**CSE523 Machine Learning**

**Prof. Mehul Raval**

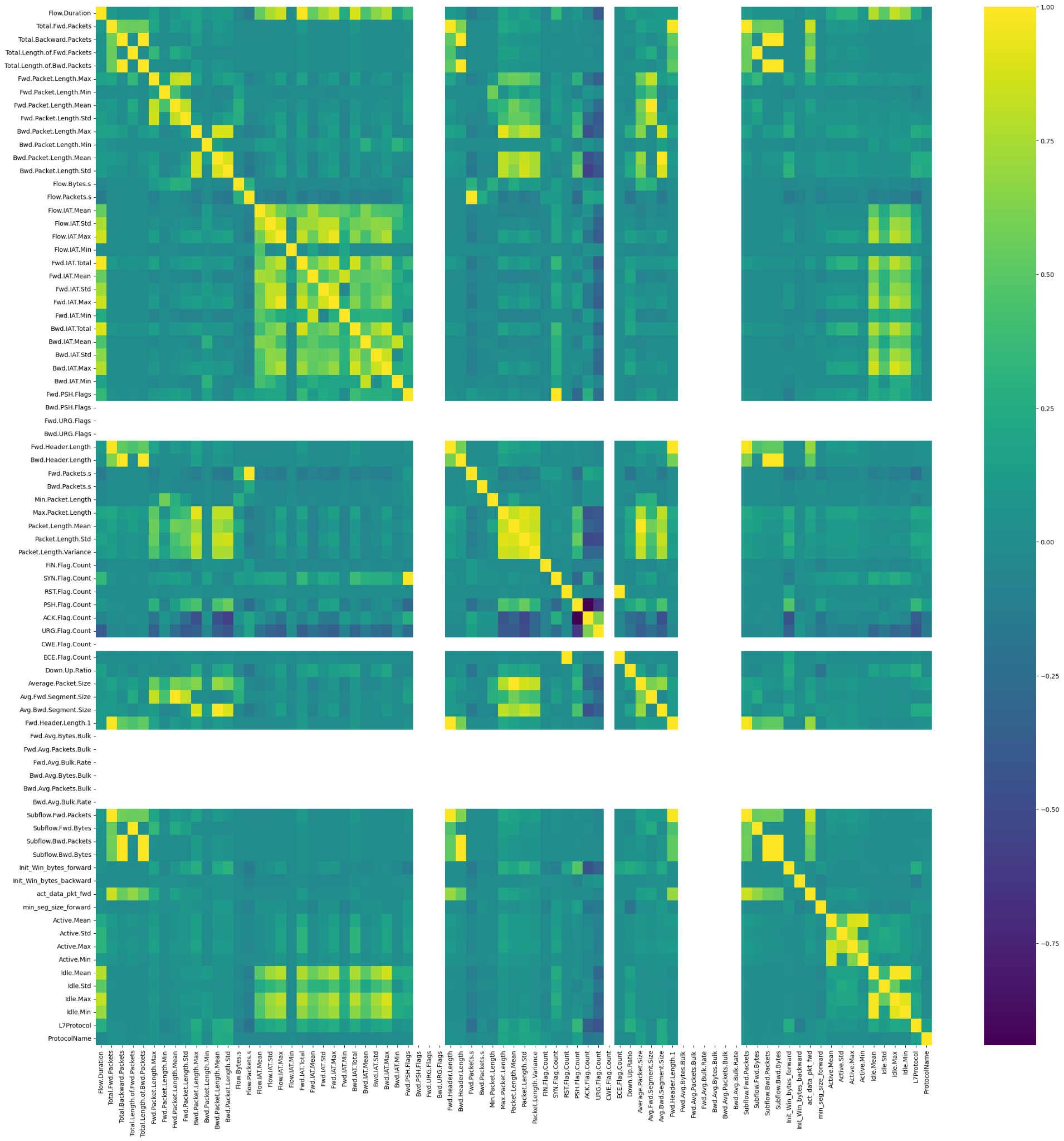
**Analyzing TCP Network Packets**

**Weekly Report**

| **Name** | **Enrolment Number** |
| --- | --- |
| Nityam Dixit | AU2040140 |
| Tej Thakar | AU2040262 |
| Yagnesh Patel | AU2040025 |
| Kavan Gondalia | AU2040030 |

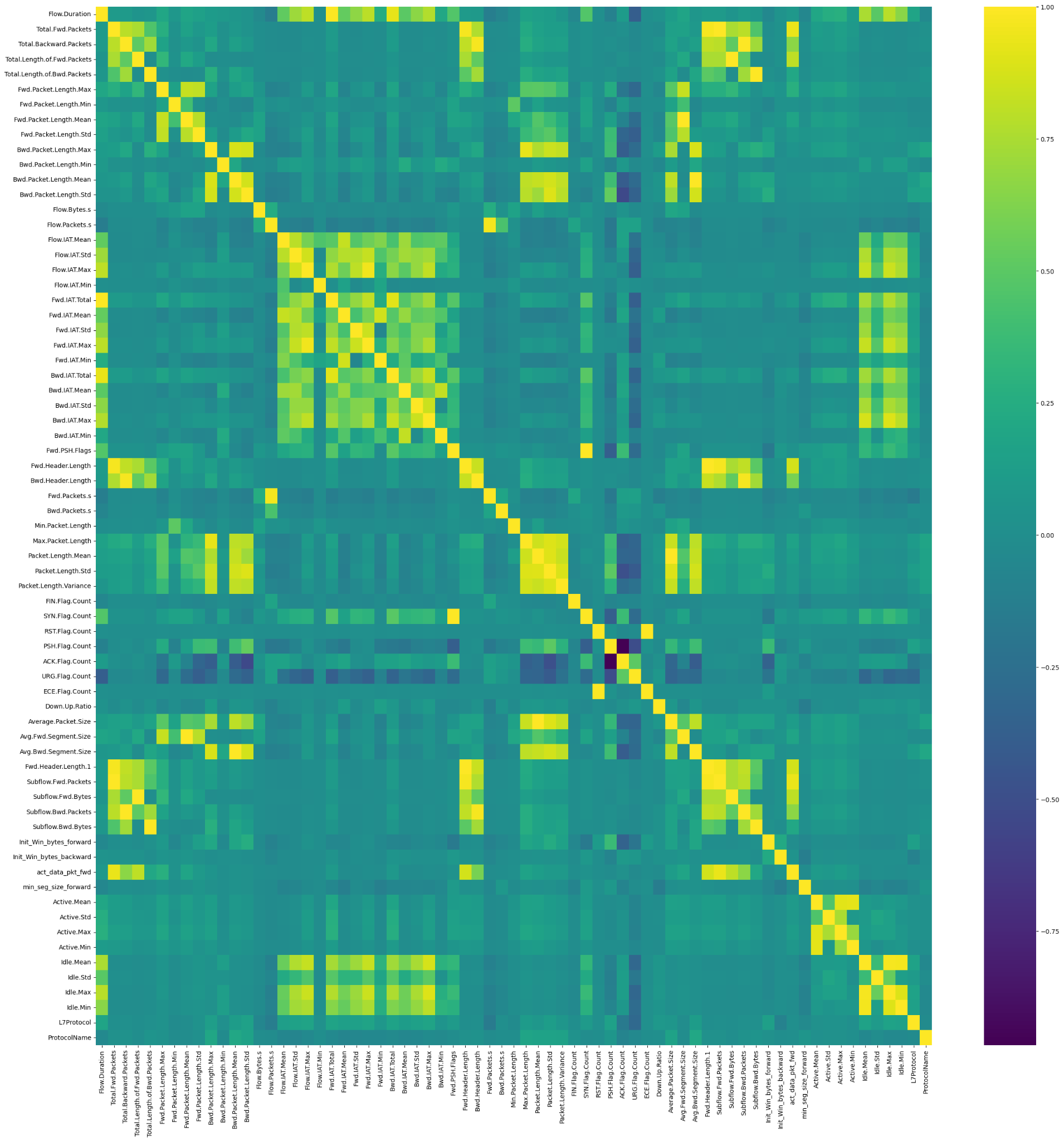
Kaggle Link: <https://www.kaggle.com/code/kavangondalia/predict-protocol-name>

* Previously the dataset contains zero values and resulted in the following correlation matrix



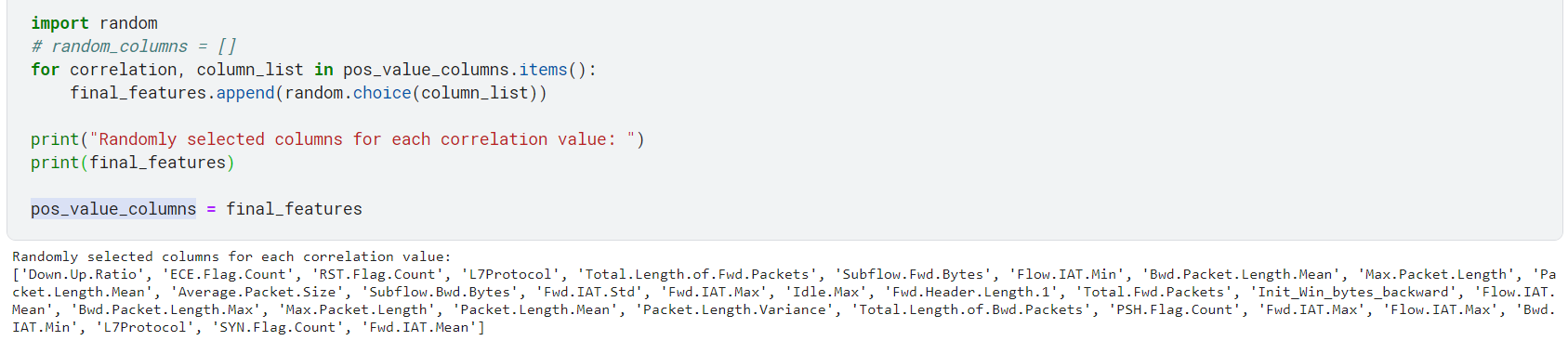
* After removing the columns/features which do not contribute to the prediction (the ones which have zero values for each value)



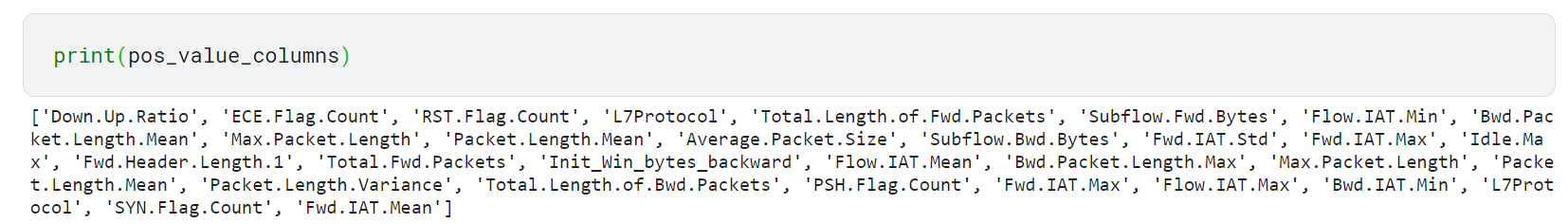


Now we map the features having the same correlation values and after mapping them we are selecting one feature for each correlation value at random.





Final\_features:



This is the accuracy we got by applying random forest on the Final features :

