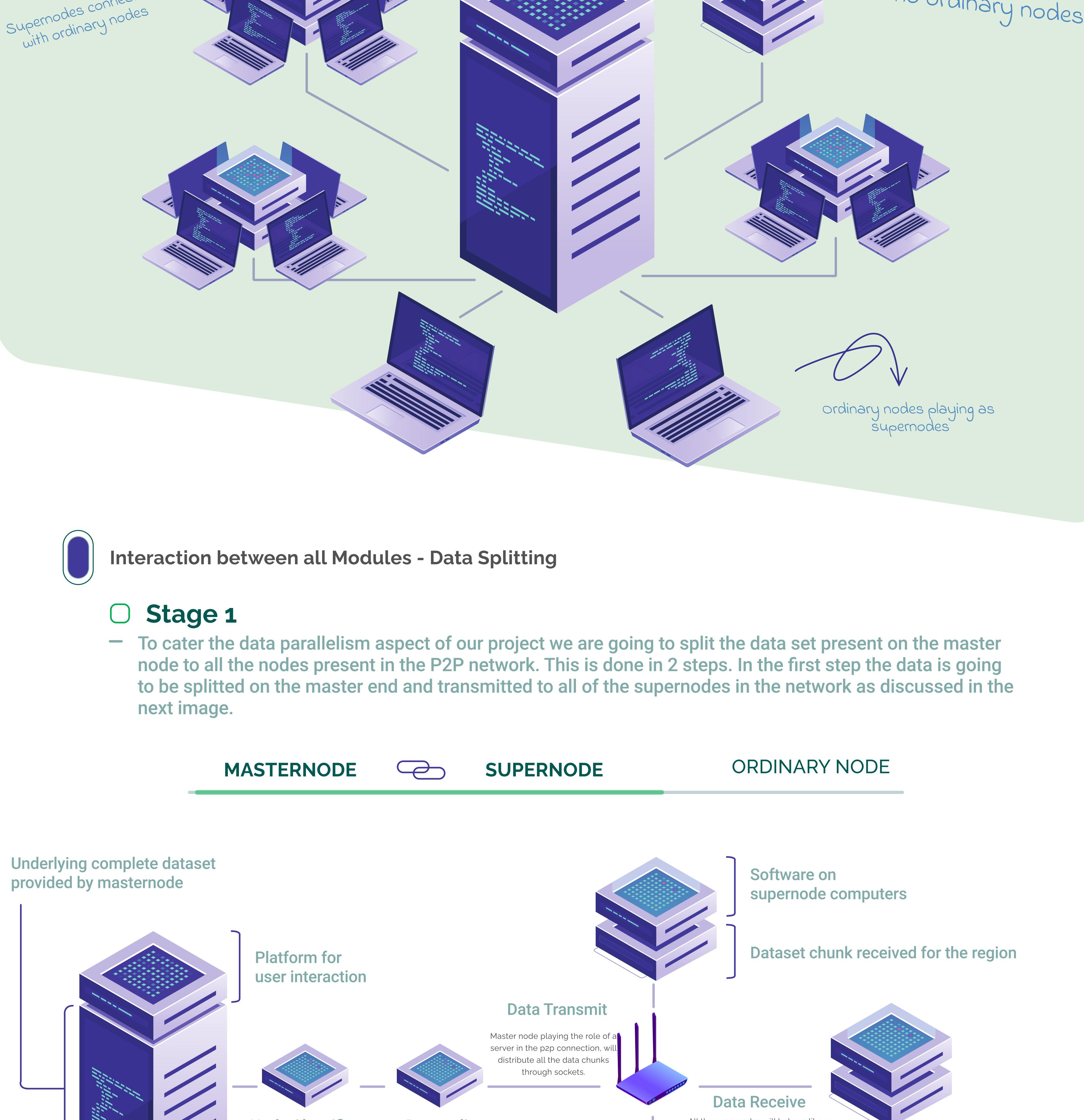


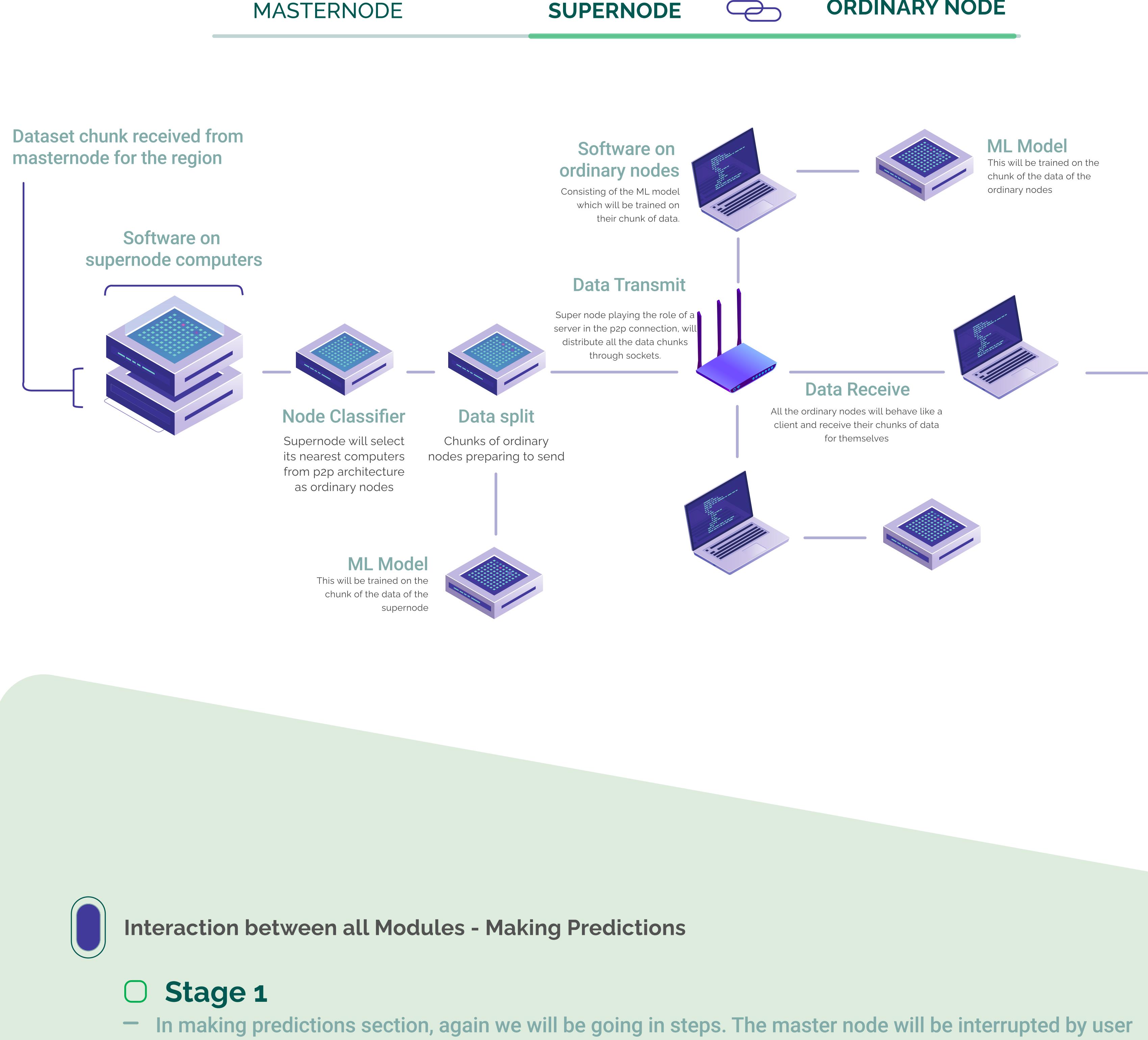
Overview of Whole System



Interaction between all Modules - Data Splitting

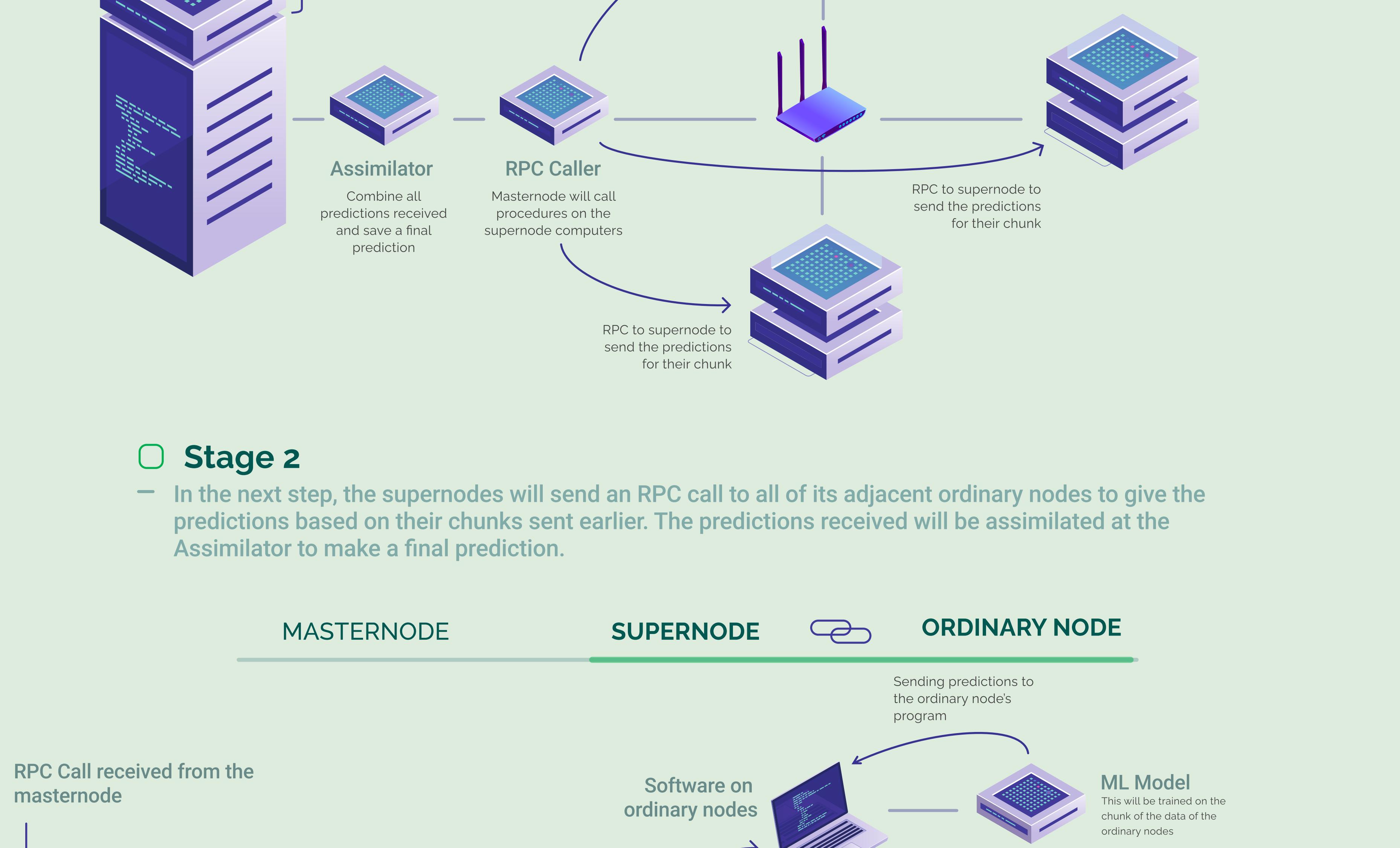
Stage 1

- To cater the data parallelism aspect of our project we are going to split the data set present on the master node to all the nodes present in the P2P network. This is done in 2 steps. In the first step the data is going to be splitted on the master end and transmitted to all of the supernodes in the network as discussed in the next image.



Stage 2

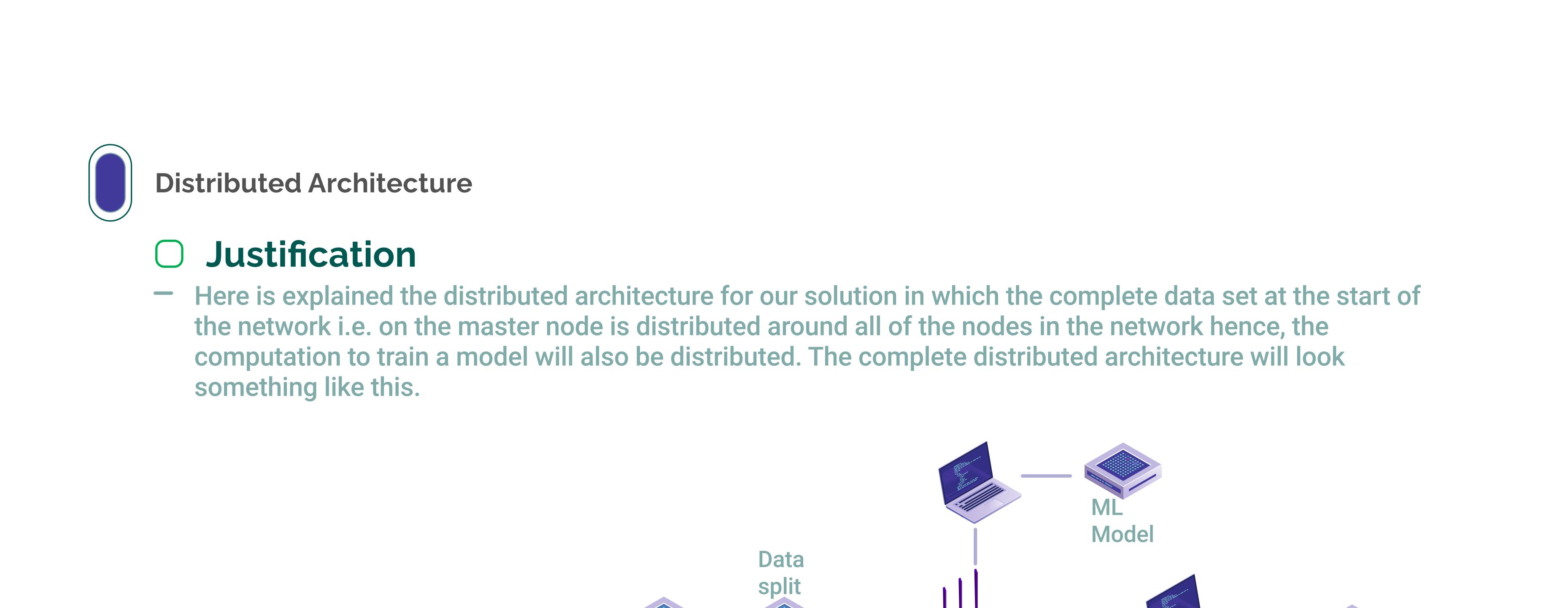
- In the next stage the data chunk collected by the supernode computer will get splitted into parts and be ready for distribution. The supernode will keep his own data chunk while transmit the chunks of all the ordinary nodes to their respective softwares which will then run ML model on that chunk.



Interaction between all Modules - Making Predictions

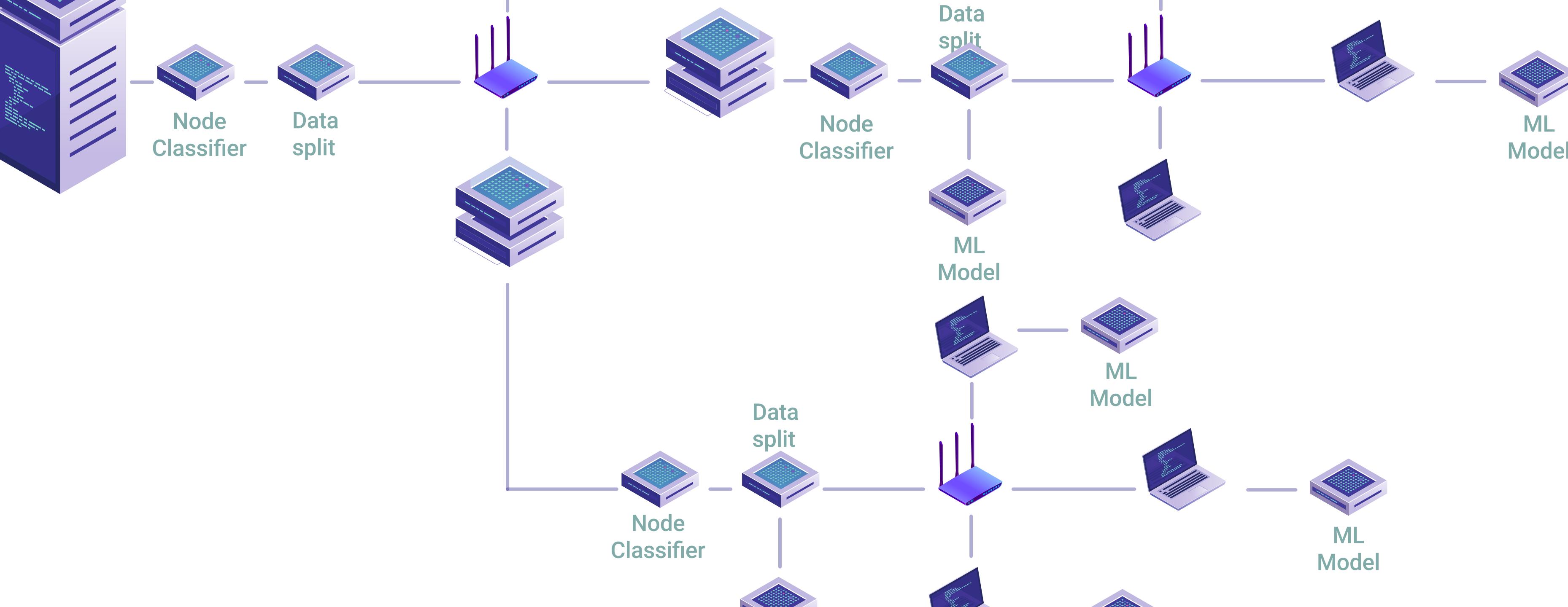
Stage 1

- In making predictions section, again we will be going in steps. The master node will be interrupted by user action to give predictions. Resultantly, in the first step, the master node will send an RPC call to all of its adjacent supernodes to give the predictions based on their chunks sent earlier. These predictions will be assimilated at the Assimilator.



Stage 2

- In the next step, the supernodes will send an RPC call to all of its adjacent ordinary nodes to give the predictions based on their chunks sent earlier. The predictions received will be assimilated at the Assimilator to make a final prediction.



Distributed Architecture

Justification

- Here is explained the distributed architecture for our solution in which the complete data set at the start of the network i.e. on the master node is distributed around all of the nodes in the network hence, the computation to train a model will also be distributed. The complete distributed architecture will look something like this.

