**Web Page phishing**

Data Format

The data is provided in CSV format, with each row representing a website and each column representing a feature. The last column contains the label for each website.

Features

This dataset contains the following features:

url\_length: The length of the URL.

n\_dots: The count of ‘.’ characters in the URL.

n\_hypens: The count of ‘-’ characters in the URL.

n\_underline: The count of ‘\_’ characters in the URL.

n\_slash: The count of ‘/’ characters in the URL.

n\_questionmark: The count of ‘?’ characters in the URL.

n\_equal: The count of ‘=’ characters in the URL.

n\_at: The count of ‘@’ characters in the URL.

n\_and: The count of ‘&’ characters in the URL.

n\_exclamation: The count of ‘!’ characters in the URL.

n\_space: The count of ’ ’ characters in the URL.

n\_tilde: The count of ‘~’ characters in the URL.

n\_comma: The count of ‘,’ characters in the URL.

n\_plus: The count of ‘+’ characters in the URL.

n\_asterisk: The count of ‘\*’ characters in the URL.

n\_hastag: The count of ‘#’ characters in the URL.

n\_dollar: The count of ‘$’ characters in the URL.

n\_percent: The count of ‘%’ characters in the URL.

n\_redirection: The count of redirections in the URL.

phishing: The Labels of the URL. 0 is phishing and 1 is legitimate.

**Output Column:**

Phishing