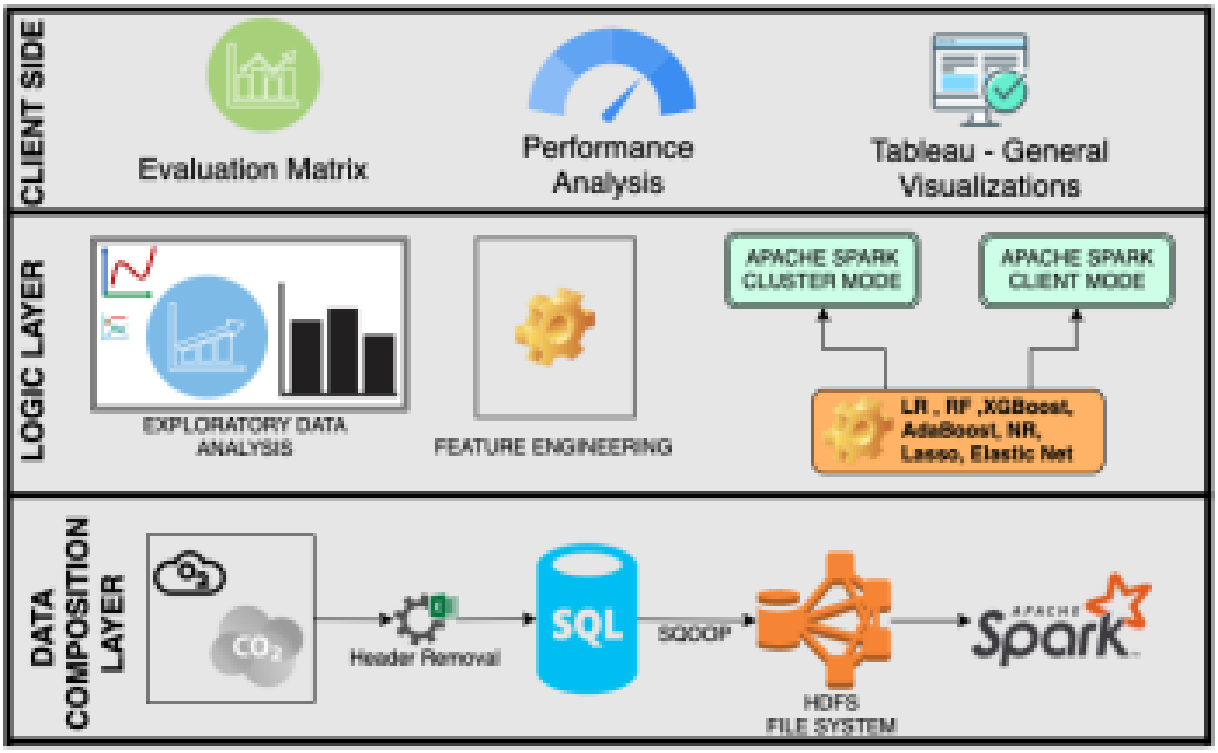
PROJECT: AIR QUALITY ANALYSIS IN TN.

ARCHITECTURE DIAGRAM



The process flow architecture was designed for the air quality in Tamil nadu to increase the robustness and accuracy of the research. The architecture has a three tier structure. The first layer in the architecture is data composition layer where the data is imported and pre-processed followed by the logic layer which acts as the core of the entire project and finally the third layer called as the client side is where the results obtained from the implementation part is presented and discussed using various tools. The main goal of the research is to predict the AQI with respect to the concentration of PM2.5 in the atmosphere of Tamil naduwhich is only possible with a strong architecture design.

KDD methodology, the traditional design is modified according to the steps followed in this research . KDD is one of the methodology used in the area of machine learning to develop different models. Data mining is a part of the KDD process. Following a clearly modified methodology approach helps in building a robust model by proper implementation of various stages at the proper time, The model can only be developed with prior understanding of the domain and client goals. The modified methodology consist of five stages. The methodology is structured in such a way that each stage is given equal importance, starting from data collection which is the first stage of the research, second is the data pre-processing, third data transformation which includes feature engineering, fourth the implementation stage and finally the result