Thermal Sensation Prediction with Feedforward Neural Network

Ömer Faruk Önder
Middle East Technical University
omer.onder@metu.edu.tr

Abstract

The data obtained from buildings become important day by day. Although obtaining meaningful data from buildings requires a tremendous amount of time and money, recent studies show how valuable this massive exertion is. One of the parameters about a room and/or building is the feeling of the climate of the spaces, which is also called thermal sensation. The thermal sensation is a parameter that represents the thermal condition feeling of a place. Many features directly affect the feeling of warm, cold, or neutral climate in a room. In this study, the data obtained from Building Genome Project is used to predict the thermal sensation with a feedforward neural network structure. The results are also compared with Miller's implementation of classification according to the accuracies of the models.

1 Introduction

The amount of recorded sensor data is growing rapidly to understand the building operation and manage the energy usage in buildings [Habib et al., 2015]. Data-driven classification models are at the forefront of methods in energy consumption analysis on buildings [3]. In literature, collected data is used for several purposes as energy analysis, anomaly detection, building type classifications, thermal comfort classifications, and the development of clustering-based models [4]. The Building Genome Project is an open dataset provides a foundation for the building energy analysis by 3,053 energy meters from 1,636 buildings. The Building Data Genome Project is led by the Building and Urban Data Science Lab (buds_lab) of the National University of Singapore. The research group aims to leverage the data sources to understand the built environment and improve human comfort. Clayton Miller, head of buds_lab, refers to the dataset and exemplifies the developed models on the dataset [Miller et al., 2015, 2020, 2017]. The dataset is assessed in the competition of Great Energy Predictor hosted by the American Society of Heating, Refrigeration, and Air-Conditioning Engineers (ASHRAE) in October-December 2019 [Miller et al., 2020]. The dataset includes the information about the occupants as; sex, clothe, thermal sensitivity beside the environmental data: ['Year', 'Season', 'Climate', 'City', 'Country', 'Building type', 'Air temperature (C)', 'Relative humidity (%)', 'Air velocity (m/s)']. Hence, the same dataset is proposed to be used with a new perspective to understand the relation between the features better.

In this study, thermal sensation values of specific buildings/rooms are predicted with a neural network model. The data is obtained from Building Data Genome Project, which is a detailed study about data acquisition from buildings, conducted by Clayton Miller and Forrest Meggers. The neural network model used for the prediction of thermal sensation values is a feedforward neural network.

The thermal comfort's influence on health and productivity finds more evidence day by day. The study of Seppänen et. al. [2006] denotes that the performance of users is increasing with temperature up to 21-22 ºC and decreases above 21-22 ºC. 22 ºC is the best option for the highest productivity. Moreover, an indoor temperature of 30 ºC may cause a decrease in productivity up to 8.9%.

Moreover, Vimalanathan and Babu's study [Vimalanathan et al., 2014] states that thermal comfort influences the performance of the workers more than illumination. The contribution of thermal comfort to performance and productivity is 38.56% according to their study.

2 Background

2.1 Building Data Genome Project

Building Data Genome Project is a project aimed at creating an open building dataset for benchmarking. The study started in 2015 and over 60 million smart meters were installed only in the United States. The majority are from university campus buildings. The buildings are mainly non-residential and the data is open. The project started with interviews proceeded with teams of building operations working in various campuses around the world. The main idea is to collect hourly data from the selected buildings at least for one full year.

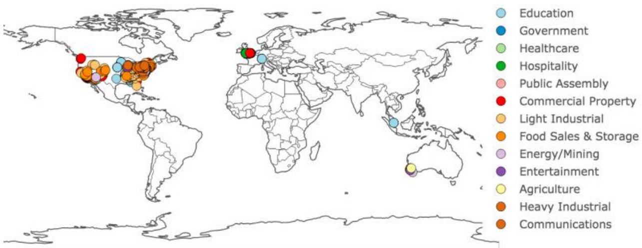

Figure 1. Locations of 1238 case study buildings collected from across the world

Two main data sources existed during the study. The first one is site visits for case studies. The sites visits are done mainly in the United States and to universities since universities enable access to several buildings with different conditions and usage types with minimum effort. The second data source is online open case studies, which is the online data on the Internet.

After data collection, buildings are divided into subsets under some top primary use types: Offices, Primary/Secondary Schools, Universities, Laboratories, University Classrooms, or Dormitories. After data cleaning and organization, a shareable set of building performance-related data from 507 buildings is ready to use. Each building includes hourly, whole-building electrical meter data, and various characteristic meta-data such as gross floor area, primary use type, weather information, and industry. The dataset is available for download and collaboration within a GitHub repository (https://github.com/buds-lab/the-building-data-genome).

2.2 Thermal Sensation

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) defines the term "thermal comfort as "the condition of mind that expresses satisfaction with the thermal environment. The thermal sensation is a standard parameter in most thermal experiments. This metric is also useful for parametrizing the feeling of the thermal comfort of a place (any indoor or outdoor places such as halls, rooms, gardens, etc.) according to its users." [Velt et al., 2017]

The thermal sensation is a scale for defining the feeling of the climate of a space. The thermal sensation is defined in ISO 7730 and ASHRAE standard 55 with a scale ranging from -3 (cold) to +3 (hot). Moreover, the scale can be extended to [-4,4] to have a more detailed classification. The Building Data Genome Project uses the [-3,3] interval. This interval is the same as the ASHRAE's 7 points thermal sensation scale. [Velt et al., 2017]

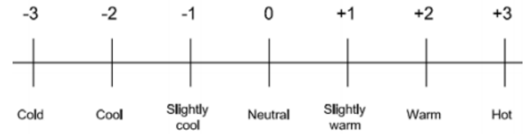

Figure 2. ASHRAE 7 points thermal sensation scale [Velt et al., 2017]

3 Methods

In the preprocessing process, the data columns 'Year' and 'Building Type' are dropped since these parameters influence thermal sensation lower than others. One hot encoding is applied to the rest of the parameters, which creates a large input dataset of 96 columns and 43448 rows.

Miller's implementation is constructed over a random forest classifier with 100 estimators, 2 minimum samples required to be at a leaf node and the test size is 30 %. The accuracy of Miller's model is 48.59 %. The aim is to have at least Miller's accuracy and improve that value if possible.

A feedforward neural network (FNN) is trained to predict the thermal sensation values. The sequential model starts with input nodes of 102. The model continues with 5 hidden layers (1 layer with 96 nodes, 2 layers with 192 nodes, 1 layer with 96 nodes, and 1 layer with 24 nodes). After every layer, Batch normalization s applied to the layer to stabilize the data. After every dense layer, a dropout with a rate of 0.1 is applied. The activation function is ReLU in every dense layer and for the output, the activation function is Softmax. The pipeline is shown in Figure 3.

The output layer has 7 nodes. The output values are -3, -2, -1, 0, +1, +2, +3 according to the ASHRAE Thermal Sensation scale. To reshape the output types for ReLU, the output values shifted by adding 3 to the thermal sensation values. So that the output values become 0, 1, 2, 3, 4, 5, 6. These are the shifted output values and the FNN model's outputs are according to this notation and interval.

The loss function is sparse categorical crossentropy. An Adam optimizer with a learning rate of 0.001 is selected as the optimizer of the model. The model is fitted to the train sets with 50 epochs. The feedforward neural network implemented on Jupyter Notebook with TensorFlow.

4 Results

Although the random forest classifier is easy to use and obtains great results in a relatively short time, neural networks can be superior in terms of accuracy.

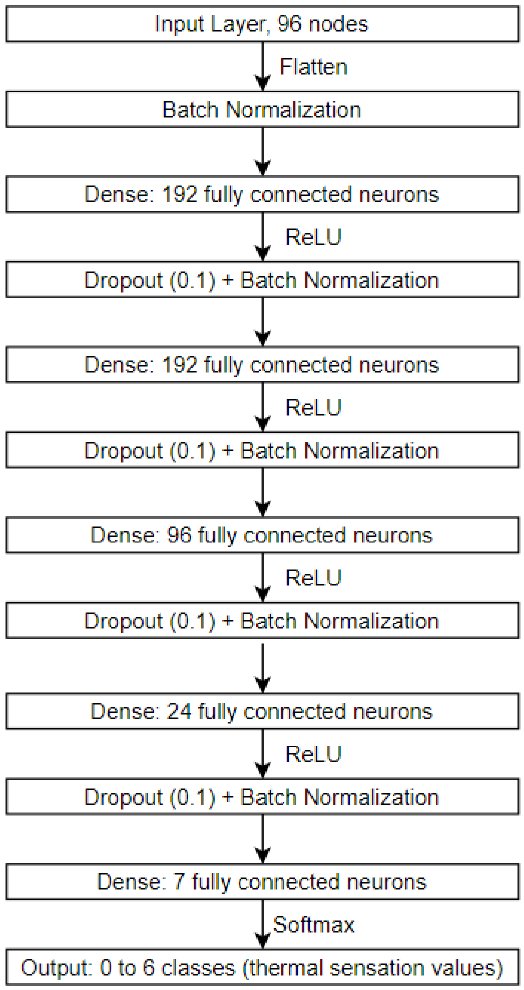

Figure 3. FNN pipeline

After 50 epochs, an accuracy of 48.93% is obtained from the FNN built for this study. The results of the FNN by means of accuracies and loss values are presented in Figure 4.

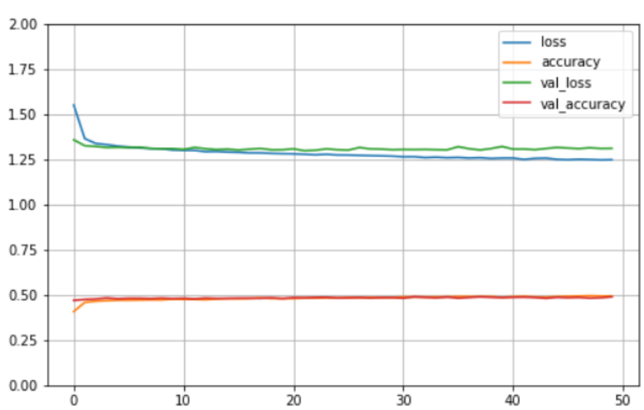

Figure 4. Accuracy and loss results of the model on training and test/validation sets

For this dataset, the feature importances obtained from the random forest classifier by the Miller's implementations. Figure 5 shows the feature importances of the first 15 important features. According to the table, air temperature, relative humidity, air velocity, and clothing are the most important parameters that influence thermal sensation. Some other informative outputs also shown in the .ipynb file.

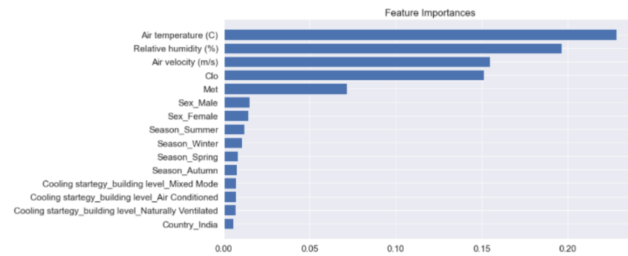

Figure 5. Feature importance graph of first 15 features that influences the data most

5 Conclusion and Future Work

In this study, the prediction of thermal sensation values using a feedforward neural network with the pipeline is shown in Figure 3. Miller's implementation used one of the most popular approaches, random forest classifier, and obtained an accuracy of 48.59 %. The FNN trained in this study has an accuracy of 48.93%, which is 0.34% better than Miller's implementation. On the other hand, the time consumption of the FNN is more than Miller's approach. The accuracy may be higher but if the one have concerns about time, a random forest classifier may be more beneficial.

Even if thermal sensation seems a data-dependent notion, it is quite subjective and human-dependent. Every person's perception of "hot weather" or a "chilled air" is different. The metabolism performance of the person, his/her emotional and hormonal state during the questionnaire, and many other physiological parameters have a significant influence on thermal sensation.

For example for the study proceeded by Velt and Daanen, 11 healthy Dutch subjects (7 males and 4 females) attended the experiment. In the experiment, winter and summer climates are simulated. In addition, a reference climate chamber is included in the setup. Moreover, the subjects wore suitable clothes for the climate that is simulated. The subjects are asked to stay for a while in one of the rooms and move to the other and their responses are collected with questionnaires. Such questionnaires give subjective data, which may be untrustworthy [Velt et al., 2017]. However, human perception is also subjective. So that, the sensory data in the Building Data Genome Project may be combined with more human-based data to create models with better performance parameters. More questionnaires may be proceeded to understand human perception of thermal comfort and the climate of the space. In addition, psychological adaptation may be an important factor to create a preferable model. Psychological adaptation may be included in such a model with the time consumed by the user in the specified room/space. Thermal comfort might be assumed proportional to the time consumed in the specified space. This approach may be helpful for the adaptive control units.

Moreover, the usage of other methods may be helpful to obtain better accuracy. These methods can be convolutional neural networks or K-Means-based structures. For example, a recent study used bag of words representation with clustering to find different patterns in buildings data [Habib et al., 2015. Different techniques and more fine-tuning would lead to predictions that are more accurate.

Machine learning and data science take a greater role in building science and engineering day by day. These approaches not only reduce the budget by optimizing the needs but also improve the life quality of the users.

These technologies not only important during the construction but also have crucial roles during the usage of the buildings. In general, HVAC (Heating Cooling, and Air Conditioning) systems contain the most energy consuming, noisy, and cumbersome devices in the building. Optimizing the working schedules and conditions of these devices may be a game-changing approach. Moreover, the optimization may be made according to the users' real-time feedback to the system actively and after enough data acquisition from the users, the system may adjust itself without any feedback.

References

[Habib et al., 2015] Habib, U., & Zucker, G. (2015, December). Finding the different patterns in buildings data using bag of words representation with clustering. In 2015 13th International Conference on Frontiers of Information Technology (FIT) (pp. 303-308). IEEE.

[Miller et al., 2015] Miller, C., Nagy, Z., & Schlueter, A. (2015). Automated daily pattern filtering of measured building performance data. Automation in Construction, 49, 1-17.

[Miller et al., 2017] Miller, C., & Meggers, F. (2017). The Building Data Genome Project: An open, public data set from non-residential building electrical meters. Energy Procedia, 122, 439-444.

[Miller et al., 2019] Miller, C. (2019). What's in the box?! towards explainable machine learning applied to non-residential building smart meter classification. Energy and Buildings, 199, 523-536.

[Miller et al., 2020] Miller, C., Kathirgamanathan, A., Picchetti, B., Arjunan, P., Park, J. Y., Nagy, Z., & Meggers, F. (2020). The building data genome project 2: Hourly energy meter data from the ASHRAE great energy predictor III competition. arXiv preprint arXiv:2006.02273.

[Vimalanathan et al., 2014] Vimalanathan, K., Ramesh Babu, T. The effect of the indoor office environment on the work performance, health and well-being of office workers. J Environ Health Sci Engineer 12, 113 (2014).

[Seppänen et al., 2006] Seppänen, O., Fisk, W., & Lei, Q. (2006). Room temperature and productivity in office work. Lawrence Berkeley National Laboratory. Retrieved from https://escholarship.org/uc/item/9bw3n707

[Velt et al., 2017] K.B. Velt, H.A.M. Daanen, Thermal sensation and thermal comfort in changing environments, Journal of Building Engineering, Volume 10, 2017, Pages 42-46, ISSN 2352-7102