ChatGPT Usage Analysis Report

1. Introduction

This report provides a comprehensive analysis of ChatGPT usage, focusing on user interactions, categorized message trends, sentiment analysis, and behavioral patterns. The analysis leverages structured data from conversation logs and applies categorization, visualization, and sentiment analysis techniques. The goal is to uncover insights into how ChatGPT is utilized for various tasks, including homework, learning, and technical assistance.

2. Data Collection and Preprocessing

- **Source**: Conversation logs were extracted from ChatGPT interactions, containing metadata such as timestamps, author roles, and message content.
- **Preprocessing Steps**:
    1. Missing timestamps were handled by dropping invalid entries.
    2. Message content was standardized to ensure consistent formatting.
    3. Messages were categorized into meaningful groups using keyword-based heuristics.
    4. Short messages (fewer than three words) were identified as a separate category for detailed subanalysis.

3. Key Findings and Observations

3.1 Final Category Distribution

Figure: Final Categories Pie Chart

This chart illustrates the overall distribution of messages across predefined categories:

- **Homework/Projects**: Represents interactions aimed at solving academic or project-based tasks, accounting for a significant portion of usage.
- **Learning**: Captures questions related to understanding new topics or concepts.
- **Technical Help**: Involves queries about installations, debugging, or configuration issues.
- **ChatGPT-Related**: Queries specifically about ChatGPT's functionality and AI models.
- **Short Messages**: Brief messages categorized separately for detailed subanalysis.
- **Other Categories**: Miscellaneous messages that could not be grouped meaningfully.

Observation: Homework/Projects and Learning dominate the usage, highlighting ChatGPT's role in academic and self-directed learning support.

3.2 Short Message Distribution

Figure: Short Messages Pie Chart

This chart delves into the distribution of short messages, breaking them into:

- **Questions**: Short queries like "What?", "How?", or "Why?".
- **Responses**: Simple acknowledgments such as "Yes", "No", or "OK".
- **Emoji Messages**: Messages composed primarily of emojis.
- **Other Short Messages**: Miscellaneous brief communications.

Observation: Questions form the majority, indicating that users frequently use short interactions for quick inquiries.

3.3 Temporal Trends

Figure: Messages Over Time Line Chart

This visualization depicts daily trends in message activity:

- Peaks in message activity correspond to specific deadlines or project-heavy days.
- Consistent usage over time suggests a steady reliance on ChatGPT.

Figure: Hourly Message Activity Bar Chart

This chart shows the distribution of messages by the hour of the day:

- Activity peaks in the late evening, coinciding with typical study or work schedules.

Observation: Users heavily engage with ChatGPT during evening hours, indicating its utility as a late-hour productivity tool.

3.4 Sentiment Analysis

Figure: Sentiment Distribution Bar Chart

This chart illustrates the polarity of user messages across the dataset:

- Sentiment ranges from highly positive to highly negative, with a majority of messages being neutral.

Figure: Average Sentiment by Category Bar Chart

This bar chart compares average sentiment across categories:

- Homework/Projects show slightly neutral-positive sentiment, reflecting the problem-solving nature of the category.
- ChatGPT-Related interactions exhibit positive sentiment, likely due to users' curiosity and satisfaction with the AI's capabilities.

Observation: Sentiment trends align with the functional goals of each category, with positive sentiment linked to exploratory and technical support inquiries.

3.5 Correlation Analysis

Figure: Hour of Day vs Sentiment Scatter Plot

This scatter plot explores the relationship between the hour of the day and sentiment polarity:

- Sentiments remain largely neutral across hours but show minor variations during late-night hours.

Figure: Message Length vs Sentiment Scatter Plot

This scatter plot examines the relationship between message length and sentiment polarity:

- Longer messages are often neutral, reflecting detailed questions or explanations.

Observation: Sentiment is relatively unaffected by time or message length, underscoring the neutral nature of most interactions.

4. Insights and Recommendations

- **Insight**: ChatGPT is primarily used for academic and learning support, with significant activity during evening hours.
- **Recommendation**: Customizing ChatGPT for academic settings could enhance its utility, such as integrating with project management tools or providing advanced debugging capabilities.
- **Future Directions**:
    - Developing dynamic dashboards for real-time interaction tracking.
    - Enhancing sentiment analysis with pre-trained transformer models for greater accuracy.

5. Conclusion

This analysis reveals ChatGPT's extensive use in academic and learning contexts, supported by detailed temporal and sentiment trends. The findings emphasize its value as a productivity and support tool. Future iterations of this analysis could incorporate predictive modeling to further understand user behavior and optimize ChatGPT's functionality.

Acknowledgments

Special thanks to ChatGPT for providing the dataset and tools used in this analysis.