Enhancing Indoor Air Quality Assessment in Mexico with Big Data & ML

Ariel Isaac Posada Barrera,
Laura Margarita Rodríguez Peralta,
arielisaac.posada@upaep.edu.mx
lauramargarita.rodriguez01@upaep.mx

Abstract

In this study spanning 2022 and 2023, we developed predictive models to forecast indoor air pollutant levels (CO2, TVOC, PM2.5, PM10) using environmental variables from Puebla and Morelos, Mexico, including temperature, humidity, occupancy, and ventilation. Employing Machine Learning Models—Regression Forest and Gradient-Boosted Tree—with the Gradient-Boosted Tree model showing the best performance, and statistical analyses (Levene test, ANOVA, correlation), we identified significant influences on air quality, aiming to mitigate respiratory diseases (like COVID-19) and Sick Building Syndrome (SBS). Our analyses confirmed that the ideal temperature for indoor environments is between 18 and 28 °C, with a relative humidity range of 30 to 50%, to maintain optimal air quality. The research underscores the importance of these optimal environmental conditions and continuous monitoring to prevent poor indoor air quality, following Environmental Protection Agency guidelines. These findings offer a novel approach for policy-making and indoor air quality standards, demonstrating the feasibility of predicting hazardous levels of indoor pollutants to enhance public health.

Keywords

Big Data, COVID-19, Sick Building Syndrome, Data Analysis. Data Cleansing, Information Technologies. Monitoring Variables