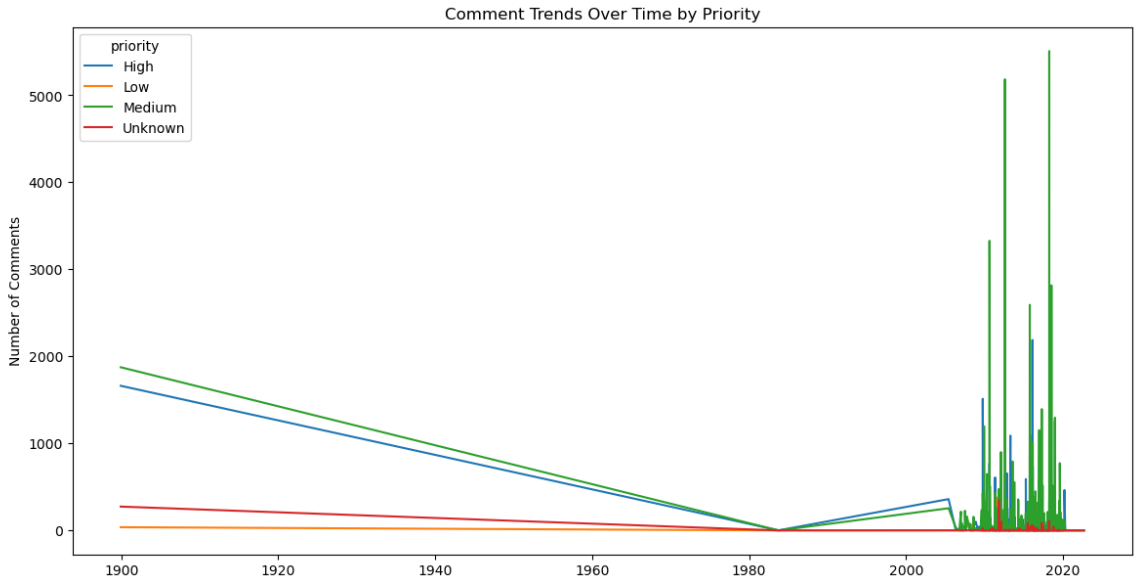


**Visualization:** A stacked bar chart shows the distribution of categories within each priority level.

- Each priority level had a mix of categories, with specific categories being more prevalent in High priority and others in Medium and Low priority.

## Comment Trends Over Time by Priority

**Visualization:** Line chart showing the trend of comment counts over time by priority level.

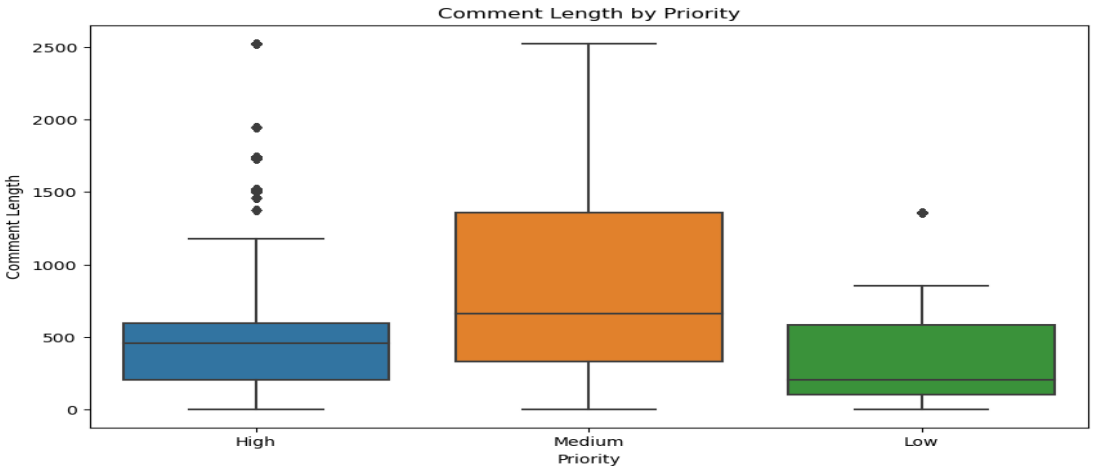



**Findings:**

- A significant increase in medium-priority comments has been observed recently.

## Comment Length by Priority

**Visualization:** Box plot comparing the length of comments across different priority levels.

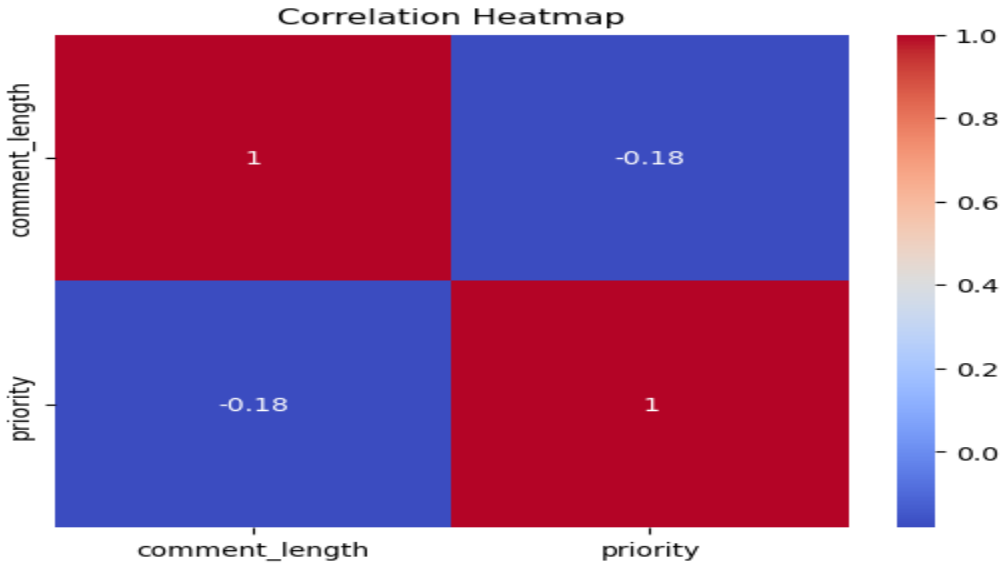


**Findings:**

- Medium priority comments tended to be longer on average, with a more excellent range of lengths.

## Correlation Analysis

**Visualization:** Heatmap showing the correlation between comment length and priority.




**Findings:**

- A weak negative correlation was observed between comment length and priority, suggesting that longer comments were slightly more likely to be categorized as Medium priority.

## Conclusions

**Effective Triage:** The triage system successfully prioritized and categorized comments, providing valuable insights into the nature and urgency of the comments.

**Balanced Categorization:** The comments were well-distributed across different categories, with a notable concentration in Legal/Contractual Issues and Data Analysis.

**Priority Insights:** The distribution of priorities highlighted that most comments are of Medium urgency, with fewer comments requiring immediate attention.

## Next Steps

## Graph Neural Networks (GNNs) for Priority Triage

- **Why It is Novel:** GNNs are innovative because they model relationships between data points (e.g., comments) like traditional models do not. They understand how the importance of one comment can influence others based on their connections.

- **How It Works:** GNNs represent comments as nodes in a graph and learn from the connections (edges) between them. This allows the model to prioritize comments by considering the context provided by related comments, making it more context-aware.
- **Why It is Different:** Unlike traditional models that treat each comment independently, GNNs consider the interdependencies between comments, providing a more holistic approach to prioritization.

### Reinforcement Learning for Adaptive Priority Triage

- **Why It is Novel:** Reinforcement Learning (RL) offers a dynamic and adaptive approach to prioritization, where the model learns from the outcomes of its actions, improving over time with feedback.
- **How It Works:** An RL model adjusts comment priorities based on rewards and penalties from previous decisions, allowing it to learn which comments require more urgent attention in real time.
- **Why It is Different:** RL is different from static models because it continuously learns and adapts, providing a flexible system that evolves with user needs and feedback.

### Multitask Learning for Integrated Triage and Classification

- **Why It is Novel:** Multitask Learning (MTL) leverages shared knowledge across related tasks, improving performance on each task by learning them together rather than separately.
- **How It Works:** MTL simultaneously trains on tasks like categorization, prioritization, and sentiment analysis, using shared representations to enhance the model's overall accuracy and efficiency.
- **Why It is Different:** Unlike single-task models focusing on one task at a time, MTL combines multiple related tasks, leading to a more comprehensive and efficient system.

### Conclusion

In this phase, we have built an advanced triage system that effectively prioritizes and categorizes comments using cutting-edge NLP techniques, revealing critical insights into comment urgency and content. Our analysis showed a balanced distribution across categories, with most comments given medium priority. Next, we plan to explore Graph Neural Networks, Reinforcement Learning, and Multitask Learning to enhance the system, making it more adaptive, context-aware, and comprehensive. These advancements will push our triage system towards greater efficiency and effectiveness.