Deep Learning Based Occupants' Activity Prediction Towards Designing A Smart Building Assistant System

Ankur Sarker[◇], Fan Yao[◇], Haiying Shen[◇], Huiying Zhao[†], Haoran Zhu[◇], Haroon Lone[◇], Laura Barnes[◇], Brad Campbell[◇], and Mitchel Rosen[◇]
University of Virginia[◇] and Beijing Jiaotong University[†]
{as4mz,fy4bc,hs6ms,hz3fr,hl7ck,lb3dp,bradjc,mcr4y}@virginia.edu; 15114197@bjtu.edu.cn

***Abstract*—Nowadays, smart building infrastructures are equipped with hundreds of sensors to detect and identify different activities of occupants to monitor and provide smart solutions toward occupants comfortability. In this paper, we propose a smart building assistance system consisting of different sensors data analysis and deep neural network (DNN)-based prediction model to increase the comfortability of the building occupants rates at the same time. First, we collect a year-long smart building dataset from four different data sources (i.e., sensors, calendar, weather, and survey). Second, we perform extensive feature engineering (i.e., concretization, one-hot encoding, and multiple feature combination) to be used in the prediction models. Third, we propose a support vector regression based prediction model and a hybrid DNN model which consists of several recurrent neural network blocks and a feed-forward DNN block so that different physical features (e.g., lighting, shadding, air quality, temperature, and so on) can be predicted for different building activities (e.g., meeting, lunch, research activities, and so on). Fourth, we conduct extensive experimental studies to evaluate the performance of the proposed prediction models compared to other existing machine learning models in terms of accuracy. We also conduct experimental studies using our collected data to figure out the activity-wise comfortability rates. From the experiments, we find that the proposed smart building assistance system is able to increase occupants' comfortability.**

***Index Terms*—Activity prediction model, Hybrid Deep Neural Network, Smart Building Assistance, LSTM**

I. Introduction

Nowadays, smart-buildings equipped with fully or partially automated systems to control the physical environments are becoming increasingly pervasive. These buildings can potentially provide environmental conditions that meet users' needs and improve the performance, comfort, health, and well-being of the occupants in particular. The physical working environments in the office buildings, such as illumination, air quality, temperature, and humidity have long been recognized as the key features in the ambient work environments that drive human comfort and thus improve their health conditions [1], [2].

A lot of research focus has been put on the adverse influences from the indoor exposures, whereas the *optimized building indoor environments* can actually provide benefits to the human health. In addition, another study [3] has also suggested that *proper indoor-environmental settings* can greatly reduce the energy usage and thus, create more environmental-friendly buildings. Usually buildings account for about 30% to 40% of the total energy consumptions and carbon dioxide emissions worldwide [4]. For example, buildings use up to 40% of energy in the U.S.A. along. To reduce the energy consumption of buildings and make them more energy-efficient can greatly contribute to lowering the overall carbon dioxide emission and thus, helping alleviate the impacts of global warming. These are the key factors that need to be taken into consideration when we design a smart building system to adjust indoor-environment settings automatically.

Usually, building occupants are not fully-aware to adjust physical features of the building, or it would be onerous and demanding for them to adjust the environmental settings according to different factors. Thus it is ideal that there is an automated system which can adjust the features based on their personalized needs/schedule.Thus, it would be ideal if there is an automated system that can adjust indoor-environment features based on occupants' personalized needs/schedule. One set of existing works [5]–[8] on feedback-based systems tries to improve the comfortability of the occupants. These works utilize the survey from the occupants to increase the usability and comfortability of the buildings. Another set of existing works [9]–[12] predicts occupants' coarse-grained activities without any consideration of occupants' comfort and usability. These works utilizes occupants' rough activity-wise preferences of physical features to control the physical features automatically.

In this paper, we design a system which bridges the gap between above two sets of existing works. We propose a smart building assistance system which automatically controls the building settings to increase occupants' comfort level. Specifically, the proposed system predicts the physical features for current activity so that occupants feel comfortable and their levels of desire are satisfied based on their previous activities and the corresponding physical features.

We first utilize the building data from an office environment to realize the proposed system. The dataset consists of different physical features (lighting, shading, temperature, air quality, noise level, etc.) of the building environment, calendar events, common office activities, public holidays, and weather information. Also, we collect occupants' preferences to predict the building settings. The dataset is collected from May 2018 to August 2019 of an office room. To accurately predict different physical features, we carefully choose different data analysis techniques and feature selection procedures to exploit the expressiveness of the collected dataset. More specifically, we perform feature discretization, feature combination, and regularization to enhance the expressive nature of the dataset. Then, we propose a support vector regression (SVR) based prediction model and a hybrid neural network (DNN) model. The proposed hybrid DNN model leverages a mixtures of a feed-forward DNN block and several recurrent neural network (RNN) blocks to achieve higher prediction accuracy. Furthermore, we conduct extensive experimental studies using the collected datasets to show that the predominant advantages of the proposed assistance system in terms of comfortability. From the experiments, we conclude that the proposed system is able to increase the comfortability of the occupants.

The following lists the major contributions of this paper:

(1) **Smart building data:** We have used real smart building data such as lighting, cooling, shading, temperature, humidity, office calendar events, occupancy, and historical weather data for our study. Also, we collected users' preferences to predict the building settings. We focus on the singly occupied room (i.e., faculty room). A detail description of used dataset is provided in Section III.

(2) **Physical features predictions:** Based on dataset just mentioned above, we applied a SVR model and a mixture of feed foward DNN and RNN models to predict different physical features. We also pre-processed the data and chose different features to increase the prediction accuracies. Based on the predicted results, we recommend the preferred settings for different physcial featues that maximize occupants' comfort level.

(3) **Extensive experimental studies:** We conducted extensive experimental studies based on the building's indoor environment data along with other calendar events and weather related datasets. From the experiments, we found that the proposed smart building assistance system is able to increase occupants' comfortability by adaptively changing the indoor environmental features (e.g., temperature, humidity, and so on) and reducing the gaps between indoor and outdoor environments.

The rest of the paper is organized as follows. Section II discusses the existing literature. Section III presents the system design of the proposed predictive model. Section IV evaluates our proposed models through extensive simulation studies. Finally, Section V concludes this paper with remarks on the future work.

II. Related Works

Researchers have been working for over a decade to make the indoor environments increasingly intelligent in different ways. We can divide the existing works in two different groups as described below.

Optimization of physical features of the smart buildings. For improving occupants' satisfaction, the methods based on occupants' feedback gave rise to early works in improving the physical features of buildings [5]–[8]. Jazizadeh *et. al* [5] proposed an intermediary occupants communication platform, which aims to estimate and model different comfort levels of the occupants separately in order to enable a personalized comfort driven heating, ventilation, and air conditioning (HVAC) operations for all occupants. Winkler *et. al* [7] developed a voting-based interface to adjust the HVAC system. Through a 40 week user study of 61 university employees across 3 buildings, the authors showed that feedback systems can be used to increase user satisfaction with improved thermal condition and reduce energy consumption caused by HVAC systems. In the work of [8], the authors presented several occupants' feedback collection methods to design a comfort voting application with the consideration of occupants-environment interaction. García *et. al* [13] proposed a context-aware collaborative learning framework for home management systems which learns context-aware occupants' desired physical factors and tries to influence occupants to reduce energy usage. However, these occupants-feedback based methods are limited by the requirements of manually controlling the the HVAC system by the occupants themselves. Therefore, some occupancy-detection based methods have attracted much attention from researchers during recent years so that the HVAC systems can be adjusted automatically based on the detection mechanism. In the work of De Silva *et. al* [14], a lighting system is integrated with occupancy detectors and daylight sensors for fully automated operations of adjusting the lighting system. Zou *et. al* [15] presented a wireless occupancy-driven lighting control system to reduce energy consumption while simultaneously preserving the lighting comfort of occupants. Huang *et. al* [16] integrated CO2 and light sensors with a wireless sensor platform and developed an occupancy detection method which can achieve higher accuracy while keeping low cost and non-intrusiveness. Ambient sensor systems [17] use collections of sensors such as temperature, light, door switches, or electricity consumption to infer the presence of people in a room. Das *et. al* [18] presented an occupancy counting solution of the building that only relies on information already available from existing modes of sensing in buildings including electrical energy demand, water consumption and number of wireless network connected devices. Using a large-scale building dataset from a major university campus, they showed that these three modalities are strong indicators of accurate number of occupancy estimates using machine learning based algorithm such as clustering. Li *et. al* [19] proposed two methods: a new moving-window inhomogeneous Markov model based on change point analysis and an integrated hierarchical probabilistic sampling model to obtain flexible occupancy models of various kinds of temporal (e.g. intra-hour and hourly) and prediction horizons (e.g. hour-ahead and day-ahead). However, all the works discussed above do not consider the individual occupants' environmental preference among different indoor activities.

Occupants' activities prediction. Another set of works [9]–[12] predicts occupants' activities using different machine learning approaches (e.g., Bayesian Network (BN), Markov Model (MM), and others). Alam and Reaz [9] proposed a MM-based algorithm, SPEED, to predict the next household activity (e.g., turning light on/off, watching TV, eating, and sleeping) from the historical traces where inputs are the time stamp and previous activities and output is the current activity. Wilke *et. al* [12] proposed a bottom-up modelling approach together with a set of calibration methodologies to predict residential building occupants' time-dependent activities (e.g., relaxing, sleeping, studying, dancing/party, eating, gardening, and cooking), for use in dynamic building simulations where the inputs are the previous time dependent activities and output is the probability of next activity. In another work [10], the authors presented an improved prediction model using BN for the occupants' household activity prediction. Unlike the traditional BN, the proposed model utilizes current features (e.g., activity location, activity time of day, activity day of week) and next features to classify the next activity (e.g., bathing, sleeping, eating), whereby the immediate next activity is predicted using the current activity. Consequently, in [11], real-time sensor data has also been used with BN network for recognizing and predicting multiple human activities (e.g., bathing, working in leaving/bed room, eating, cooking, watching TV, sleeping) in the building at the same time. The authors utilized previous sequence of activities as the inputs of the BN network to predict the conditional probability of residents' current activities.

We should note however that approach used in previous studies that focus on indoor activities in residential buildings is not suitable for predicting activities of office individual occupant. Even though the work of Tabak *et. al* [20] considered the office activities, they mainly focused on predicting the frequency and occupancy of the intermediate activities (which interrupt the planned activities) conducted by office occupants but neglected the primary activities (i.e., pre-planned). The authors used intermediate activities (e.g., walking to printer/mailbox, having lunch, getting a drink, smoking, taking a break), time of the day, and day of the week as inputs to the S-curve and probabilistic prediction model to predict the activity at current time. Similarly, Zhao *et. al* [21] only achieved a four-type office occupant behaviors (i.e., occupied computer-based work, occupied non-computer-based work, unoccupied remote work, and unoccupied) prediction model using support vector machine, locally weighted learning methods where inputs are the electricity consumption data of computers, computer monitors, task lights, and other office appliances. Peng *et. al* [22] provided an in-depth analysis of occupants' stochastic behavior (e.g., entering/leaving the office room, working in the desk, and office appliance usages) within an office building. The authors proposed a demand-driven control strategy that automatically responds to occupants' energy-related behavior and stops the office appliances automatically for reducing energy consumption and maintains room temperature for occupants. In this paper, we use the previous physical features and calendar events of a office room occupant to predict the next calendar event and physical features.

III. System Design

In this section, we first describe the overall architecture of the proposed system (Section III-A). Then, we describe the

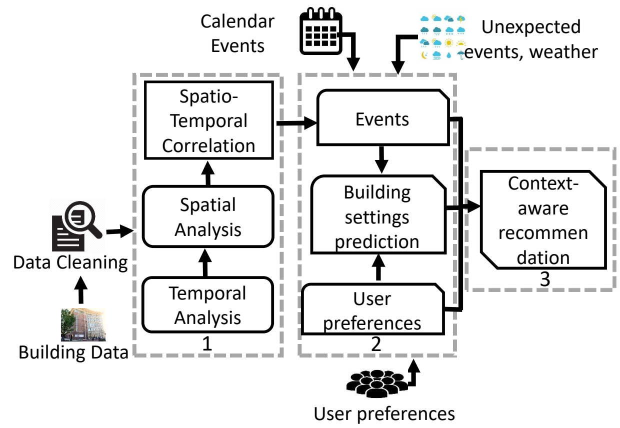


Fig. 1: Overview of the proposed smart building system.

data and its analysis we used in our system (Section III-B). Finally, we present the prediction and classification parts of the proposed system (Section III-C).

A. Overview

The proposed smart building assistance system consists of three parts: data collecting and data processing, physical features and activity prediction module, and context-aware recommendation. Fig. 1 shows the overview of the proposed smart building system. First, we collect smart-building sensors data such as lighting, shading, cooling, temperature, humidity, and occupancy. These are time-series sensor data. We pre-process and clean the data in 15-minutes window basis to combine with the occupants' events (e.g., meetings, classes, lunch hours) as well as with the calendar events (i.e., organizational activities). We also collect the users' preferences (e.g., preferred lighting intensity, indoor temperature, window shade, air quality) for physical features of the the room. We analyse and categorize the collected data based on spatio-temporal correlations. Here, spatial analysis refers to the categorization of events and physical features with respect to the different locations (e.g., inside office room, meeting room, outside). After cleaning and preprocessing the data, we combine those to predict physical features. For predicting the physical features, we utilize occupants' preferences as well. Finally, our model recommends different physical features which are suitable and comfortable with respect to the predicted event.

B. Data Analysis

Here, we discuss the data sources and data analysis of the proposed system.

1) Data sources: We have the following data sources from the smart building environment:

(i) *Physical features:* We collected the smart building physical features data like temperature, humidity, lighting, air quality, door opening/closing, number of people. The physical feature data is collected from a faculty room in a three-semester long period from April 2018–May 2019. Table I shows the different physical features collected from the faculty room. The temperature, humidity, and illumination features are the most perceptible ones

TABLE I: Different physical feaetures.

Physical Factors	Term	Description
Temperature ˆ	Temperature*	Indication of mean temperature in 15 mins time period
Humidity ˆ	Humidity*	Indication of mean humidity in 15 mins time period
Lighting ˆ	Illumination*	Luminous flux per squared meters area; it describes light intensity
Air quality ˆ	CO2*	Carbon dioxide concentration
	Concentration*	Air substance concentration (i.e., presence of carbon monoxide, lead, nitrogen dioxide, ozone, particulate matter of different size fractions, and sulfur dioxide in the air)
	VOC*	Volatile organic compounds (i.e., organic chemicals from paints, varnishes, and wax are emitted as gases)
Door open	Contact	Indication of door opening and closing in 15 mins time period
Existing people	PIR	Passive infrared sensor value (i.e., motion object detection); it indicates if there are persons in the room
Number of occupants	Ground truth from the person	How many people in the room

ˆ These are physical environment factors.
* These are used as inputs to the prediction of activity.

TABLE II: Adjustment ways for physical factors.

Physical Factors	Control systems	Adjustment ways
Temperature	Heating; Air Conditioning	Turn on/off; Turn up/down
Humidity	Humidifier; Air Conditioning	Turn on/off; Turn up/down
Lighting	Lamp	Turn up/down
Air quality	Window; Door; Air Conditioning	Close/open; Turn on/off; Turn up/down

that directly affect the comfortability of the occupants. They are measured by corresponding sensors at every 15 minutes. And, air quality feature has direct impact on the occupants' physical healths. For example, the concentration of the air is associated with the conditions of the respiratory systems. The higher the concentration, the more deleterious the air is. Other physical features are for the smart building assisting purposes. For example, the number of occupants is negatively correlated to the desired indoor temperature, as the increase in the number of people leads to the rise in room temperature. We can see that for different data items, the table also shows the meaning of the data item. For example, the air quality is represented by CO2, concentration, and violative organic compounds (VOC) where each of the terms has different meaning as presented in the third column of Table I. We used the contact sensor for monitoring door opening and the infrared sensor for detecting moving objects. The event/activity-wise number of people inside the room is collected from occupants verbally. Later, it is confirmed by the readings from the motion sensor. Furthermore, Table II shows the ways to adjust different factors in the faculty room. For instance, occupants can adjust the lighting condition gradually from dark to fully bright.

TABLE III: Office activities need to be predicted

Index	Activity
1	Calling
2	Office hour
3	Somebody visiting
4	Preparing for classes
5	Lunch time
6	Individual meeting
7.1	Weekly team meeting
7.2	Project meeting
7.3	Speaker
7.4	Business trip
7.5	Lunch or dinner with visitor
7.6	Faculty meeting
7.7	Teaching
7.8	Night time
8	Research activity
9	Not real event

(ii) *Calendar events:* We also collected the occupants professional calendar event data from April 2018 to May 2019 to understand different activities throughout the week. We combined these data with different department activities and game day events information to understand the patterns of different events occupants have to attend and the relationship between these common events and the physical features of the room. Table III shows the different activities we pulled from the calendar data. We can see that there are nine types of main activities from table and most of the event happen inside the office room. We further categorized the outside events (index 7) into eight different categories.

(iii) *Weather:* We also collected external weather information from a publicly available historical weather information API, in order to take the impacts of external weather on the desired indoor environment into consideration. For example, when outdoor temperature falls below a certain threshold, indoor temperature should be adjusted to a higher degree. Historical weather data is recorded every six hours. The collected data includes weather conditions (e.g., cloudy, rainy, sunny, fog, clear, thunderstorm),

TABLE IV: The range of physical factors.

Range	Low	Mid: Low	Mid	Mid: High	High
Temperature/°C	<18	[18,22)	[22,25)	[25, 30)	>30
Humidity/%	<28	[28,38)	[38,48)	[48,58)	>58
Lighting/lx	<200	[200,300)	[300,400)	[400,500)	>500
CO2/ppm	<500	[500,600)	[600,700)	[700,800)	>800
Concentration/ ppm	<650	[650,780)	[780,910)	[910,1040)	>1040
VOC/ ppb	<405.8	[405.8,532.6)	[532.6,659.4)	[659.4,913)	>913

current temperature (in Celsius), pressure (in Hg), wind speed (in mph) and humidity (in %).

(iv) *Preference:* We interviewed occupants of the office room monthly to get to know his/her personal presences during different activities retrieved from the calendar event. Additionally, we asked about their preferred range of different physical features during a certain activity.

2) Data preprocessing and feature selection: After data collection, we pre-processed our datasets and combined them together. This process was accomplished in Python with the help of some machine learning packages such as pandas and sklearn. Table IV shows the range of physical features along with three groups. It also shows the units of different physical features. We first converted the raw data into a data frame with size 23808×22 where 23808 is the number of instances for each physical feature within a 15-min time window and 22 is the number of physical features(e.g., the number of occupants in a room, indoor activity, outdoor temperature). Then, we considered the distribution of each feature separately to rule out abnormal observation values.

For feature engineering and selection, we performed the following tasks successively:

- (i) *Discretization and one-hot encoding:* Here, we treat categorial features and numerical features differently. For each categorical feature (e.g., activity, PIR, time slot, week day, and weather), we use one-hot coding to generate multiple binary features to replace the original feature. For each numerical feature, we segment its value into 10 bins and generate its one-hot representation while keeping the original feature, in case that some tree-based models may utilize the expressiveness of numerical features.
- (ii) *Design combined features:* To enhance the expressive power of the data set, we design some useful combined features (e.g., Cartesian product of two categorical features) based on common knowledge. For example, the combination of humidity and outdoor temperature is a stronger indicator for the indoor temperature setting than either of these two features separately as the combined feature can effectively capture the extreme conditions (for example, when both the humidity level and outdoor temperature are extremely high, the indoor temperature should be set to a low level almost for sure).
- (iii) *Statistical features:* In addition, we extract some statistical values (e.g., minimum, maximum, median) to expand the feature space. In particular, for each feature x_i in a sample vector x, we calculate the minimum, maximum, median, mean and standard variation values for all the samples which share the same timeslot with x_i but in different days. These statistical values are added in x as the new feature columns.
- (iv) *Time-series features:* From our initial observations, we find that different physical feature values exhibit periodicity. Therefore, we utilize different values of physical features over time as time-series features. First, we arrange the datasets in ascending order with respect to timeslots. Let and let $x^{(t)}$ be the feature vector obtained after one-hot encoding, feature combination and statistical extraction for the t-th sample(The length of $x^{(t)}$ is 305). Then, we fix a time step window k and concatenate every k adjacent $x^{(t)}, t = T-k+1, \cdots, T-1, T$, to form the complete input for predicting $y^{(t)}$.
- (v) *Feature selection:* We apply $\mathbb{L}_1$ regularization to select different features because it exhibits better performance than other feature selection approaches (e.g., sequential forward/backward selection). However, we feed the whole feature set into deep learning models without explicitly selecting any particular feature.

C. Physical Features Prediction Model

For the physical feature prediction model, we utilize a DNN model leveraging a mixture of LSTM blocks and a feed-forward DNN model where inputs are the time series data of different physical features. The detailed descriptions of the proposed hybrid DNN model is explained below.

1) Support Vector Regression: Support vector regression is a variation of support vector machine(SVM) that is used as a regression method, maintaining all the main features that characterize the algorithm (maximal margin). SVR uses the same principles as SVM does for classification. A margin of tolerance (i.e., epsilon) is set in approximation to the SVM which would have already requested from the problem. The main purpose is the same, to minimize error by individualizing the hyperplane which maximizes the margin, keeping in mind that part of the error is tolerated. To train a standard SVR for data $(x_i, y_i)_{i=1}^N$, we need to solve the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|\omega\|^2, \\ \text{subject to} \quad & y_i - \omega x_i - b \leq \varepsilon \\ & \omega x_i + b - y_i \leq \varepsilon, \end{aligned}$$

where ε is a free parameter that serves as a threshold: all predictions have to be within an ε range of the true predictions. Slack variables are usually added into the above problem to allow for errors and to allow approximation when the above problem is infeasible.

2) Feed-forward neural network: A feed-forward neural network represents the function $f(x;\theta)$ where x is the input vector and θ is a set of parameters. Specifically, f is a composition of functions. The smallest unit of a neural network is a so called neuron. It maps the weighted sum $\sum_{i=1}^{k} x_i w_i$ to an activation value via the function $f_{act}(x^T w)$ where x is the vector of inputs for the neuron and w is a vector of parameters denoted as weights. A layer of the network is a set of neurons that usually use the same activation function. In this case, a layer i can be represented as function $f^{(i)}(x; W^{(i)}) = f_{act}(W^{(i)^T} x)$ where x is the input vector of the layer, $W \subset \mathcal{R}^{k \times l}$ is the matrix that contains the weights of l neurons (i.e., each column of W represents the weights of a neuron). Putting all together, such a feed-forward neural network represents a composition of the layer functions with the parameters $\theta = \{W^{(1)}, W^{(2)}, ..., W^{(n)}\}$. The first layer is called input layer and is only there to receive the input with the identity function $f^{(0)}(x) = x$. The last layer $f^{(n)}$ is called output layer. All other layers are called hidden layers.

3) Long short term memory network: Recurrent neural network (RNN) is a powerful model for time series data modeling. Specifically, given an input temporal sequence $x = (x_1, x_2, ..., x_T)$, where in our case x_i is the i-the input instance of a physical feature, the hidden states of a recurrent layer $h = (h_1, h_2, ..., h_T)$ and the output $y = (y_1, y_2, ..., y_T)$ of a RNN model can be obtained as:

$$h_t = \theta_h(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \tag{1}$$

$$y_t = \theta_y(W_{ho}h_t + b_o) \tag{2}$$

where W_{x_h}, W_{h_h}, and W_{h_o} are connection weight matrices for input layer, hidden layer, and output layer; b_h and b_o are bias values, and θ_h and θ_y are activation functions.

LSTM, a version of RNN, is with forget gates and peephole connections. The key point of LSTM is the cell state c that capacitates RNNs to memorize by removing or adding information to it. This manipulation is mainly regulated by three modules, namely the input gate (i), forget gate (f), output gate (o). LSTM proceeds by the following functions:

$$i(t) = \sigma((W_{xi}x^{(t)} + W_{hi}h^{(t-1)} + W_{ci}c^{(t-1)} + b_i), \tag{3}$$

$$f(t) = \sigma(W_{xf}x^{(t)} + W_{hf}h^{(t-1)} + W_{cf}c^{(t-1)} + b_f), \tag{4}$$

$$c(t) = f^{(t)} \cdot c^{(t-1)} + i^{(t)} \cdot (W_{xc}x^{(t)} + W_{hc}h^{(t-1)} + b_c), \tag{5}$$

$$o(t) = \sigma(W_{xo}x^{(t)} + W_{ho}h^{(t-1)} + W_{co}c^{(t)} + b_o), \tag{6}$$

$$h(t) = \phi(c^{(t)}) \cdot o^{(t)}; \tag{7}$$

in which σ is an element-wise application of the logistic sigmoid function, ϕ is an element-wise application of the *tanh* function and $\cdot$ denotes element-wise multiplication.

4) Proposed hybrid DNN model:: To utilize both the time-series nature and the structured form of input data, we propose a hybrid neural network that consists of several standard LSTM blocks and a feed-forward 3-layer neural network block on top of the LSTM blocks. The architecture of the proposed model is shown in Figure 2. There are k LSTM blocks to utilize previous k time stamps of the inputs. The input of the LSTM blocks at different time stamp t is the concatenation of $X^{(t)}$ and $Y^{(t)}$ where $t = T - k, T - k + 1, \cdots, T - 1$. $X^{(t)}$ is the feature vector with 305 dimensions and $Y^{(t)}$ is a 5-class one-hot label vector, both at time slot t. Here if $t < 0$, $X^{(t)}$ and $Y^{(t)}$ are padded with zero vectors. The input of the feed-forward neural network is the concatenation of $X^{(T)}$ and $h(T)$, where $h(T)$ is the output of the LSTM block. The output of the hybrid

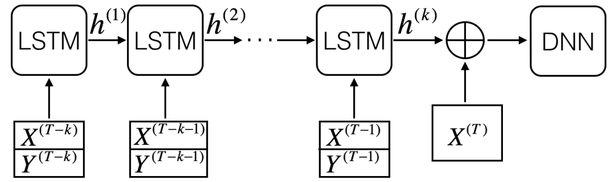


Fig. 2: The proposed hybrid deep learning model.

DNN model is $\hat{Y}^{(T)}$, a 5-class softmax probability vector. The loss function is given by the cross-entropy between true label $Y^{(T)}$ and the network output $\hat{Y}^{(T)}(\theta, X)$ as given below:

$$Loss(\theta, X) = -\sum_{T=1}^{N} \sum_{i=1}^{5} Y_i^{(t)} \log(\hat{Y}_i^{(T)}(\theta, \hat{X}^{(T)})) \tag{8}$$

where N is the number of training instances, $\hat{X}^{(T)}$ is the hybrid input for predicting $Y_i^{(t)}$ and θ represents the weights of the network.

Hyperparameters. The model is implemented using *Keras* with input feature space dimension being 305, input time-series length k being 21, and the output length of LSTM block being 5. Therefore, the numbers of neurons in each layer of the DNN is 305, 30, 10, and 5, respectively. In the proposed hybrid DNN model, we use the ReLU activation function. We also use the Adam optimizer instead of other optimizers (e.g., Adamax, RMSprop, Adagrad) to train our model where we set the learning rate as 10^{-3} in the optimizer. The number of epochs is set to 200. We use the cross validation to test the trained model and reduce the overfitting of the trained model.

IV. Performance Evaluations

In this section, we first present the experimental settings we used to evaluate the proposed system. Then, we present the experimental evaluations to validate the proposed approach.

A. Experimental Settings

For the experiment, we predicted 6 indoor physical factors including temperature, humidity, lighting, CO2, concentration, and VOC based on the sensor data. The followings summarize the experimental settings of different approaches:

1) Inputs: The inputs are time phase of a day, day status (e.g., weekday or weekend), day of the week, number of occupants, weather, outdoor temperature, school event status, department events status, holiday status, and activity types. After applying data preprocessing and feature engineering methods as described in Section III-B, we obtained an input sequence $X^{(t)}, t = 1, 2, \cdots, N$ where $N = 23808$. Each $X^{(t)}$ is a vector with dimension of 305 and contains all the information for a specific time slot.

2) Outputs: The outputs are six physical features including temperature, humidity, lighting, CO2, concentration, and VOC. For each factor i, we set a numerical objective $y^{(i)}$ using its original value and a categorical value $Y^{(i)}$ by discretizing its numerical into five intervals (see Table. IV for more details).

TABLE V: The hyper parameters in model

Model	implementation	hyper-parameter settings	info
Linear Regression	sklearn	Default	standard LR without regularizations
LASSO	sklearn	$\alpha = 0.002$	α :Constant that multiplies the $\mathbb{L}_1$ term
SVR	sklearn	C=10	C: Penalty parameter of the error term
GBRT	sklearn	n=50	n: The number of boosting stages to perform
SVM	sklearn	C=20	C: Penalty parameter of the error term
Softmax Regression	Keras	Default	implemented as 1-layer neural network
DNN	Keras	Default	3-layer neural network with size $30 \times 10 \times 5$
LSTM	Keras	Default	see III(C)(2)
Hybrid DNN	Keras	Default	see III(C)(3)

3) Comparative methods: In the case of regression, we used four different methods: linear regression, LASSO regression, support vector regression (SVR), and gradient boosting regression tree (GBRT). In the case of classification, we used five different methods: softmax regression, support vector machine (SVM), DNN, LSTM, and, the our proposed hybrid DNN model.

All the comparative methods were implemented using sklearn or Keras. The hyper parameters of the comparative models were all well-tuned to obtain the best possible performances. The optimization solver was Adam with learning rate of 10^{-3}, and the termination criteria is 200-epoch iteration on the training set. Some of the major hyperparameters are presented in Table. V. The parameter settings for neuron network models are already introduced in Section III-C. Note that in Table. V, by saying default hyper-parameter setting, we mean to apply the default implementation in these packages.

4) Evaluation metrics: For the comparative analysis, we mainly applied four different evaluation metrics: mean absolute percentage error (MAPE), accuracy, and comfortability.

1) We utilized MAPE to evaluate the regression models. To evaluate MAPE, we formalized the prediction task as a regression problem and take the mean value of the percentage deviation for each predicted data point as follows:

$$MAPE = \frac{1}{N}\Sigma_{i=1}^{N} \mid \frac{y_i - p_i}{y_i} \mid, \tag{9}$$

where N is the total number, y is the actual value, and p is the predicted value. We found that MAPE is a better metric compared with root mean square error (RMSE) because it alleviates the bias introduced by the magnitude of different objectives.

2) We used accuracy to evaluate the classification models. To evaluate accuracy, we formalized the prediction task as a classification problem by discretizing each targeted column into 5 levels based on quantile statistics. The accuracy represents the fraction of truly predicted physical factor level among all time slots as follows:

$$Accuracy = \frac{\sum_{i=1}^{N} \mathbb{I}(Y_i = P_i)}{N}, \tag{10}$$

where Y is the actual level and P is the predicted level. $\mathbb{I}(Y_i = P_i) = 1$ if and only if $Y_i = P_i$. Note that these two metrics are both calculated in 5-fold validation.

3) We consider comfortability level to evaluate the predicted level with respect to the personal preferences. We define $CR = \{F|f_i \in [f_i^{(L)}, f_i^{(H)}], i = 1,2,3,4,5,6\}$, the comfortable range for the 6 predicted physical factors $F = (f_1, f_2, f_3, f_4, f_5, f_6)$ by empirical measurements as shown in Table. VI. Then, the average comfortability level is calculated as follows:

$$Avg.\ comfortability\ level = \frac{1}{N}\sum_{i=1}^{N} CL(F_i), \tag{11}$$

where

$$CL(F) = \begin{cases} 0, & if\ F \in CR, \\ 1, & otherwise. \end{cases}$$

The binary function CL measures whether the given physical factor configuration falls into the pre-defined comfortable range. And when calculating the overall comfortability level for six physical factors, Equation (11) assigns equal weights to each factor.

TABLE VI: The comfortable range for physical factors

Factors	Notation	Lowest value $f_i^{(L)}$	Highest value $f_i^{(H)}$
Temperature/°C	f_1	15	25
Humidity/%	f_2	40	80
CO2/ppm	f_3	0	600
Concentration/ ppm	f_4	700	900
VOC/ ppb	f_5	300	1000
Illumination/lx	f_6	300	1000

B. Experimental Evaluations

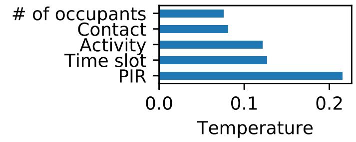
In this section, we describe the experimental evaluations in terms of prediction accuracy, and comfortability, respectively. First, the prediction accuracy of different algorithm is shown in Table. VII and VIII. As we can see from Table VII, the

TABLE VII: The MAPE of predicted physical factors.

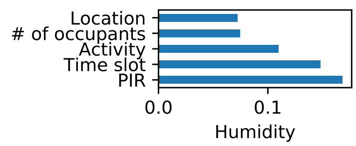
Objectives	Linear regression	LASSO	SVR	GBRT
Temperature/°C	0.06	0.05	0.03	0.07
Humidity/%	0.13	0.13	0.07	0.18
CO2/ppm	0.06	0.05	0.04	0.06
Concentration/ ppm	0.06	0.06	0.05	0.07
VOC/ ppb	0.22	0.21	0.17	0.26
Illumination/lx	0.08	0.08	0.10	0.11

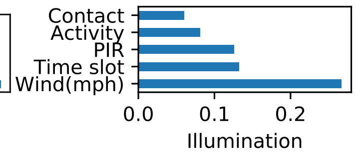
TABLE VIII: The prediction accuracy of physical factors.

Objectives	Softmax regression	SVM	DNN	LSTM	Hybrid Model
Temperature/°C	0.76	0.79	0.84	0.84	0.85
Humidity/%	0.76	0.78	0.82	0.84	0.85
CO2/ppm	0.89	0.88	0.91	0.92	0.92
Concentration/ ppm	0.85	0.85	0.86	0.84	0.86
VOC/ ppb	0.80	0.82	0.92	0.97	0.97
Illumination/lx	0.97	0.97	0.98	0.98	0.99

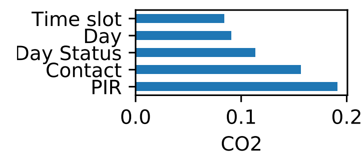


(a) Top 5 relevant feature for Temperature

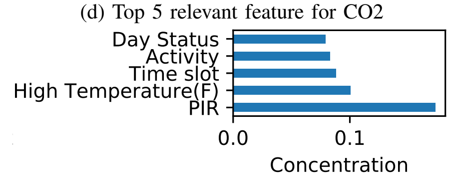




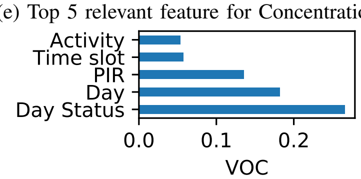
(b) Top 5 relevant feature for Humidity (c) Top 5 relevant feature for Illumination



(d) Top 5 relevant feature for CO2



(e) Top 5 relevant feature for Concentration



(f) Top 5 relevant feature for VOC

Fig. 3: Top five feature importance scores for different physical features.

proposed approach SVR performs better than all other methods except for VOC. The MAPEs for all the other five physical factors reaches a level under 10%. For VOC, the prediction performances are relatively poor for all comparative methods. It is due to the fact that the VOC factor is less correlated with the collected indicators.

Similarly, from Table VIII, we can see that the best performance of classification comes from the proposed hybrid DNN model. It is due to the fact that the proposed hybrid DNN model exploits both the temporal structure of the data and the high-level feature-generating power of neural network architecture. With the proposed Hybrid DNN model, the accuracy for each physical factor exceeds 85%, which seems totally acceptable in practice.

Figure 3 shows the top five most important features in predicting each physical feature. The importance of a feature is measured by calculating the decrease in the model's prediction error after permuting that feature. A feature is important if changing its value decreases the model error because, in this case, the model relied on that feature for further prediction. From the figure, we can see that for every different target feature, different features play key roles in during the prediction process. This is primarily because each target feature depends on a different set of features. Specifically, CO2 mainly depends on the number of people in the office instead of what people are doing in the office while temperature has a close relation with the activity people conduct in the office as activity determines the amount of heat people emit.

Despite the differences in these top five lists, they generally match the intuitive sense in daily life. For example, in the prediction of CO2 concentration in the office, PIR and contact have the largest feature importance. This matches the intuition as CO2 from humans exhalation is the primary source of CO2 in an indoor environment. Thus, the more people in the office (and the higher PIR consequently), the higher the CO2 concentration. Moreover, Since the office door controls the air exchange between the office and outside environment, contact, an indication of the door opening and closing in a time period, has the second-highest feature importance. Day status, day and time slot also have major impacts on the number of people in the office. For instance, intuitively, there are more people in the office on weekdays than on weekends. These features indirectly influence the CO2 concentration in the office and thus have relatively high feature importance.

On the other hand, there are features that play important roles in multiple predictions. In general, data with a wide range tends to have an effect on multiple target features. Take activity as an example, there is a diverse range of activities, each linked with a unique scenario. Some of the scenarios involve people coming in the office and some involve people coming out of the office. There are multiple people in some scenarios while there is only a single person or even no people (e.g., lunchtime, business trip, and so on) in other scenarios. As a result, a shift from one activity to another might introduce a dramatic change to the scenario inside the office, thus exerting an impact on target features.

From the above figures and discussion, we can conclude that the proposed automatic smart building assistance is able to predict the activity-wise different physical features throughout the year to increase the comfortability of the building occupants.

V. Conclusions

In this paper, we proposed a smart building assistance system to increase the comfortability of the building occupants and save more energy at the same time. First, we collected and analyzed one year long data from a smart building room from different sources (i.e., sensors, calendar, weather, and survey). Second, we combined all the datasets together and preformed feature engineering (i.e., discretization, one-hot encoding, and multiple features combination) for further usage in different machine learning models. Third, we proposed a hybrid DNN model composed of several LSTM blocks and a feed-forward DNN block to predict different physical features for different activities. Fourth, we compared the proposed hybrid DNN model with other existing models in terms of accuracy and MAPE. We also conducted extensive experiments using real data to find the levels of comfortability and energy saving rates for different activities. From the experiments, we concluded that the proposed smart building assistance system is able to increase occupants' comfortability by adaptively changing the temperature, humidity, and other factors. The proposed system also reduces the difference between indoor and outdoor temperatures and aids to save energy by smartly reducing the workloads of control systems like air-conditioners when there are no occupants. We expect future research to focus on the following tasks: (1) Study the interpretability of our automatic control system to make each adaptation of factors understandable to humans and even to provide reasons for the changes. (2) Integrate human knowledge or customized settings in separated systems. For example, users in different rooms may have distinct preferences. We can further explore different machine learning algorithms to take personal preferences into account, which can then be used as inputs to generate the recommended indoor environment. (3) Improve the robustness of our system. Generally, DNN-based applications are sensitive to minor changes in their inputs and thus vulnerable to adversarial attacks. Improving the robustness of the system and thus avoiding malicious manipulations of indoor factors might be essential concerns in some safety-critical scenarios like government agencies and military forces.

References

[1] J. G. Allen, P. MacNaughton, J. G. C. Laurent, S. S. Flanigan, E. S. Eitland, and J. D. Spengler, "Green buildings and health," *Current Environmental Health Reports*, vol. 2, no. 3, 2015.

[2] M. P. Deuble and R. J. de Dear, "Green occupants for green buildings: the missing link?" *Building and Environment*, vol. 56, 2012.

[3] H. Chen, P. Chou, S. Duri, H. Lei, and J. Reason, "The design and implementation of a smart building control system," in *Proc. of e-Business Engineering*, 2009.

[4] B. Dean, J. Dulac, K. Petrichenko, and P. Graham, "Towards zero-emission efficient and resilient buildings.: Global status report," 2016.

[5] F. Jazizadeh and B. Becerik-Gerber, "Toward adaptive comfort management in office buildings using participatory sensing for end user driven control," in *Proc. of BuildSys*, 2012.

[6] V. L. Erickson and A. E. Cerpa, "Thermovote: participatory sensing for efficient building hvac conditioning," in *Proc. of SenSys*, 2012.

[7] D. A. Winkler, A. Beltran, N. P. Esfahani, P. P. Maglio, and A. E. Cerpa, "Forces: Feedback and control for occupants to refine comfort and energy savings," in *Proc. of UbiComp*, 2016.

[8] ——, "Forces: feedback and control for occupants to refine comfort and energy savings," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 1188–1199.

[9] M. R. Alam, M. B. I. Reaz, and M. M. Ali, "Speed: An inhabitant activity prediction algorithm for smart homes," *IEEE TSMC*, vol. 42, 2012.

[10] E. Nazerfard and D. J. Cook, "Crafft: an activity prediction model based on bayesian networks," *JAIHC*, vol. 6, 2015.

[11] Z.-H. Wu, A. Liu, P.-C. Zhou, and Y. F. Su, "A bayesian network based method for activity prediction in a smart home system," in *Proc. of SMC*, 2016.

[12] U. Wilke, F. Haldi, J.-L. Scartezzini, and D. Robinson, "A bottom-up stochastic model to predict building occupants' time-dependent activities," *Building and Environment*, vol. 60, pp. 254–264, 2013.

[13] Ó. García, J. Prieto, R. Alonso, and J. Corchado, "A framework to improve energy efficient behaviour at home through activity and context monitoring," *Sensors*, vol. 17, no. 8, p. 1749, 2017.

[14] U. De Silva, A. Lertsinsrubtavee, A. Sathiaseelan, and K. Kanchanasut, "Named data networking based smart home lighting," in *Proc. of SIGCOMM*, 2016.

[15] H. Zou, Y. Zhou, H. Jiang, S.-C. Chien, L. Xie, and C. J. Spanos, "Winlight: A wifi-based occupancy-driven lighting control system for smart building," *Energy and Buildings*, vol. 158, pp. 924–938, 2018.

[16] Q. Huang and C. Mao, "Occupancy estimation in smart building using hybrid co2/light wireless sensor network," *Journal of Applied Sciences and Arts*, vol. 1, no. 2, p. 5, 2017.

[17] Z. Yang and B. Becerik-Gerber, "Cross-space building occupancy modeling by contextual information based learning," in *Proc. of BuildSys*, 2015.

[18] A. K. Das, P. H. Pathak, J. Jee, C.-N. Chuah, and P. Mohapatra, "Non-intrusive multi-modal estimation of building occupancy," in *Proc. of SenSys*, 2017.

[19] Z. Li and B. Dong, "Short term predictions of occupancy in commercial buildingsłperformance analysis for stochastic models and machine learning approaches," *Energy and Buildings*, vol. 158, pp. 268–281, 2018.

[20] V. Tabak and B. de Vries, "Methods for the prediction of intermediate activities by office occupants," *Building and Environment*, vol. 45, no. 6, pp. 1366–1372, 2010.

[21] J. Zhao, R. Yun, B. Lasternas, H. Wang, K. P. Lam, A. Aziz, and V. Loftness, "Occupant behavior and schedule prediction based on office appliance energy consumption data mining," in *CISBAT 2013 Conference-Clean Technology for Smart Cities and Buildings*, 2013, pp. 549–554.

[22] Y. Peng, A. Rysanek, Z. Nagy, and A. Schlter, "Using machine learning techniques for occupancy-prediction-based cooling control in office buildings," *Applied Energy*, vol. 211, pp. 1343 – 1358, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261917317129