

I. INTRODUCTION

Cyberbullying is an increasing form of online bullying that has a negative impact on mental health, especially among adolescents. As of April 2024, Indonesia ranked fourth in the world with 24.69 million active users of the social media platform X (formerly known as Twitter) [1]. This platform allows users to share content, send messages, and interact through comments. However, it is often misused for acts of cyberbullying, such as insults, threats, or spreading slander.

Cyberbullying is defined as an act of intimidation or bullying committed online through social media, chat, email, or websites by spreading slander, insulting, or revealing one's personal secrets [2]. To overcome this problem, preventive measures are needed, one of which is developing a cyberbullying detection system. What makes this research different from other research is the implementation of a hybrid CNN-RNN approach combined with TF-IDF and FastText methods on Indonesian datasets.

Despite the wide use of social media, Indonesian text presents unique linguistic challenges, such as the use of slang, abbreviations, and mixed languages, making it difficult for traditional statistical methods alone to capture the true intent behind the words. Semantic integration is crucial because cyberbullying often relies on implicit meanings, wordplay, and contextual understanding, which cannot be adequately addressed by statistical methods like TF-IDF alone.

The novelty of this research is the combination of TF-IDF and FastText, where FastText is used for extraction and expansion features applied to Indonesian text, integrated with a hybrid CNN-RNN model. TF-IDF is used to weight words based on their frequency of occurrence [3], while FastText is used to capture contextual meaning and semantic relationships between words [4]. In addition, FastText enriches the feature representation through word expansion to handle different word variations and forms in text. The purpose of this research is to create a mechanism that can identify cyberbullying by combining TF-IDF and FastText methods in a CNN-RNN model.

The combination of TF-IDF and FastText is expected to overcome the limitations of feature representation, such as TF-IDF's inability to capture context, while improving the performance of the cyberbullying detection system. Word embedding is widely used in feature extraction like TF-IDF, but to our knowledge, no one has used FastText as a feature extraction method in cyberbullying research or combined the two. This research is designed to improve upon previous research by integrating a CNN-RNN hybrid model, which utilizes the power of CNN in capturing spatial features from text and the ability of RNN to understand sequence and temporal context