**Model Development Phase Template**

|  |  |
| --- | --- |
| Date | 28 September 2024 |
| Team ID | LTVIP2024TMID24973 |
| Project Title | Detection of Phishing Websites from URLs  Using Machine learning |
| 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** |
| URL Length | The total number of characters in the URL | Yes | Longer URLs are often indicative of phishing attempts, as they may contain unnecessary parameters. |
| Use of HTTPS | Whether the URL uses HTTPS or not | Yes | Legitimate websites usually employ HTTPS, while many phishing sites do not. |
| **Number of Subdomains** | **Count of subdomains present in the URL** | **Yes** | **Phishing sites often use multiple subdomains to mimic legitimate sites, making this a useful feature.** |
| **Use of IP Address** | **Whether the URL contains an IP address instead of a domain name** | **Yes** | **URLs that use IP addresses instead of domain names are often suspicious and associated with phishing.** |
| **Presence of Special Characters** | **Count of special characters (e.g., “@”, “-”, “\_”)** | **Yes** | **High counts of special characters may indicate attempts to obfuscate the URL, common in phishing attempts.** |