**Model Development Phase Template**

|  |  |
| --- | --- |
| Date | 15 October 2024 |
| Team ID | LTVIP2024TMID24818 |
| Project Title | Toxic Comments Classification For Social Media |
| Maximum Marks | 5 Marks |

**Feature Selection Report Template**

In the forthcoming update, each feature will be accompanied by a brief description. Users will indicate whether it's selected or not, providing reasoning for their decision. This process will streamline decision-making and enhance transparency in feature selection.

|  |  |  |  |
| --- | --- | --- | --- |
| **Feature** | **Description** | **Selected (Yes/No)** | **Reasoning** |
| Word Count | Total number of words in the comment | Yes | Indicates comment length; correlates with toxicity levels |
| Character Count | Total number of characters in the comment | Yes | Similar to word count, also suggests intensity |
| Exclamation Count | Exclamation Count | Yes | Higher counts suggest heightened emotion |
| Toxicity Lexicon Count | Count of words from a toxicity lexicon | Yes | Directly indicates potential toxicity. |
| Negation Count | Number of negation words (e.g., "not") | Yes | Strong correlation with negative sentiment |
| Special Characters | Count of special characters (e.g., !, ?) | No | Minimal impact on classification accuracy. |
| URL Count | Number of URLs present in the comment | No | Does not significantly impact toxicity detection |