



Cyberbullying Detection Report

Generated on: April 10, 2025 02:12:10

Executive Summary

This report analyzes 12 comments for potential cyberbullying content. The analysis found that **25.0%** of the comments contain language consistent with cyberbullying. A total of **3** comments were identified as potentially problematic.

Key findings:

- Most common terms in problematic comments: yap, louise, mariele, liber, stfu
- Primary sentiment in flagged comments: Positive

Based on these findings, immediate attention is recommended. Detailed analysis and specific recommendations are provided in the following sections.

Analysis Details

Analysis Summary

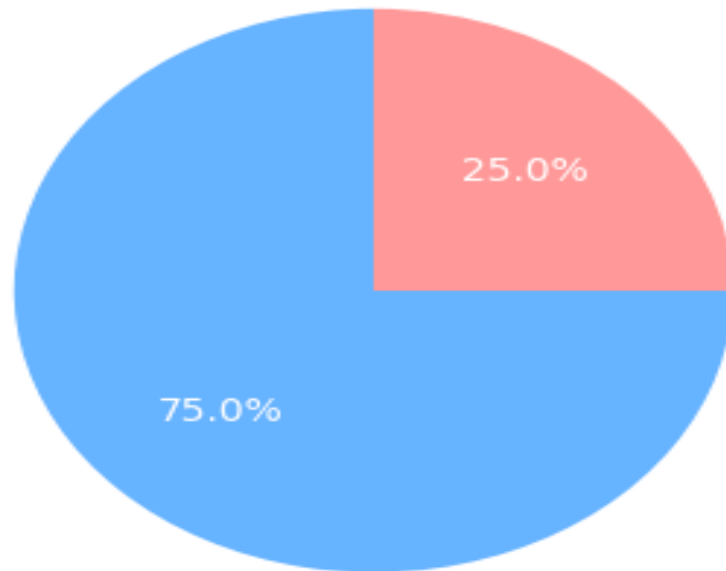
Total Comments Analyzed:	12
Cyberbullying Comments:	3 (25.0%)
Normal Comments:	9 (75.0%)
High Confidence Predictions (>90%):	12 (100.0%)
Comments in Selection List:	0

The analysis detected some cyberbullying content (25.0%), but the majority of comments appear to be normal. The model shows high confidence in its predictions, with 100.0% of predictions having over 90% confidence.

Visual Analysis

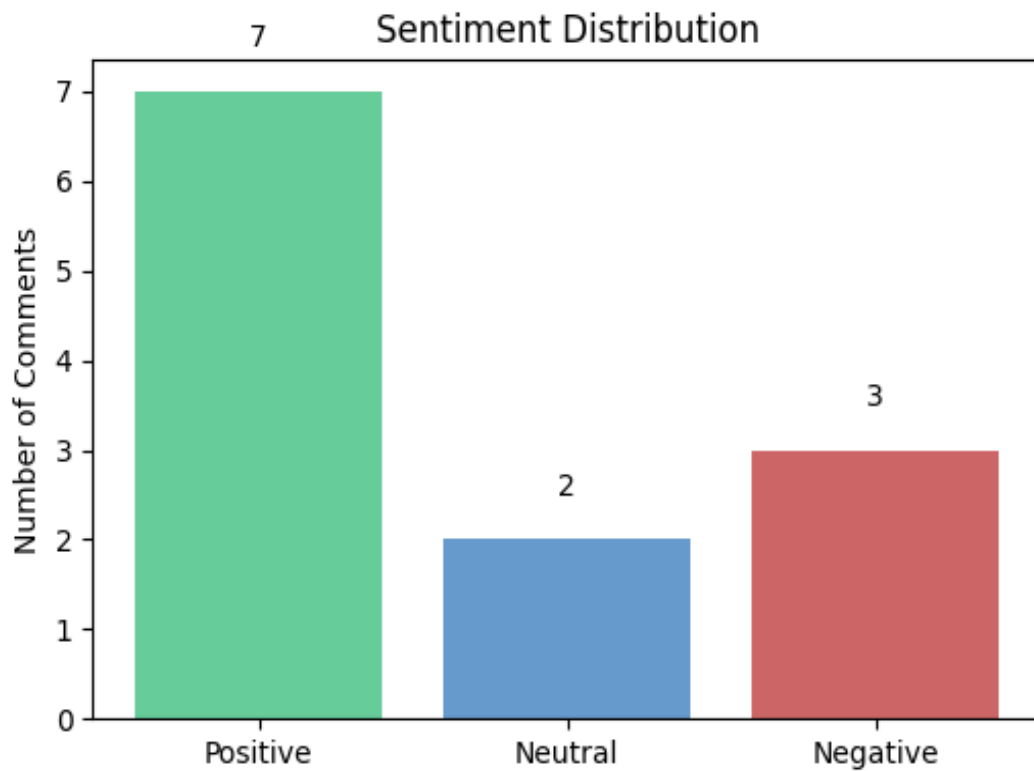
Classification Distribution

Comment Classification Distribution



The chart shows a moderate level of concerning content with 25.0% of comments classified as cyberbullying. This indicates the need for enhanced monitoring and targeted interventions.

Sentiment Analysis



The sentiment analysis shows a predominantly positive tone, consistent with the low levels of cyberbullying content detected. This suggests a generally healthy communication environment.

Word Frequency Analysis

Arya must listen to my wife	Normal	97.11%
Collin Lee ILOVEYOUUU, honeyyy	Normal	92.87%

Recommendations

Moderate cyberbullying content detected (25.0%)

1. **Enhanced Monitoring:** Increase the frequency of comment section review to identify and address problematic content quickly.
2. **Targeted Moderation:** Focus moderation efforts on:
 - Comments containing the most frequent harmful terms identified in the word cloud
 - Threads or discussions that may be escalating in negativity
3. **Community Engagement:** Actively engage with the community to promote positive interactions and discourage negative behavior.
4. **Clear Guidelines:** Ensure your community guidelines are visible and clearly communicate expectations for respectful communication.
5. **Follow-up Analysis:** Conduct a follow-up analysis within 2-4 weeks to monitor for improvement or deterioration.

Technical Notes

- This report was generated using an AI-powered cyberbullying detection system
- Classification confidence is based on the prediction model's probability score
- Word clouds exclude common stop words in both English and Tagalog
- Sentiment analysis utilizes a specialized model designed for social media content
- The system may not catch all instances of cyberbullying, particularly those using coded language or novel terms

For additional support or questions about this report, please contact the system administrator.

Report generated on April 10, 2025 at 02:12 AM