

## **Preface**

Online communication has become a crucial aspect of our daily lives in today's linked world. We may voice our opinions, participate in deep conversations, and connect with a wide spectrum of people through social media platforms, forums, and comment sections. This digital environment is not without its difficulties, though.

Online toxicity is on the rise, causing risks to people's wellbeing, preventing open discourse, and promoting a toxic and polarizing online environment. Comments that are vile, nasty, and disparaging have the capacity to cause emotional injury, support discrimination, and foster a hostile environment. For online interactions to be secure, inclusive, and of high caliber, this issue must be solved.

This study focuses on the classification of comment toxicity using deep learning methods in response to these difficulties. The goal is to create a smart system that can recognize toxic remarks and classify them depending on their severity, encompassing elements like obscenity, insults, identity hate, and more. By reliably identifying toxic remarks, we hope to equip users, content moderators, and platform administrators with the means to lessen harm, encourage fruitful debate, and advance a more positive online community.

We want to overcome the complexity and nuances of human language by combining powerful natural language processing and deep learning techniques, enabling the system to comprehend context, recognize subtle indications, and differentiate between harmful and non-damaging comments. The Long Short-Term Memory (LSTM) architecture, which is well-known for its capacity to record temporal dependencies, can be used to describe and analyze sequential data.

This study emphasizes the ethical ramifications and societal effects of encouraging responsible digital interactions in addition to addressing the technical difficulties of comment toxicity categorization. With the completion of this project, we intend to stimulate additional study, cooperation, and innovation in the area of comment toxicity classification, ultimately fostering a digital environment that promotes respect, empathy, and productive debate.