Introduction

The basic idea of internet connection is as follows:

a. A connection is initiated from the port of a machine (the source) to a port on another or the same machine (the destination).
b. Information is transmitted along this connection in form of ***packets***. These packets may vary in size.
c. Sometimes, the packets may be very large, in which case they are broken down into ***fragments*** or ***segments*** via a process known as ***fragmentation***.
d. Packets may be sent either source-to-destination (i.e., forward direction) or destination-to-source (i.e., backward direction).
e. Among these packets, there is usually a special one that describes and allows for authentication of all the other packets. This special packet is known as the ***header***.
f. During or at the end of packet transmission in either the forward or backward direction, the status of the connection or packet receipt is designated via a set of ***flags***. They include: *FIN, PSH, WAIT, SYN, REST, and ACK*. A flag might be raised depending on the status of the packets or connection.

Features

The features of interest in the dataset are outlined below:

1. Src_Port, - Originating port for connection
2. Dst_Port, - Destination port for connection
3. Protocol, - Connection protocol (HTTP, HTTPS, UDP, TCP etcetera)
4. Flow_Duration, - Duration of connection
5. Tot_Fwd_Pkts, - Total number of packets transmitted forward over connection lifetime
6. Tot_Bwd_Pkts, - Total number of packets transmitted backward over connection lifetime
7. TotLen_Fwd_Pkts, - Total size of packets transmitted forward over connection lifetime
8. TotLen_Bwd_Pkts, - Total size of packets transmitted backward over connection lifetime
9. Fwd_Pkt_Len_Max, - Size of largest packet transmitted forward over connection lifetime
10. Fwd_Pkt_Len_Min, - Size of smallest packet transmitted forward over connection lifetime
11. Fwd_Pkt_Len_Mean, - Mean size of packets transmitted forward over connection lifetime
12. Fwd_Pkt_Len_Std, - Standard deviation of size of packets transmitted forward over connection lifetime
13. Bwd_Pkt_Len_Max, - Size of largest packet transmitted backward over connection lifetime
14. Bwd_Pkt_Len_Min, - Size of smallest packet transmitted backward over connection lifetime
15. Bwd_Pkt_Len_Mean, - Mean size of packets transmitted backward over connection lifetime
16. Bwd_Pkt_Len_Std, - Standard deviation of size of packets transmitted backward over connection lifetime
17. Flow_Byts/s, - Overall connection speed in bytes per second
18. Flow_Pkts/s, - Overall connection speed in number of packets per second
19. Flow_IAT_Mean,
20. Flow_IAT_Std,
21. Flow_IAT_Max,
22. Flow_IAT_Min,
23. Fwd_IAT_Tot,
24. Fwd_IAT_Mean,
25. Fwd_IAT_Std,
26. Fwd_IAT_Max,
27. Fwd_IAT_Min,
28. Bwd_IAT_Tot,
29. Bwd_IAT_Mean,
30. Bwd_IAT_Std,

31. Bwd_IAT_Max,
32. Bwd_IAT_Min,
33. Bwd_PSH_Flags, - PSH flag raised during backward transmission?
34. Fwd_Header_Len, - Size of header for forward packets.
35. Bwd_Header_Len, - Size of header for backward packets.
36. Fwd_Pkts/s, - Forward connection speed in number of packets per second
37. Bwd_Pkts/s, - Backward connection speed in number of packets per second
38. Pkt_Len_Min, - Size of smallest packet transmitted over connection lifetime
39. Pkt_Len_Max, - Size of largest packet transmitted over connection lifetime
40. Pkt_Len_Mean, - Mean packet size over connection lifetime
41. Pkt_Len_Std, - Standard deviation of packet sizes over connection lifetime
42. Pkt_Len_Var, - Variance of packet sizes over connection lifetime
43. FIN_Flag_Cnt, - Number of *FIN* flags raised over connection lifetime.
44. SYN_Flag_Cnt, - Number of *SYN* flags raised over connection lifetime.
45. RST_Flag_Cnt, - Number of *RST* flags raised over connection lifetime.
46. PSH_Flag_Cnt, - Number of *PSH* flags raised over connection lifetime.
47. ACK_Flag_Cnt, - Number of *ACK* flags raised over connection lifetime.
48. Down/Up_Ratio, - Ratio of
49. Pkt_Size_Avg, - Average packet size over entire connection lifetime
50. Fwd_Seg_Size_Avg, - Average size of forward packet segments over connection lifetime
51. Bwd_Seg_Size_Avg, - Average size of backward packet segments over connection lifetime
52. Subflow_Fwd_Pkts,
53. Subflow_Fwd_Byts,
54. Subflow_Bwd_Pkts,
55. Subflow_Bwd_Byts,
56. Init_Bwd_Win_Byts,
57. Fwd_Act_Data_Pkts,
58. Active_Mean, - Average packet/connection active time
59. Active_Std, - Standard deviation of packet/connection active time
60. Active_Max, - Maximum packet/connection active time
61. Active_Min, - Minimum packet/connection active time
62. Idle_Mean, - Average packet/connection idle time
63. Idle_Std, - Standard deviation of packet/connection idle time
64. Idle_Max, - Maximum packet/connection idle time
65. Idle_Min, - Minimum packet/connection idle time