**Description of Dataset and Problem Statement**

**Occupancy Detection Data Set**: This dataset is an experimental data used for binary classification (room occupancy) from Temperature, Humidity, Light and CO2. Ground-truth occupancy was obtained from time stamped pictures that were taken every minute.

**date**: Contains time stamp as date time year-month-day hour:minute:second.

**Temperature**: Room temperature in terms of degree Celsius.

**Humidity**:Relative Humidity in %.

**Light**: Light in the room measured in Lux.

**CO2**: Carbon dioxide content, measured in ppm.

**HumidityRatio**: Derived quantity from temperature and relative humidity, in kgwater-vapor/kg-air.

**Occupancy**: Categorical data, 0 for not occupied, 1 for occupied status.

--------------------------------------------------------------------------------------------------------------------

Perform the data cleaning if required. Perform the descriptive analytics to understand the data and infer from the data.

Perform the predictive analytics (classification) on predicting the whether the room is occupied or not. Use different classification techniques and compare the results. Perform the data pre-processing (normalization, standardization, correlation analysis & feature section, dimension reduction using PCA) and compare the results of classification with unprocessed data.

Consider 70% of data from each class for training and remaining 30% of data from each class for testing.

Infer the results obtained from descriptive and predictive analytics.