About the data:

Data is a txt file containing 7 columns and 24724 rows representing human presence/absence in a room in the binary(1/0) format for a timeframe of 18 minutes.

Col1: Timestamp
Col2: Stream from Node1(0/1)
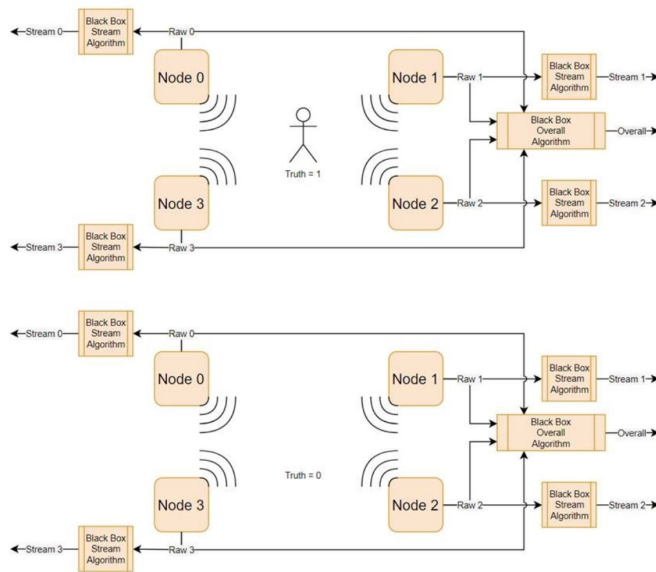Col3: Stream from Node2(0/1)
Col4: Stream from Node3(0/1)
Col5: Stream from Node4(0/1)
Col6: Overall stream generated by the black box
Col7: Truth Value of human presence/absence

The Stream columns show decisions based on the data from a single data stream. The Overall column shows the output of a black box non-ml algorithm based on past/current stream data. The Truth column shows when someone was actually present.

Overview of the notebook (Submission.ipynb):

- Data exploration and analysis of stream data from 4 different sensor nodes to infer something insightful and make some conjectures about the black-box algorithm.
- Attempt to generate my own detection output that approaches Truth.
- Time series prediction using LSTM.
- Time series classification using Random forest and Logistic Regression.