Abstract

This project addresses the growing challenge of cyberbullying on social media by proposing a scalable, tool-driven framework designed to automate content moderation. Traditional approaches often struggle with the sheer volume and dynamic nature of social media data, prompting the need for a solution that combines efficiency, adaptability, and ease of deployment. The developed framework tackles these issues by integrating Apache Spark and Docker to create a robust pipeline capable of processing large-scale datasets while maintaining consistency across environments.

At its core, the system processes data from eight diverse social media datasets stored in HDFS, using Apache Spark to ingest and clean text efficiently. Spark’s distributed computing power ensures rapid handling of missing values, noise, and inconsistencies, even with terabytes of data. From there, raw text is transformed into actionable insights through natural language processing techniques embedded in Spark’s MLlib. Tokenization, HashingTF, and TF-IDF convert unstructured posts into structured features, enabling the framework to identify patterns indicative of aggressive behavior. A logistic regression model trained on this data achieves an AUC of 0.845, demonstrating strong performance in distinguishing harmful content from benign interactions.

Deployment is streamlined through Docker, which packages the entire workflow into portable containers, eliminating environment-specific conflicts and simplifying collaboration. GitHub Workflows automate testing and updates, while cluster management tools optimize memory allocation in Spark and secure communication via SSH authentication. By prioritizing scalability and reproducibility, this open-source framework not only addresses the immediate challenge of cyberbullying detection but also provides a flexible foundation for future enhancements, such as integrating advanced machine learning models or expanding to new platforms. The result is a practical, tool-centric solution that bridges the gap between experimental research and real-world implementation in content moderation.