|  |  |  |
| --- | --- | --- |
| **Categories** | | **Description** |
| **General Features** | 1.use\_of\_ip | If the URL contains IP address |
| 2.is\_https | If the URL uses 'http' or 'https' |
| 3.is\_redirection | Find how many '//' in the URL |
| 4.have\_- | If the URL contents '-' symbol |
| 5.have\_@ | If the URL contents '@' symbol |
| 6.dir\_depth | Get the depth of the URL |
| **URL Component Lengths** | 7.url\_len | The length of the whole URL |
| 8.subdomian\_len | The length of the URL's subdomain |
| 9.tld\_len | The length of the top-level domain |
| 10.fld\_len | The length of the domain and TLD |
| 11.url\_path\_len | The length of the path |
| 12.hostname\_len | The length of the URL’s hostname |
| **Count Features** | 13.url\_alphas | A count of the alpha/characters within the URL |
| 14.url\_digits | A count of the digit/numbers within the URL |
| 15.url\_puncs | A count of the punctuation within the whole URL |
| 16.count’.’ | A count of the number of ’.’ within the URL |
| 17.count’@’ | A count of the number of ‘@’ within the URL |
| 18.count’-’ | A count of the number of ‘-’ within the URL |
| 19.count’%’ | A count of the number of ‘%’ within the URL |
| 20.count’?’ | A count of the number of ‘?’ within the URL |
| 21.count’=’ | A count of the number of ‘=’ within the URL |
| 22.count\_dirs | A count of the number of directories within *url\_path* |
| 23.first\_dir\_len | A count of the length of the first directory within *url\_path* |
| **Ratio Features** | 24.pc\_alphas | The percentage of letters in the URL |
| 25.pc\_digits | The percentage of numbers in the URL |
| 26.pc\_puncs | The percentage of punctuations in the URL |
| **Binary Label** | 27.result | A binary label where 0 denotes benign and 1 denotes malicious |
| **Domain Features** | 28.domain\_Age | Calculate the time period the domain lives |
| 29.domain\_End | Calculate how soon the domain expires |
| 30.email | Get the email information from domain information |
| 31.google\_index | If the URL is in Google's index or not |
| 32.web\_traffic | This feature measures the popularity of the website by Alexa |