**Table S6.** Feature for both datasets with the highest scores based on TU and PBCC

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dataset 1 | | | Dataset 2 | |
| # | Features | Score | Features | Score |
| 1 | NumDash | 0.3722 | SSLfinal\_State | 0.715 |
| 2 | PctNullSelfRedirectHyperlinks | 0.3428 | URL\_of\_Anchor | 0.693 |
| 3 | PctExtNullSelfRedirectHyperlinksRT | 0.3064 | Prefix\_Suffix | 0.349 |
| 4 | NumDots | 0.2941 | web\_traffic | 0.346 |
| 5 | PctExtHyperlinks | 0.2597 | having\_Sub\_Domain | 0.298 |
| 6 | NumSensitiveWords | 0.2552 | Request\_URL | 0.253 |
| 7 | PathLevel | 0.2295 | Links\_in\_tags | 0.248 |
| 8 | FrequentDomainNameMismatch | 0.1810 | SFH | 0.221 |
| 9 | ExtMetaScriptLinkRT | 0.1696 | Google\_Index | 0.129 |
| 10 | HostnameLength | 0.1692 | age\_of\_domain | 0.121 |
| 11 | NumDashInHostname | 0.1504 | Page\_Rank | 0.105 |
| 12 | NumQueryComponents | 0.1474 |  |  |
| 13 | SubmitInfoToEmail | 0.1126 |  |  |