* Hyperparameter training - using Bayesian optimization
* Evaluate the model - StratifiedKFold - n\_splits=10. Random seed - 7
  + No overfitting
* Train the model (train test split 0.2, shuffle true, random seed - 7)

All selected features were transformed using MinMaxScaler and brought into the range of 0 to 1, without distorting the shape of the original features. Transformed features and the selected model were used to tune the hyperparameters of the model using Bayesian Optimization, which outputs the best-optimized list of parameters while learning from previous iterations in each iteration. The tuned model was evaluated using StratifiedKFold cross validation, which represents each fold as a representation of the whole dataset in the evaluation process. The results generated by the above-mentioned cross validation confirms the fact that the model is less biased, performing well in unseen data and not overfitting. Based on the results of the evaluation process, the model is trained using 80% of the shuffled dataset and validated against the rest. The above process is repeated for all 16 routers separately and 16 separate models were trained, one per each router.

Cross Validation

Cross validation is a resampling process to evaluate the performance of the machine learning model.

KFold Cross Validation shuffles the dataset and splits it into k subsets, then trains on k-1 and evaluates on the other set iteratively (till each subset is used as the test set.)

StratifiedKFold Cross Validation shuffles the dataset and splits in into k subsets by class and use a subset from each in the test set emulating a representation of the entire dataset in each fold both for training and validation.