**1 BACKGROUND AND RELATED WORK**

Проблема загрязненности воздуха в закрытых пространствах, таких как офисные здания, учебные аудитории, концертные залы существует давно. Еще в 2001 году Dimitroulopoulou разработал динамическую модель для описания физических процессов, определяющих концентрации загрязняющих веществ в воздухе помещений в зависимости от концентрации загрязняющих веществ на открытом воздухе (Dimitroulopoulou C. et al. Modelling of indoor exposure to nitrogen dioxide in the UK //Atmospheric Environment. – 2001. – Т. 35. – №. 2. – С. 269-279.)

За последние несколько лет потребности людей в хорошем качестве воздуха внутри помещений многократно возросли, поскольку пандемия COVID-19 простимулировала множество людей перейти на удаленную работу в городских центрах, в которых загрязнение наружного воздуха является одной из крупнейших в мире проблем со здоровьем и окружающей средой (Salmatonidis A. et al. Workplace exposure to nanoparticles during thermal spraying of ceramic coatings //Annals of work exposures and health. – 2019. – Т. 63. – №. 1. – С. 91-106.)

(Lasithiotakis M. et al. Heavy metals inhalation exposure analysis from particulate matter emitted from dry and wet recycling processes of waste electrical and electronic equipment //Environmental Progress & Sustainable Energy. – 2019. – Т. 38. – №. 6. – С. e13265.)

Кроме этого, существуют неправильно спроектированные здания, в которых используется лишь малая доля уличного воздуха, а основные воздушные массы просто рециркулируются в процессе использования

(Quang T. N. et al. Influence of ventilation and filtration on indoor particle concentrations in urban office buildings //Atmospheric environment. – 2013. – Т. 79. – С. 41-52.)

(Jurado S. R., Bankoff A. D. P., Sanchez A. Indoor air quality in Brazilian universities //International journal of environmental research and public health. – 2014. – Т. 11. – №. 7. – С. 7081-7093.)

Для оценки качества воздуха внутри помещений (IAQ) используются множество параметров, такие как влажность и температура (Wolkoff P., Kjærgaard S. K. The dichotomy of relative humidity on indoor air quality //Environment international. – 2007. – Т. 33. – №. 6. – С. 850-857.), взвеси твердых частиц (Horemans B., Van Grieken R. Speciation and diurnal variation of thoracic, fine thoracic and sub-micrometer airborne particulate matter at naturally ventilated office environments //Atmospheric Environment. – 2010. – Т. 44. – №. 12. – С. 1497-1505.) и органические смеси (Fraser A. J. et al. Polyfluorinated compounds in dust from homes, offices, and vehicles as predictors of concentrations in office workers' serum //Environment international. – 2013. – Т. 60. – С. 128-136.)

(Watkins D. J. et al. Associations between PBDEs in office air, dust, and surface wipes //Environment international. – 2013. – Т. 59. – С. 124-132.)

(Lappalainen S. et al. Indoor air particles in office buildings with suspected indoor air problems in the Helsinki area //International journal of occupational medicine and environmental health. – 2013. – Т. 26. – №. 1. – С. 155-164.)

(Salonen H. J. et al. Airborne concentrations of volatile organic compounds, formaldehyde and ammonia in Finnish office buildings with suspected indoor air problems //Journal of Occupational and Environmental Hygiene. – 2009. – Т. 6. – №. 3. – С. 200-209.)

В настоящей работе рассматривается один из компонентов загрязнения воздуха, органическая смесь диоксида углерода.

Бурный рост промышленности и человеческая деятельность оказывают существенное влияние на увеличение концентрации углекислого газа в воздухе (Pachauri R. K. et al. Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change. – Ipcc, 2014. – С. 151.)

(Sinyak Y. Global climate and energy systems //Science of the total environment. – 1994. – Т. 143. – №. 1. – С. 31-51.)

Исходя из исследования IPCC [1] концентрация углекислого газа в атмосфере увеличилась с 300 до 390 ppm.

Увеличение концентрации углекислого газа в атмосфере влечет за собой увеличение концентрации углекислого газа внутри зданий, а соответственно, и увеличение энергетических затрат на обслуживание помещений (Leivo V. et al. Indoor thermal environment, air exchange rates, and carbon dioxide concentrations before and after energy retro fits in Finnish and Lithuanian multi-family buildings //Science of the Total Environment. – 2018. – Т. 621. – С. 398-406.)

Существует три типа вентиляционных систем, предназначенных для контроля за IAQ в офисных зданиях и коммерческих помещениях. Первый тип — это естественная конвекционная вентиляция, обеспечивающая движение воздушных масс через двери и окна. Второй тип — это механические вентиляционные системы, называемые HVAC (heating and ventilating air conditioning systems). Третий тип является смесью первых двух типов, т.е. включает в себя как механическую вентиляцию помещений посредствам HVAC так и естественную конвекционную вентиляцию, производимую через окна и двери (Irga P. J., Torpy F. R. Indoor air pollutants in occupational buildings in a sub-tropical climate: comparison among ventilation types //Building and Environment. – 2016. – Т. 98. – С. 190-199.)

**Экономическая забота**

В настоящей работе планируется использовать третий тип вентиляционных систем, поскольку HVAC системы являются наилучшими для обеспечения хорошего самочувствия сотрудников. Однако, использование HVAC влечет за собой экономические издержки.

По данным Управления энергетической информации США до 35% потребления энергии приходится на промышленные, коммерческие и жилые помещения строительного сектора (Pang Z. et al. How much HVAC energy could be saved from the occupant-centric smart home thermostat: A nationwide simulation study //Applied Energy. – 2021. – Т. 283. – С. 116251.)

Наибольшими потребителями электроэнергии являются heating, ventilation, and air conditioning (HVAC) systems (Papadopoulos S. et al. Rethinking HVAC temperature setpoints in commercial buildings: The potential for zero-cost energy savings and comfort improvement in different climates //Building and Environment. – 2019. – Т. 155. – С. 350-359.)

Очевидно, что HVAC это источник избыточных энергетических трат, которые могут быть предотвращены путем оптимизации режимов работы и контролем производительности без нанесения этими действиями ущерба для indoor occupants.

Опираясь на результаты (Zhu M. et al. An occupant-centric air-conditioning system for occupant thermal preference recognition control in personal micro-environment //Building and Environment. – 2021. – Т. 196. – С. 107749.) эффективная работа HVAC системы в большей мере определяется оптимизацией ее параметров управления. Исследователи (Jazizadeh F., Joshi V., Battaglia F. Adaptive and distributed operation of HVAC systems: Energy and comfort implications of active diffusers as new adaptation capacities //Building and Environment. – 2020. – Т. 186. – С. 107089.) сделали вывод, что наиболее оптимальным, надежным и выгодным решением является улучшение алгоритма управления системой HVAC, чем замена его на более энергоэффективные и дорогие аналоги. Авторы (Schito E. et al. Multi-objective optimization of HVAC control in museum environment for artwork preservation, visitors’ thermal comfort and energy efficiency //Building and Environment. – 2020. – Т. 180. – С. 107018.) показали, что умное управление HVAC системой может экономить до 30% электроэнергии при сохранении уровня комфорта сотрудников. Усовершенствование алгоритмов работы HVAC систем является приоритетной задачей при решении оптимизационных задач повышения энергоэффективности вентиляционных систем (Reena K. E. M., Mathew A. T., Jacob L. A flexible control strategy for energy and comfort aware HVAC in large buildings //Building and Environment. – 2018. – Т. 145. – С. 330-342.)

(Taheri S. et al. Stochastic framework for planning studies of energy systems: a case of EHs //IET Renewable Power Generation. – 2020. – Т. 14. – №. 3. – С. 435-444.)

Существует множество методов контроля режимов работы HVAC систем. Всю совокупность существующих методов обычно классифицируют на три больших группы (Taheri, S., & Razban, A. (2021). Learning-based CO2 concentration prediction: Application to indoor air quality control using demand-controlled ventilation. *Building and Environment*, *205*, 108164. [https://doi.org/https://doi.org/10.1016/j.buildenv.2021.108164](https://doi.org/https:/doi.org/10.1016/j.buildenv.2021.108164)): CCM (classical control methods), ICA(intelligent control approaches), MPC (model predictive control). В группу CCM входят такие методы как:on-off, proportional, PI (proportional-integral), PID (proportional-integral derivative controllers. Классические методы контроля используются для for indoor temperature control [Yang S. et al. Model predictive control with adaptive machine-learning-based model for building energy efficiency and comfort optimization //Applied Energy. – 2020. – Т. 271. – С. 115147.], dynamic control of supply air pressure [Zhao T. et al. Online differential pressure reset method with adaptive adjustment algorithm for variable chilled water flow control in central air-conditioning systems //Building Simulation. – Tsinghua University Press, 2021. – Т. 14. – №. 5. – С. 1407-1422.], cooling coil unit control [Hurt R., Wang G., Song L. Experimental Validation of Cooling Coil Control Valve Performance with Cascade Control //ASHRAE Transactions. – 2020. – Т. 126. – №. 2.], management of supply air temperature [Ahamed M. D. S. et al. Gray-box virtual sensor of the supply air temperature of air handling units //Science and Technology for the Built Environment. – 2020. – Т. 26. – №. 8. – С. 1151-1162.], evaporator supply heat control [Wu C., Xingxi Z., Shiming D. Development of control method and dynamic model for multi-evaporator air conditioners (MEAC) //Energy conversion and management. – 2005. – Т. 46. – №. 3. – С. 451-465.], and control of variable air volume unit temperature [Li X. et al. Predication control for indoor temperature time-delay using Elman neural network in variable air volume system //Energy and Buildings. – 2017. – Т. 154. – С. 545-552.]. Существенным недостатком классического метода управления является невозможность работать с временными задержками, вследствие этого такие системы начинают работать несогласованно. Для решения данной проблемы используются методы, основанные на искусственном интеллекте [Fakhari I. et al. Design and tri-objective optimization of a hybrid efficient energy system for tri-generation, based on PEM fuel cell and MED using syngas as a fuel //Journal of Cleaner Production. – 2021. – Т. 290. – С. 125205.], [Fakhari I. et al. Comparative double and integer optimization of low-grade heat recovery from PEM fuel cells employing an organic Rankine cycle with zeotropic mixtures //Energy Conversion and Management. – 2021. – Т. 228. – С. 113695.]. Нейросетевой подход применялся авторами в [Deng Z., Chen Q. Reinforcement learning of occupant behavior model for cross-building transfer learning to various HVAC control systems //Energy and Buildings. – 2021. – Т. 238. – С. 110860.] для прогнозирования температурного мультизонного комфорта и поведения occupants. В целях энергосбережения компании часто нарушают стандарты качества воздуха в помещениях и из-за этого в офисных помещениях, учебных аудиториях и концертных залах часто наблюдаются повышенные концентрации углекислого газа.

**Забота о здоровье сотрудников**

Большинство людей проводят более 90% своего времени в помещениях, это может быть как работа в офисных зданиях, так и работа на дому (McCreddin A. et al. Personal exposure to air pollution in office workers in Ireland: measurement, analysis and implications //Toxics. – 2013. – Т. 1. – №. 1. – С. 60-76..)

(Dimitroulopoulou C. et al. Modelling of indoor exposure to nitrogen dioxide in the UK //Atmospheric Environment. – 2001. – Т. 35. – №. 2. – С. 269-279. – было 1 в )

 В настоящее время большинство зданий представляют собой герметичные оболочки для сохранения и сбережения тепла и энергии внутри, поэтому крайне актуально иметь в этих зданиях адекватную систему вентиляции (<https://www.sciencedirect.com/science/article/pii/S0360132316302657>)

(<https://www.sciencedirect.com/science/article/pii/S0360132318301410>)

(<https://www.sciencedirect.com/science/article/pii/S0306261917317610>)

Кроме этого, нельзя забывать и о старых восстановленных зданиях. Результат исследования (Fauzan N. H., Jalaludin J., Chua P. Indoor Air Quality and Sick Building Syndrome (SBS) among Staff in Two Different Private Higher Learning Institution Settings in Kuala Lumpur and Selangor //International Journal of Applied Chemistry. – 2016. – Т. 12. – №. 1. – С. 57-61.) показал, что в старых восстановленных зданиях уровень концентрации CO2, PM10, bacteria существенно выше, чем в новых зданиях, соответственно, необходимость в адекватной вентиляционной системе для таких типов зданий огромная

Неэффективная система вентиляции может привести к скоплению углекислого газа в помещениях и по достижению уровня концентрации СО2 выше чем 1000ppm существенно усиливать психоэмоциональные симптомы, такие как головные боли, трудности концентрации внимания, слабость и раздражительность (N. Muscatiello, A. McCarthy, C. Kielb, W.H. Hsu, S.A. Hwang, S. Lin

Classroom conditions and CO2 concentrations and teacher health symptom reporting in 10 New York State Schools

Indoor Air-Int. J. Indoor Air Qual. Clim., 25 (2) (2015), pp. 157-167, 10.1111/ina.12136.)

[Persily A. Challenges in developing ventilation and indoor air quality standards: The story of ASHRAE Standard 62 //Building and Environment. – 2015. – Т. 91. – С. 61-69.]

Более того, последние научные исследования показали [Lu X. et al. Energy and ventilation performance analysis for CO2-based demand-controlled ventilation in multiple-zone VAV systems with fan-powered terminal units (ASHRAE RP-1819) //Science and Technology for the Built Environment. – 2020. – Т. 27. – №. 2. – С. 139-157.] что классические стратегии управления системами вентиляции не могут обеспечить регламентированный ASHRAE Standard 62.1 2019  уровень качества воздуха в помещениях. Еще одной проблемой является избыточная концентрация углекислого газа в переполненных учебных аудиториях школ в Калифорнии, США [Chan W. R. et al. Ventilation rates in California classrooms: Why many recent HVAC retrofits are not delivering sufficient ventilation //Building and Environment. – 2020. – Т. 167. – С. 106426.]

В работе (Erdmann C. A., Steiner K. C., Apte M. G. Indoor carbon dioxide concentrations and sick building syndrome symptoms in the BASE study revisited: Analyses of the 100 building dataset. – Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2002. – №. LBNL-49584.)

было исследован датасет из 100 зданий в США, в результате были получены статистически значимые результаты о наличии взаимосвязи между концентрацией углекислого газа в помещениях и симптомами SBS.

Исходя из всего вышеперечисленного крайне актуальной является проблема автоматического динамического регулирования уровня углекислого газа в помещениях, как одного из источников загрязнения воздуха.

Решение данной проблемы включает в себя решения двух подзадач: задача прогнозирования уровня углекислого газа в помещениях и задача своевременного регулирования уровня концентрации углекислого газа исходя из результатов прогноза.

**1 BACKGROUND AND RELATED WORK**

Задача прогнозирования уровня углекислого газа решалась несколькими коллективами ученых. Один из подходов решения данной задачи описан в [Ryu S. H., Moon H. J. Development of an occupancy prediction model using indoor environmental data based on machine learning techniques //Building and Environment. – 2016. – Т. 107. – С. 1-9.]. Авторы использовали технику обучения с учителем. Обучение с учителем происходит с использованием учебных данных с метками о принадлежности к классу. Каждый пример учебных данных является парой, состоящей из входного объекта и желаемого выходного результата. В данном исследовании был выбран алгоритм машинного обучения называемый The Classification and Regression Tree (CART) для предсказания количества occupants и the Hidden Markov Model (HMM) для определения количества углекислого газа в воздухе. Алгоритм CART осуществляет построение бинарного дерева решений, каждый узел которого при разбиении имеет только двух потомков.

Рассмотрим алгоритм CART подробнее. Для алгоритма CART поведение объектов выделенной группы означает долю модального значения выходного признака. Выделенные группы - доля модального значения выходного признака достаточно высока. На каждом шаге построения дерева в узле формируется правило, которое делит заданное множество примеров на две части: в первой части правило выполняется, её называют правым потомком, часть, в которой правило не выполняется называют левым потомком. Оценка качества модели базируется на оценочной функции, основанной на идее уменьшения энтропии в узле. Одним из примеров оценочных функций является индекс Gini. Если набор данных T содержит данные n классов, тогда индекс Gini можно определить

|  |  |
| --- | --- |
|  | (1) |

, где вероятность класса i в T. Если набор T разбивается на две части и с числом параметров в каждом и соответственно, тогда показатель качества разбиения будет равен

|  |  |
| --- | --- |
|  | (2) |

Наилучшим считается то разбиение, для которого минимально. Обозначим N – число примеров в узле-предке, L, R – число объектов соответственно в левом и правом потомке, и ri – число экземпляров i-го класса в левом и правом потомке. Тогда качество разбиения можно оценить по следующей формуле:

|  |  |
| --- | --- |
|  | (3) |

Для уменьшения объема вычислений данную формулу преобразуют к виду

|  |  |
| --- | --- |
|  | (4) |

Так как умножение на константу не играет роли при минимизации:

|  |  |
| --- | --- |
|  | (5) |
|  |  |
|  | (6) |
|  |  |
|  | (7) |
|  |  |

В результате, лучшим будет то разбиение, для которого величина максимальна. Это значит, что при построении CART ищется такой вариант ветвления, при котором максимально уменьшается значение

Помимо CART в данном исследовании использовались HMM.

Рассмотрим HMM подробнее. HMM это статистическая марковская модель, которая моделирует поведение системы, представляющей собой марковский процесс с неизвестными параметрами.

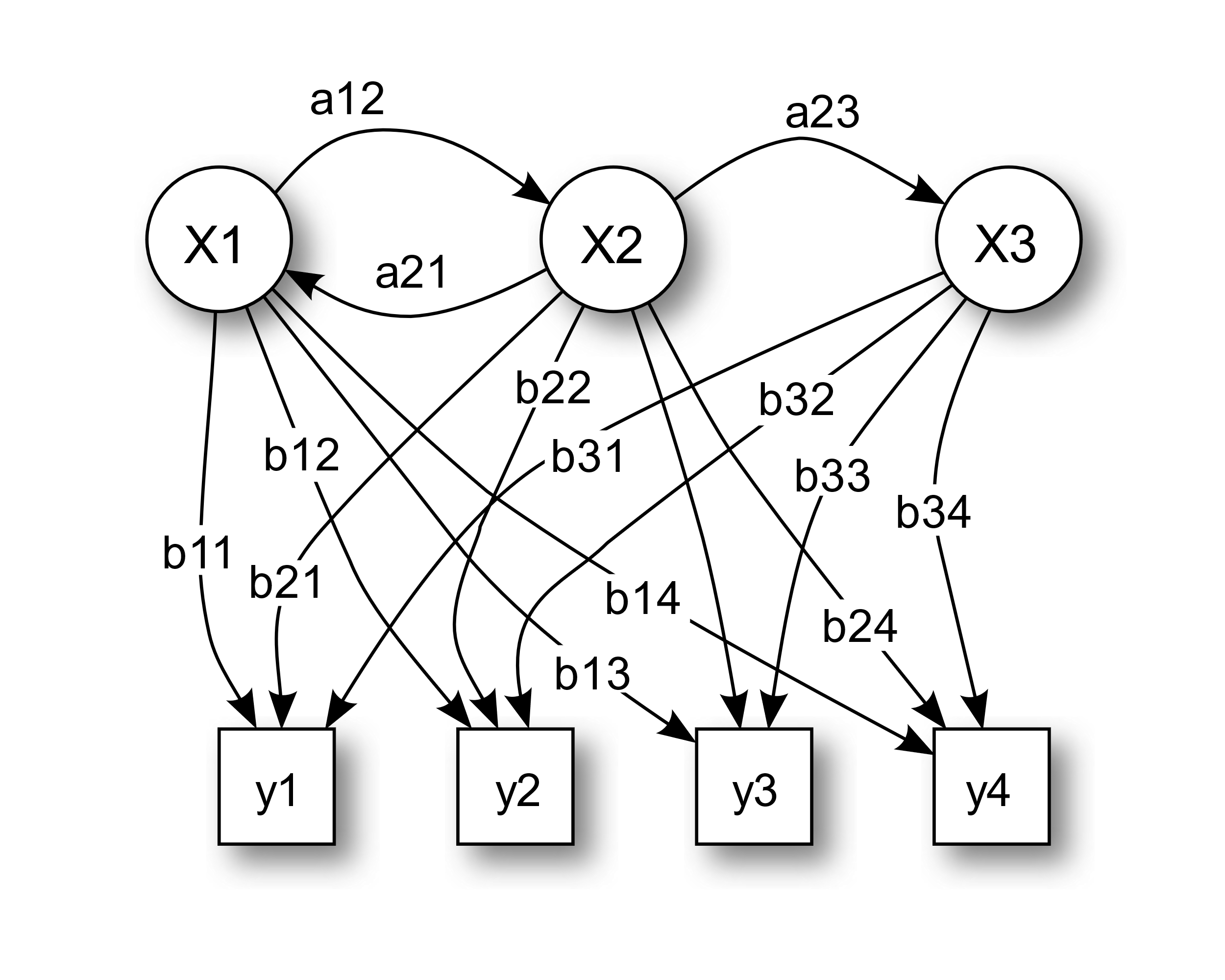


Fig2.  Probabilistic parameters of a hidden Markov model (example) [https://en.wikipedia.org/wiki/Hidden\_Markov\_model]

*X* — states  
*y* — possible observations  
*a* — state transition probabilities  
*b* — output probabilities

The diagram below shows the general architecture of an instantiated HMM. Each oval shape represents a random variable that can adopt any of a number of values. The random variable *x*(*t*) is the hidden state at time *t* (with the model from the above diagram, *x*(*t*) ∈ { *x*1, *x*2, *x*3 }). The random variable *y*(*t*) is the observation at time *t* (with *y*(*t*) ∈ { *y*1, *y*2, *y*3, *y*4 }). The arrows in the diagram (often called a trellis diagram) denote conditional dependencies.

From the diagram, it is clear that the conditional probability distribution of the hidden variable x(t) at time t, given the values of the hidden variable x at all times, depends only on the value of the hidden variable x(t − 1); the values at time t − 2 and before have no influence. This is called the Markov property. Similarly, the value of the observed variable y(t) only depends on the value of the hidden variable x(t) (both at time t).

Рассмотрим стандартную HMM, скрытые состояния являются дискретными, наблюдаемые параметры могут быть также дискретными, если представляют собой категориальные распределения или непрерывными, если представляют собой непрерывные распределения, например, распределение Гаусса. Есть два типа параметров в HMM, первый тип это transition probabilities. Этот тип параметров отвечает за вероятности перехода из состояния со временем t-1 в состояние t. Второй тип параметров называется emission probabilities (also known as output probabilities) эти параметры определяют вероятность получения данной наблюдаемой в зависимости от состояния системы.

Скрытое пространство состоит из N возможных значений, имеющим категориальное распределение. Это означает, что для каждого из N возможных значений из скрытого пространства в момент времени t существует N transition probabilities в каждое скрытое состояние с моментом времени t+1, таким образом всего существует N2 transition probabilities. Таким образом, мы имеем Марковскую матрицу размера NxN transition probabilities Because any transition probability can be determined once the others are known, there are a total of N\*(N-1) transition parameters. Для каждого из N возможных значений существует набор of emission probabilities определяющий вероятность возникновения наблюдаемой переменной в зависимости от того, в каком скрытом состоянии находится данная система в этот момент времени. Для оценки сложности рассмотрим для примера M возможных наблюдаемых параметров, распределенных дискретно. Тогда в системе будет M-1 отдельных параметра, а общее число emission parameters будет равно N(M-1), если М-мерный вектор наблюдаемых параметров распределен как многомерное Гауссовское распределение, то у нас будет M параметров для контроля за средними значениями и параметров для контроля за матрицей ковариации, таким образом всего параметров для контроля , что приводит к сложности порядка при контроле за emission parameters. Поскольку сложность зависит от M квадратично, то при больших значениях M накладывают искусственные ограничения на матрицу ковариаций, подразумевая, например, что элементы независимы друг от друга.

В рассматриваемой статье скрытыми состояниями марковской модели являлись спрогнозированные величины количества occupancy посредствам CART. Основным преимуществом HMM является наличие временной корреляции во временных рядах. По этой причине для прогнозирования потребления углекислого газа в данном исследовании была выбрана модель HMM. Авторам удалось получить точность предсказания концентрации CO2 в indoor на уровне 89.5%. Однако, у этого исследования есть ряд недостатков. Один из которых, это исключительные условия эксперимента, проводимые на испытательном стенде Building Integrated Control Test-bed (BICT) at Dankook University, Korea Fig. 1.



Fig1. Building Integrated Control Test-bed (BICT) used in the development of the occupancy prediction model [последняя цитата].

Поэтому неизвестно, может ли быть применена данная модель в реальных условиях: в офисных зданиях, учебных аудиториях или концертных залах.

Еще одним подходом к решению задачи прогнозирования концентрации углекислого газа является подход, основанный на long-short term memory (LSTM) архитектуре нейронной сети. LSTM является особой разновидностью архитектуры рекуррентных нейронных сетей, способных к обучению долговременным зависимостям. LSTM сети имеют преимущество над обычными рекуррентными нейронными сетями, поскольку последние страдают проблемами исчезающего градиента[Karim F. et al. LSTM fully convolutional networks for time series classification //IEEE access. – 2017. – Т. 6. – С. 1662-1669.]. Hochreiter and Schmidhuber в своём исследовании [Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.] показали, что LSTM RNNs address the vanishing gradient problem commonly found in ordinary recurrent neural networks by incorporating gating functions into their state dynamics. При каждой итерации LSTM содержит скрытый вектор **h** и вектор памяти **m**, который отвечает за состояние обновлений выходов. Более конкретно Kawakami *[*Kawakami, Kazuya. "Supervised sequence labelling with recurrent neural networks." *Ph. D. thesis* (2008).*]*определили вычислительные итерации LSTM на временном шаге t как показано ниже:

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  | (8) |
|  |  |
|  |  |

where σ is the logistic sigmoid function, ⊙ represents elementwise multiplication, Wu,Wf,Wo,Wc are recurrent weight matrices and Iu,If,Io,Ic are projection matrices.

В своей статье [Elkhoukhi H. et al. On the use of deep learning approaches for occupancy prediction in energy efficient buildings //2019 7th International Renewable and Sustainable Energy Conference (IRSEC). – IEEE, 2019. – С. 1-6.] авторы использовали LSTM для прогнозирования уровня концентрации углекислого газа в зависимости от количества людей в помещении, обучаясь на исторических данных. Предсказания проводились для короткого горизонта предсказаний. Однако, полученная точность предсказания не превышает 70%, что является сравнительно небольшим показателем. Доработка архитектуры и методики проведения эксперимента могут быть рассмотрены как одни из возможных задач, решаемых в нашем исследовании.

В работе [Dong B., Andrews B. Sensor-based occupancy behavioral pattern recognition for energy and comfort management in intelligent buildings //Proceedings of building simulation. – 2009. – С. 1444-1451.] авторы разрабатывали и применяли алгоритмы основанные на сенсорном моделировании с помощью которых можно предсказывать пользовательское поведение в зданиях. После этого полученные закономерности внедрялись в системы управления комфортом в зданиях и проводились симуляции потребления энергии этими системами. Посредством проведённых симуляций авторы показали, что существует потенциал снижения потребления электроэнергии до 30% без потери комфортабельности рабочих мест. Схема исследования включала в себя применение HMM для предсказания энергозатрат и комфортабельности. В качестве наблюдаемых параметров использовались данные с сенсоров, представляющие собой измерения уровня CO2, температуры и относительной влажности в офисных помещениях. В результате данной работы было показано, что использование машинного обучения для прогнозирования выбросов CO2 было полезно, однако нужно учитывать, что данная статья несколько устарела (год выпуска статьи в печать 2009г), поэтому нельзя точно сказать, насколько лучше или хуже данный подход в современных реалиях. Поэтому одной из возможных задач для настоящей работы может быть задача проведения всестороннего исследования по сравнению применимости для данной проблемы современных алгоритмов машинного обучения.

Еще одна стратегия по оценке концентрации углекислого газа в помещении представлена в статье Paige WenbinTien et. al [Tien P. W. et al. A vision-based deep learning approach for the detection and prediction of occupancy heat emissions for demand-driven control solutions //Energy and Buildings. – 2020. – Т. 226. – С. 110386.]. Исследователи решили отказаться от статистического подхода оценки занятости офисных работников, фиксированного расписания работы и исторической информации о времени работы. Вместо этого, в работе был предложен подход по оценке занятости и активности сотрудников. Основная идея данной работы заключена в том, чтобы в режиме реального времени определять активность сотрудников во время рабочего дня используя камеры. Исследователи использовали модель глубокого обучения, основанную на сверточной нейронной сети. Общая архитектура нейронной сети, используемой в исследовании, представлена на картинке ниже

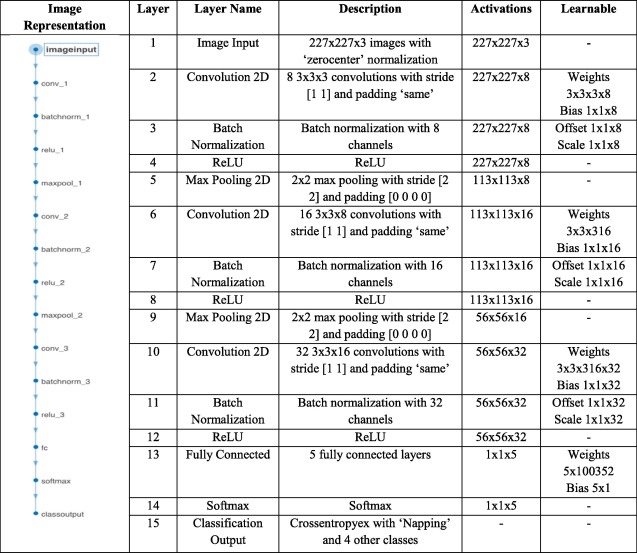


Fig3 Архитектура 15 слойной сверточной нейронной сети, используемой в исследовании[там же]

Исследователям удалось разработать метод глубокого обучения, позволяющий определять основные виды деятельности офисных работников в типичном офисном здании со средней точностью обнаружения 80.62%.

Получив достаточно высокую точность в определении численности occupants и их офисной активности авторы предложили использовать эту информацию для более точного прогнозирования и оценки концентрации уровня углекислого газа в здании. Поскольку уровень углекислого газа в помещениях является одним из факторов, определяющих качество воздуха внутри помещений [Pathirana S. M., Sheranie M. D. M., Halwatura R. U. Indoor thermal comfort and Carbon Dioxide concentration //2017 Moratuwa Engineering Research Conference (MERCon). – IEEE, 2017. – С. 401-406.], то при решении задачи оптимизации качества воздуха внутри помещений, авторы предложили использовать разработанную ими модель для оценки данного параметра.

Для тестирования методов авторы смоделировали помещение одного из офисных зданий в Nottingham, UK и, используя симуляцию офисных процессов в здании, проводили building energy simulation (BES).

Используя BES (fig 4.) моделирование автором удалось показать, что уровень концентрации углекислого газа в воздухе напрямую зависит от количества находящихся в помещении людей, а также от того, какие активности офисный работники выполняют внутри. Данный вывод указывает на высокий потенциал использования данного метода для прогнозирования и регуляции уровня углекислого газа в офисных помещениях.

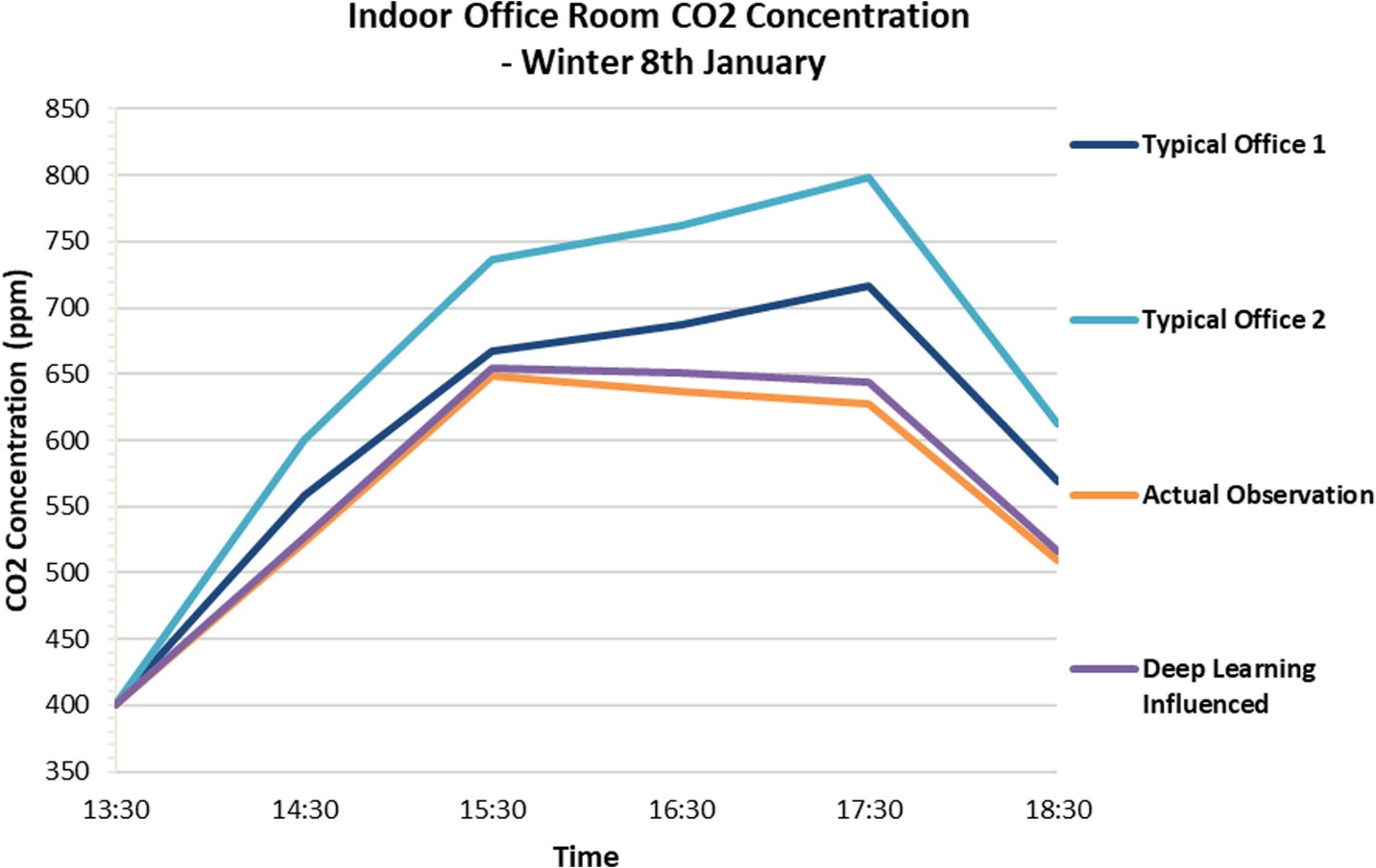


Fig. 4. CO2 Concentration for a winter day (8th January) at the open-plan office space with four occupants; based on the two Typical Office Profiles, DLIP and Actual Observation Profile[там же]

На рис.4 представлен график уровня концентрации углекислого газа в офисном помещении, рассчитанный на основании того, что внутри помещения находится 4 человека. По графику видно, что разница в уровне концентрации CO2 между реальным офисом и нейросетевым предсказанием может достигать 248.8 ppm. Этот факт указывает на то, что для оценки уровня концентрации углекислого газа в помещении недостаточно использовать только количество людей и их активность.

Другие исследователи в своей работе [Zhu H. C., Ren C., Cao S. J. Fast prediction for multi-parameters (concentration, temperature and humidity) of indoor environment towards the online control of HVAC system //Building Simulation. – Tsinghua University Press, 2021. – Т. 14. – №. 3. – С. 649-665.] рассказывают о комбинированной модели машинного обучения и алгоритмом управления вентиляционной системой для улучшения качества воздуха в помещениях. В своей работе авторы использовали искусственную нейронную сеть для предсказания уровня концентрации углекислого газа в помещениях. Алгоритмы управления вентиляционными системами представляют собой алгоритмы, базируемые на contribution ratio of indoor climate. В своей работе авторы обнаружили, что используя сильные стратегии в управлении HVAC системами, можно сэкономить до 35% потребления электроэнергии. В нашем случае эта работа может быть интересна тем, что одним из регулируемых параметров является углекислый газ, регулирование которого мы хотим осуществить в нашем исследовании. Однако, основным недостатком этой работы является то, что показатели загрязнённости и температур были симулированы, поэтому у нас нет четкого понимания применимости данных технологий в реальных средах.

Одной из наиболее интересных для нас работ является недавнее исследование Kallio et al. по прогнозированию уровня концентрации углекислого газа в офисных помещениях методами машинного обучения на годичном наборе данных [Kallio J. et al. Forecasting office indoor CO2 concentration using machine learning with a one-year dataset //Building and Environment. – 2021. – Т. 187. – С. 107409.]. В своей статье авторы использовали 4 метода машинного обучения для прогнозирования уровня концентрации CO2 внутри помещений. Это были методы Ridge regression, Decision Tree, Random Forest, and Multilayer Perceptron. В ходе данного литературного обзора мы уже рассматривали метод решающих деревьев CART. Остановимся подробнее на трех других методах машинного обучения.

Первый метод, который мы рассмотрим называется Ridge regression. Если углубиться немного в историю, то можно увидеть, что изначально метод назывался методом регуляризации Тихонова, метода назвали в честь Советского и Российского математика Андрея Николаевича Тихонова. Ridge regression это метод понижения размерности. Применяется для борьбы с избыточностью данных, когда независимые переменные коррелируют друг с другом, вследствие чего проявляется неустойчивость оценок коэффициентов многомерной линейной регрессии. Введем понятие мультиколлинеарности. multicollinearity — наличие линейной зависимости между независимыми переменными регрессионной модели. Различают полную коллинеарность и частичную или просто мультиколлинеарность — наличие сильной корреляции между независимыми переменными.

Рассмотрим пример линейной модели:

|  |  |
| --- | --- |
|  | (9) |
|  |  |

При этом имеется зависимость между переменными:

|  |  |
| --- | --- |
|  | (10) |
|  |  |

Добавим тогда к первому коэффициенту произвольное число , при этом из двух других коэффициентов вычтем это же число.

Исключая случайную ошибку, получим

|  |  |
| --- | --- |
|  | (11) |
|  |  |

Несмотря на относительно произвольное изменение коэффициентов модели мы получили исходную модель, то есть такая модель неидентифицируема. На практике чаще встречается проблема сильной корреляции между независимыми переменными. В этом случае оценки параметров модели получить можно, но они будут неустойчивыми.

Рассмотрим задачу многомерной линейной регрессии:

Дана линейная зависимость вида:

|  |  |
| --- | --- |
|  | (12) |
|  |  |

Тогда нужно найти вектор при котором достигается минимум среднего квадрата ошибки:

|  |  |
| --- | --- |
|  | (13) |

|  |  |
| --- | --- |
|  | (14) |

Находя решение методом наименьших квадратов, получим:

|  |  |
| --- | --- |
|  | (15) |

В условиях мультиколлинеарности(multicollinearity) матрица становится плохо обусловленной.

Для решения данной проблемы накладываем ограничение на величину коэффициентов:

|  |  |
| --- | --- |
|  | (16) |

Функционал с учетом ограничения принимает вид:

|  |  |
| --- | --- |
|  | (17) |

, где λ – неотрицательный параметр.

Решением в этом случае будет

|  |  |
| --- | --- |
|  | (18) |

Это изменение увеличивает собственные значения матрицы , но не изменяет её собственные вектора. В результате получаем хорошо обусловленную матрицу.

Следующий метод, который мы рассмотрим, называется Multilayer perceptron (MLP), представленный на рис 4.

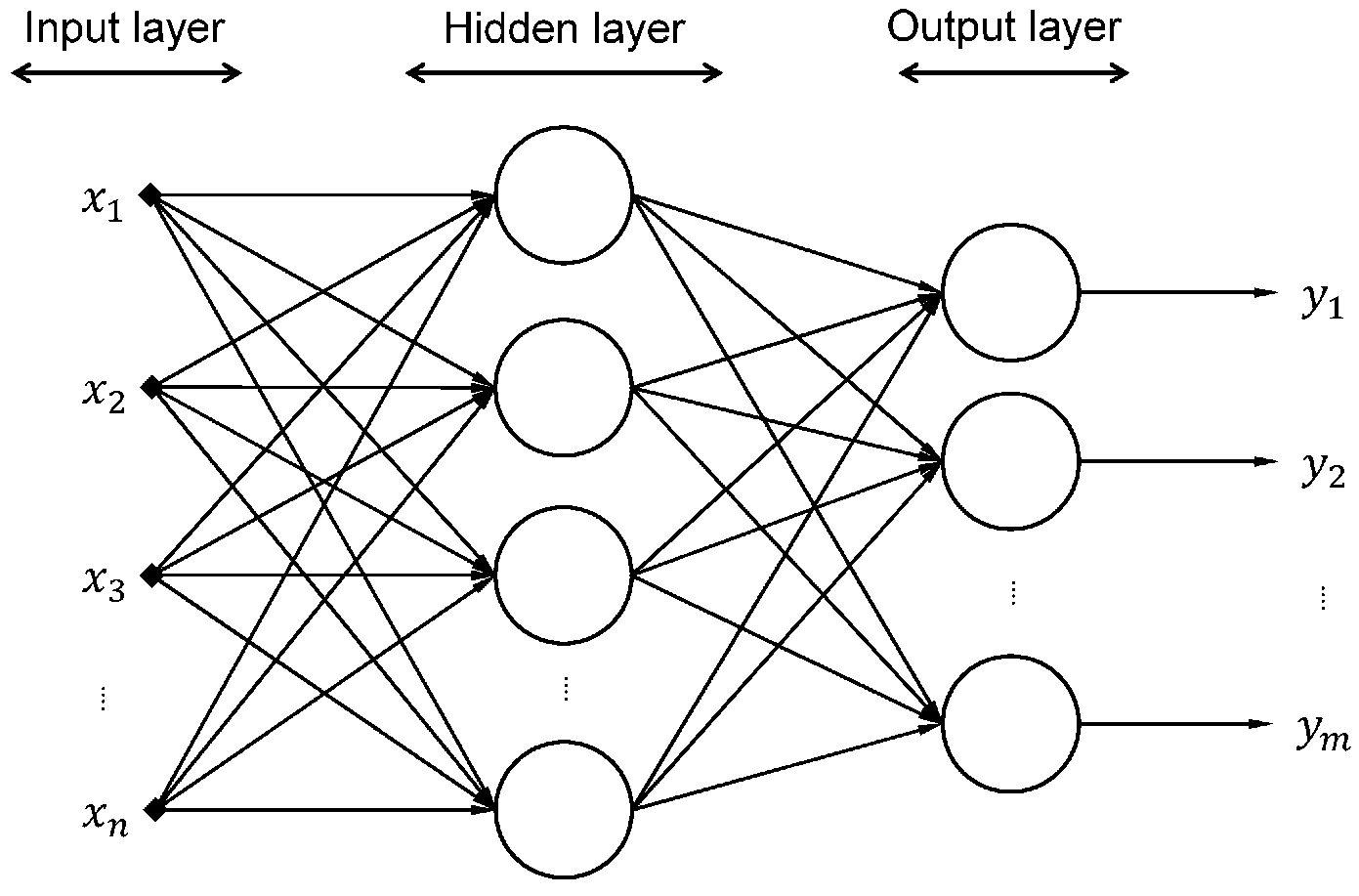


Fig 4 Multilayer perceptron. [Ryu S., Noh J., Kim H. Deep neural network based demand side short term load forecasting //Energies. – 2017. – Т. 10. – №. 1. – С. 3.]

Multilayer perceptron состоит минимум из трёх слоёв: an input layer, a hidden layer and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called backpropagation for training.

Обучение многослойного перцептрона происходит посредством изменения весов связей нейрона после обработки каждого батча данных. Изменение весов происходит путем минимизации функции ошибки, которая сигнализирует о том, насколько ожидаемый результат отличается от результата на выходе нейронной сети. Такой метод обучения называется supervised learning, метод включает в себя механизм обратного распостранения ошибки и в общем случае является обобщением метода наименьших квадратов в линейных моделях.

We can represent the degree of error in an output node in the -th training example by

|  |  |
| --- | --- |
|  | (19) |

, where is the target value and is the value produced by the perceptron. The node weights can then be adjusted based on corrections that minimize the error in the entire output, given by

|  |  |
| --- | --- |
|  | (20) |

Using gradient descent, the change in each weight is

|  |  |
| --- | --- |
|  | (21) |

, where is the output of the previous neuron and is the learning rate, which is selected to ensure that the weights quickly converge to a response, without oscillations.

The derivative to be calculated depends on the induced local field , which itself varies. It is easy to prove that for an output node this derivative can be simplified to

|  |  |
| --- | --- |
|  | (21) |

, where is the derivative of the activation function, which itself does not vary. In artificial neural networks, the activation function of a node defines the output of that node given an input or set of inputs. The analysis is more difficult for the change in weights to a hidden node, but it can be shown that the relevant derivative is

|  |  |
| --- | --- |
|  | (22) |

This depends on the change in weights of the nodes, which represent the output layer. So, to change the hidden layer weights, the output layer weights change according to the derivative of the activation function, and so this algorithm represents a backpropagation of the activation function.

У данной работы также есть ряд недостатков. К примеру, в качестве метрики качества в своём исследовании авторы используют MAE (mean absolute error)

|  |  |
| --- | --- |
|  | (23) |

, where is the prediction and the true value. Очевидно, что такая оценка неустойчива к масштабу. Поэтому, одним из вариантов развития данной тематики в моей работе может быть уточнение результатов с применением различных, устойчивых к масштабу метрик. Например, MAPE

|  |  |
| --- | --- |
|  | (24) |

А также я считаю, что для практических применений целесообразнее использовать более длительные горизонты прогнозирования с более высокими показателями метрик качества.

Экспериментальная часть

В настоящем исследовании принимали участие ### человек различных гендеров, профессиональных навыков, вероисповеданий и рас. Возрастной диапазон участников варьировался от ### до ### лет. Для участия в эксперименте не допускались подопытные, имеющие дефекты, связанные с нарушением оптической системы глаз, имеющие диагностированные психоневрологические заболевания, а также люди, имеющие цветные или тонированные контактные линзы. Кроме этого, в эксперименте не принимали участия люди, имеющие заболевания верхних и нижних дыхательных путей и лёгких.

Все участники эксперимента были проинформированы заранее и подписали информированное согласие на проведение эксперимента и обработку персональных данных.

Вывод

В этом обзоре были рассмотрены некоторые различные подходы к проблеме скопления и удаления излишек воздух загрязняющих веществ в офисных зданиях, учебных аудиториях и концертных залах. Основными задачами, решаемыми в процессе работы являются задача прогнозирования уровня углекислого газа в зависимости от числа сотрудников, их активностей, рабочего расписания и задача динамического управления HVAC системами для управления климатом в здании. Помимо этого, были рассмотрены исследования, показывающие уменьшения потребления электроэнергии при динамическом управлении HVAC до 35%. Однако, во всех рассмотренных работах авторы рассматривают поддержание комфортного климата внутри помещения посредством механической вентиляции HVAC. Однако, регуляцию внутренней среды можно также осуществлять и посредством естественной вентиляции. Очевидно, что комбинация HVAC и естественного способа вентиляции для решения задач терморегуляции и поддержки комфортного уровня концентрации углекислого газа внутри помещений энергетически наиболее эффективна, поскольку естественное движение воздушных потоков из помещения наружу и обратно не требует энергетических затрат.

Данный подход требует применения многопараметрических моделей, поскольку теперь при прогнозировании работы HVAC систем следует учитывать состояние воздуха снаружи, например, его температуру, влажность, уровень концентрации углекислого газа. Учитывать уровень концентрации углекислого газа наружного воздуха особенно важно, поскольку это может быть большой проблемой в крупных городских центрах или его скачков, вызванных экологическими причинами. Параметры наружного воздуха могут меняться динамически, поэтому применение только инженерного подхода не представляется возможным. Для анализа динамически изменяющихся систем следует использовать машинное обучение и методы искусственного интеллекта. Помимо очевидных экономических выгод, своевременная терморегуляция и поддержка комфортного уровня концентрации углекислого газа в помещении, вероятно, способна повысить продуктивность работы сотрудников, а значит, сократить экономические издержки компаний, учебных учреждений или концерт залов.

Настоящая магистерская работа подразумевает создание специального устройства для регулирования ширины открытия наружных окон, модели машинного обучения и алгоритмы, управляющие открытием\закрытием окон и HVAC системой для поддержки комфортной температуры и качества воздуха для людей, работающих в закрытых помещениях, таких как офисы, учебные аудитории или концертные залы.