# Indoor Air Quality Prediction and Automation System - General Summary

## Purpose and Scope of the Study

This study aims to predict indoor air quality in smart home automation systems using deep learning techniques and make it accessible to user control. The project seeks to address issues in current automation systems such as sensor malfunctions, short-term changes in user preferences, and lack of foresight against potential disasters.

## Dataset and Methodology

The study utilized the CN-OBEE dataset collected from a residential unit in Beijing. This dataset contains minute-by-minute recorded data from May 31, 2021, to May 31, 2022. Temperature, relative humidity, pressure, window status, and occupancy information were collected from six different rooms (cloakroom, home office, kitchen, living room, master bedroom, and secondary bedroom).

During the data preprocessing phase, missing values were filled with median values, categorical data were converted to numerical format, and data were optimized to 10-minute intervals. MinMaxScaler was used for normalization, and a 10-day lookback window with a 1-day prediction delay was applied for time series analysis.

## Model Comparison and Findings

The study compared three different Recurrent Neural Network (RNN) models: GRU (Gated Recurrent Unit), LSTM (Long Short-Term Memory), and BiGRU (Bidirectional GRU).

**Temperature Predictions:** The GRU model outperformed LSTM in all rooms. The average MAE value for GRU was 0.57°C, while it was 0.93°C for LSTM. BiGRU exhibited balanced performance between the two.

**Relative Humidity Predictions:** GRU was more successful in three out of five rooms (average MAE: 5.41%). LSTM showed lower MSE values in some extreme conditions. BiGRU had the highest error rates for this parameter.

**Air Pressure Predictions:** The BiGRU model excelled in this parameter, surpassing both GRU and LSTM. BiGRU achieved the lowest MAE, MSE, and RMSE values.

## Room-Specific Performance Analysis

The study revealed that room occupancy frequency has a significant impact on prediction performance. Rooms with high occupancy frequency (living room and kitchen) exhibited more stable environmental conditions and showed lower error rates. Rooms with low occupancy frequency (secondary bedroom and home office) presented more variable and difficult-to-predict dynamics.

## Air-Smart Controller Interface

The web-based interface developed using the Streamlit library enables users to:

* View 44-day prediction data at 10-minute intervals
* Access temperature, humidity, and pressure information from 5 different rooms
* Learn which modes the automation system will operate in
* Suspend or schedule system operations

This interface serves as a bridge between advanced prediction models and daily user experience.

## Results and Contributions

The study demonstrated that the GRU model is generally the most suitable model for predicting indoor air quality parameters. However, model selection may vary depending on specific requirements, the importance of the parameter to be predicted, and computational constraints.

The project demonstrates the potential of machine learning to improve living spaces while maintaining user autonomy and comfort. Long-term impacts on energy efficiency, occupant comfort, and overall air quality management will be valuable for quantifying the practical benefits of this approach.

## Future Work

The study's limitations include a dataset collected from a single household and a one-year duration. Recommendations for future research:

* Use of larger and more diverse datasets
* Exploration of advanced transformer-based architectures
* Development of customized models based on room characteristics
* Research on hybrid models combining the strengths of GRU, LSTM, and BiGRU architectures

This study has been published in the Journal of Artificial Intelligence and Human Sciences, contributing to advances in automation technology and air quality management.