**Model Development Phase Template**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | LTVIP2024TMID24955 |
| Project Title | SMS Spam Detection - AIML |
| 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** |
| Message Length | The total number of characters in the SMS message. | Yes | Spam messages often have distinct length patterns (e.g., longer promotional or short urgent texts). |
| Presence of Keywords | Whether specific keywords (e.g., "free," "win," "prize," "click") appear. | Yes | Spam messages frequently include enticing or misleading keywords. |
| Number of Links | The count of URLs or links within the message. | Yes | Spam messages often contain links leading to phishing sites or offers.  Special Characters |
| Special Characters | Presence of special characters (e.g., "$", "%", "!", "@"). | Yes | Spam messages may contain special characters to grab attention or bypass filters. |
| Sender Information | Metadata about the sender (e.g., phone number, unknown sender, short code). | No | Not always reliable as spammers can spoof phone numbers, and metadata may not be consistent. |
| Capitalization | Percentage of words in all caps (e.g., "FREE," "URGENT"). | Yes | Spam messages frequently use capitalization to emphasize key points or urgency. |