

修士論文

IoTセンサーネットワークと
実空間における位置情報を用いた在室状況推定

Room occupancy estimation using IoT sensor network
and location information

令和5年1月26日 提出

指導教員 佐藤周行准教授

電気系工学専攻 融合情報学コース

学籍番号 37-216538

中村有佑

要旨

近年，世界各地で電力需給がひっ迫するという事態が発生している．このような事態に対処する上で，建物におけるエネルギー消費量の削減は有効な方法である．これには，室内におけるエネルギー制御を行う必要がある．室内における在室状況を適切に把握し，空調設備や電球の制御を行うことが有効である．また，センサーの小型化やスマートデバイスがネットワークに接続されることになり，IoT センサーを使って大量のデータが収集できるようになっている．一方で，収集したビッグデータの解析には，計算コストや対象データの推定精度やセンサーの位置関係に伴うネットワークボトルネックの課題がある．本研究では，まず，IoT センサーネットワークを構築する．この時のネットワークは関連する全ての IoT センサーを一台の Raspberry Pi 4 に接続する形である．作成した IoT ネットワークを用いて，IoT センサーを用いて CO₂ 濃度，温度，湿度，照度，気圧の環境情報を収集する．データの可視化を行った後に，従来の手法の評価を行う．そして，計算量の削減，推定精度の向上を目的として，提案手法の適用を行う．具体的には，収集したデータに対して実空間における位置情報を組み合わせた前処理を施し，データセットを作成する．作成したデータセットに，機械学習アルゴリズムを適用することで，高速かつ高精度な在室状況推定を実現する．また，対照実験として外れ値検知技術の LOF を用いた外れ値検知を用いて本実験の結果の比較・評価を行う．最終的に，在室人数検出精度を誤差 ± 1 人以下にするという目標を達成した．

目次

第 1 章	序論	1
1.1	はじめに	1
1.2	研究目的	2
1.3	本論文の構成	2
第 2 章	在室人数推定と関連研究	3
2.1	IoT ネットワーク	3
2.2	異常検知	4
2.3	センサーを用いた異常検知に関する研究	5
2.3.1	IoT データを用いたリアルタイム異常検知システム	5
2.3.2	非線形グラフフィルタを用いた異常検出システム	6
2.3.3	温度センサーシステムの異常検知に Prophet を使用した時間予測駆動型アプローチ	6
2.3.4	エントロピー法を用いた異常検知システム	6
2.3.5	ウェアラブルセンサーネットワークにおけるストリーミングデータに対する異常検知システム	7
2.4	IoT ベースの環境測定システムに関する研究	8
2.5	機械学習	10
第 3 章	IoT センサーネットワークの構築とデータの可視化	12
3.1	研究の段階分け	12
3.2	IoT センサーネットワークの構築	12
3.3	データの可視化	13
第 4 章	Candanedo の方法	17
4.1	予備実験の目的と概要	17
4.2	実験環境	17
4.3	実験方法	18
4.4	実験結果	19
4.5	評価	21
第 5 章	提案手法	22
5.1	環境センサーデータを用いたデータセットの作成	22
5.2	在室推定で用いる機械学習アルゴリズムと評価指標	24
第 6 章	IoT センサーネットワークと実空間座標情報を用いた在室人数推定	25
6.1	実験環境	25
6.2	実験方法	28

6.3	実験結果	28
6.4	対照実験	32
6.4.1	LOF(Local Outlier Factor) とは	32
6.4.2	対照実験の結果	32
6.5	評価	36
第 7 章	結論	37
7.1	まとめ	37
7.2	今後の展望	37
	発表文献	42

目 次

1.1	Data utilization [1]	2
2.1	Smart city based on fog architecture [2]	3
2.2	A generic architecture for different outlier detection in IoT and WSN [3]	5
2.3	Environment Monitoring System Architecture [4]	8
2.4	Environment Monitoring System Architecture [5]	9
2.5	scikit-learn algorithm cheat-sheet [6]	10
3.1	System configuration diagram	13
3.2	Connection environment	14
3.3	Graph on PC about CO2 sensor	15
3.4	Graph on PC about Temperature / humidity sensor	15
3.5	Heat map on PC about Temperature / humidity sensor	16
4.1	Connection environment for preliminary experiment	17
4.2	Measurement profiles for 1 week	19
4.3	Pairs plot	20
5.1	An example of focus and sensor positions in the experimental environment	23
6.1	Sensor placement	26
6.2	Sensor location	27
6.3	Distribution map of CO2 concentration in the training data	33
6.4	The first-order difference in the concentration of CO2 in the training data	33
6.5	Outlier detection results in the training data	33
6.6	Distribution map of CO2 concentration in the test1 data	34
6.7	The first-order difference in the concentration of CO2 in the test1 data	34
6.8	Outlier detection results in the test1 data	34
6.9	Distribution map of CO2 concentration in the test2 data	35
6.10	The first-order difference in the concentration of CO2 in the test2 data	35
6.11	Outlier detection results in the test2 data	35

表 目 次

4.1	Computing environment	18
4.2	Sensor equipment	18
4.3	Data set description	18
4.4	Correlation matrix	19
4.5	Accuracy at each parameter for each model application	20
6.1	IoT sensor equipment	25
6.2	Data set description	28
6.3	MAE at each parameter for each model application	29
6.4	MSE at each parameter for each model application	30
6.5	RMSE at each parameter for each model application	31

第1章 序論

1.1 はじめに

近年、世界各地で電力需給がひっ迫するという事態が発生している。この背景には、covid-19からの経済回復の反動として、電力需要の世界的な増加が要因の一つとして挙げられている [7]。エネルギーは主に住居や学校、ビルをはじめとした建物で消費されている。具体的には、このような建物は世界のエネルギー消費の4割を占めている。また、電力分野でのエネルギー消費の6割を占めている [8]。建物におけるエネルギー問題解決のために、建材の品質改善に関する研究が多く行われてきた。エネルギー消費の観点から考えると、建物内の居住状況は重要な要素となる。例として、電球の制御や人数に応じた空調設備の制御を行うことによって、消費エネルギーの低減が実現できる。空調設備の制御を通して、エネルギー消費量を20%以上節約した研究がある [9]。このように室内の人数を把握することはエネルギー問題の解決、消費エネルギーの節約に対して非常に重要な役割を果たす。

居住状況把握システムを作る場合、プライバシーについて配慮する必要がある [10]。カメラを設置しての居住者検出、居住者カウントを行うことは精度的な観点からは望ましいが、プライバシー的な観点から望ましくない。そこで、非侵入型モニタリングシステムが必要となる。具体的には、環境センサー(ex. 温度センサー, 湿度センサー, CO2濃度センサー, 照度センサー)や人感センサーを用いて室内の状況を把握することによって、居住者のプライバシーに配慮しつつ、室内の状況を把握することができる。同時に、データを適切に分析する必要がある。本研究では、環境センサーネットワークと実空間におけるセンサーの位置情報を用いてデータの収集を行い、収集したデータと機械学習を用いて室内の人数の把握を行う。データの収集を行う際に、IoTというアーキテクチャを用いた。以下で、IoTについての説明を行う。

通信・ネットワーク技術の発達に伴い、スマートフォン、タブレット端末、スマート家電や車をはじめとした多くのスマートデバイスがネットワークに接続されるようになってきている。このような背景からIoTという概念が生まれた。IoTとは、Internet of Thingsの略である。具体的には、インターネットを経由して、スマートデバイス同士が通信することを意味している。従来の場合、インターネットを介して通信を行うデバイスは、コンピュータとサーバのようなIT機器であった。IoTという概念が生まれてからは、IT機器のみならず、従来は組み込みシステムで動作していたエアコン、冷蔵庫、炊飯器をはじめとした家電製品や、センシングデバイスを搭載した車、温度や湿度、CO2濃度情報の収集を行い室内情報の管理に用いるセンサーもインターネットを介して相互に通信を行うようになった。これらのIoT機器をスマートフォン上のアプリを用いて遠隔操作を行うホーム・オートメーションや人感センサーやLiDARを用いた車の運転支援に活用されるスマートシティを実現する場合の基盤となるアーキテクチャとしてIoTは基本的なものである。スマートデバイスの遠隔制御や遠隔監視をはじめとして、スマートデバイス同士の通信も可能になった。このような背景から、IoT技術の応用性は非常に高いものである。IoTデバイスが収集する大量のデータを活用することは以前から考えられてきている [1]。

IoTデバイス上で大量のデータを処理する時に、計算能力が低い点が遅延の原因となるケース

がある。また、推定を行う際には、推定精度・推定速度が高いことが望ましい.. データのパラメータ数を減らすことによって計算能力の面をカバーすることができる。機械学習アルゴリズムを適切に用いることで、推定精度と速度を高めることができる。

図 1.1 は、データの活用によって、実現できることを示したものである。本研究では、見える化・自動検出・予測を対象とする。

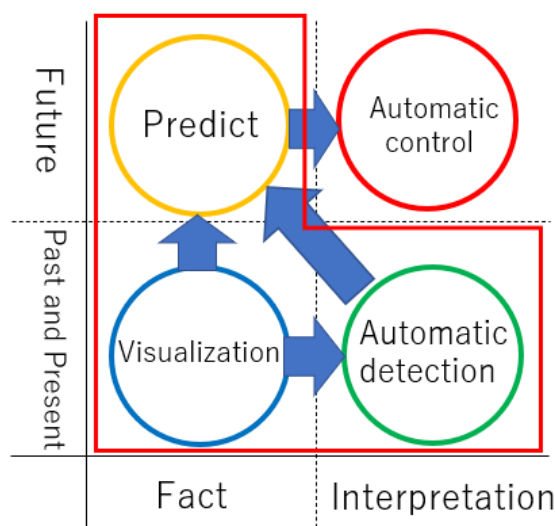


図 1.1: Data utilization [1]

1.2 研究目的

本研究の目的は、実空間座標情報に注目し、低い計算コストで高速かつ高精度に在室状況推定を行うことである。具体的には、複数の IoT センサーを用いて、部屋の CO2 濃度や温度、湿度、照度、気圧の計測を行う。そこで収集されたデータに実空間座標情報を組み合わせて前処理を行い、次元数を圧縮したデータセットの作成を行う。前処理を行う際に、伝熱工学的観点进行を参考にした提案手法を用いる。具体的な提案手法については 5 章で説明する。最後に、作成したデータセットと機械学習を用いて、在室状況推定を行う。目標は、在室人数検出精度を誤差 ± 1 人以下にすることである。一連の提案・実験を通して、関連研究において指摘されている、センサーの位置が測定値に影響を与える可能性があり、センサーやモデルを再配置するたびにモデルのトレーニングが必要となるという課題を解決する。

1.3 本論文の構成

本論文は、7 章で構成される。第 2 章では、本研究の背景知識および在室状況推定の関連研究について述べる。第 3 章では、IoT センサーネットワークの構築とデータの可視化について述べる。第 4 章では、Candanedo の方法について述べる。第 5 章では、提案手法について述べる。第 6 章では、IoT センサーネットワークと実空間座標情報を用いた在室状況推定についての実験・評価を行う。第 7 章では、本研究によって得られた実空間座標情報が在室状況推定に与える影響についてまとめ、今後の課題を述べる。

第2章 在室人数推定と関連研究

本章では，本研究における関連事項・関連研究について説明する．具体的には，第4章で構築するIoTネットワークに関する関連事項や，温度やCO₂濃度を計測するセンサーの数値データを対象とする在室人数推定のシステムに関する研究である．また，IoTベースの環境測定システムの開発も説明する．

2.1 IoT ネットワーク

IoT ネットワークを構成する要素について，説明する．まず，IoT ネットワークの構成要素は3種類に分けることができる．具体的には，クラウドレイヤー，フォグレイヤー，IoT レイヤーの3種類である．一つ目のクラウドレイヤーは，主に情報の保存や管理を行うサーバーで構成されている．本研究では，主にWebサーバ，スマートフォン，PCが該当する．二つ目のフォグレイヤーは，IoT レイヤーで収集したセンサー情報や画像の情報をコンピューター上で処理できる形式にする機能がある．これによって，IoT レイヤーとクラウドレイヤーの仲介を行う．本研究では，Raspberrypiが該当する．三つ目のIoT レイヤーでは，センサーやカメラを用いて，環境データや画像データを始めとして様々なデータの収集を行う．本研究では，温度・湿度・気圧・照度・気圧・CO₂濃度を計測する環境センサーが該当する．

図2.1は，スマートシティにおけるIoTネットワークの構造の一例である [2]．



図 2.1: Smart city based on fog architecture [2]

図2.1を参考にIoTネットワークを作成を行った．具体的には，RaspberrypiとセンサーデバイスはBLE通信で接続する．BLE通信とは，Bluetooth Low Energy通信の略である．この通信方法は，従来のBluetooth通信と比較して，より少ない電力で稼働することが可能である．消費電

力が少ないため、長期間のデータの収集に適しており、本研究で採用した。また、ラズベリーパイ、スマートフォン、タブレットとの親和性が高いため、ライブラリが豊富であり、開発環境が整っており、ソフトウェア面での実装におけるサポートが充実している。性能面では、通信速度が 1Mbps 程度と一般的な回線を使用した通信方法と比較して低速である。しかし、センサーデータのサイズを考慮しても、この程度の通信速度が出ていれば十分のため、本研究では通信速度は問題とならない。一方で、通信距離が主な課題となった。具体的には、通信距離が数メートルであるため、実験環境内で安定した通信を確立することが難しく切断が頻発してしまうことが課題となった。

2.2 異常検知

異常検知とは、データマイニングや機械学習の技術・アルゴリズムを用いて、蓄積された大量のデータとの比較や同一データ内での比較を行い、通常の動作に当てはまらない、異常と判断できるデータを検出するための技術である。実際の生活では、クレジットカードの不正使用検知 [11]、アカウントの不正アクセス検知 [12]、産業システムの故障検知 [13] の場面において用いられる技術である。このような場面で異常検知を実施することで、ネットワーク上におけるセキュリティリスクの低減、人件費の削減、ヒューマンエラー・事故の防止、業務の作業レベルの担保を実現している。用いられる手法は 3 種類に分けることができる。具体的には、外れ値検知、異常部位検出、変化点検知の 3 種類である。

無線でノード同士が接続されたセンサーネットワークにおいて、外れ値検知を用いて異常検知を実現した研究がある [3]。外れ値検知は対象とするデータ内において異常であると判断されるデータを検出する手法である。外れ値を検知することによってデータの品質を担保することができる。また、[3] の研究では、機械学習を用いることで検知精度を向上させている。一方で、ノードの移動性やネットワークの接続形式の変更に対応することが課題となっている。図 2.2 は、IoT ネットワークにおいて外れ値検知を実施する際の汎用的なアーキテクチャの一例である。動作は全部で三段階である。第一段階では、IoT センサーを用いて収集したデータを IoT ゲートウェイを介してサーバに転送する。第二段階では、収集したデータを入力データとして、外れ値検知を実施する。第三段階では、システムのエンドユーザーに向けて、アラートを含めた結果を送信する。異常部位検出は、外れ値検知の応用手法に当たる。具体的な違いとして、外れ値検知ではデータにおける異常と判断できる点を検知するのに対して、異常部位検出は異常と判断できる部分 (ex. 時系列データの一部) を検知する。変化点検知とは、時系列データにおいて急激に変化した箇所を検知する手法である。外れ値検知は異常と判断できる点を検知しているのに対し、変化点検知は急激に数値に変化が起きた点を全て検知する。センサーネットワークにおいて、変化点検知を用いて異常検知を行った研究がある [14]。また、機械学習を組み合わせることでデータのやり取りをする回数を減らした上で、異常検知精度の向上を実現した研究もある [15]。この研究で用いた機械学習アルゴリズムはベイジアンネットワークである。

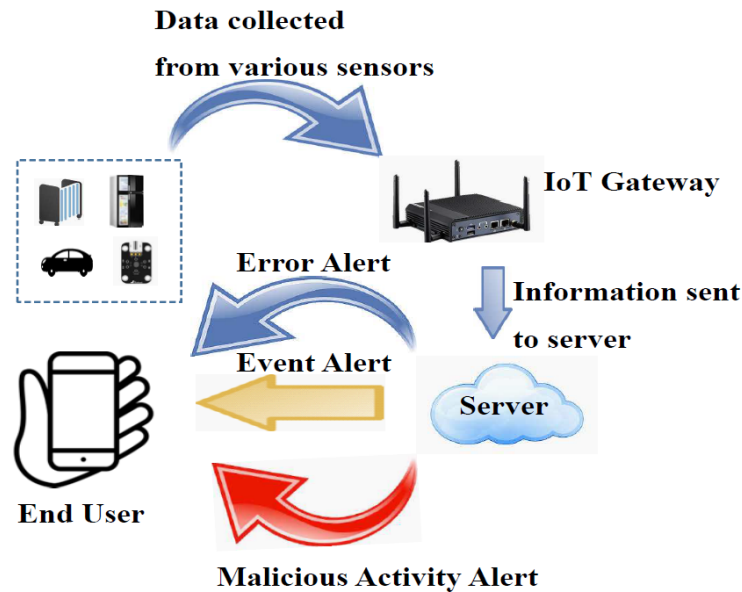


図 2.2: A generic architecture for different outlier detection in IoT and WSN [3]

外れ値検知と異常部位検出を行う上での代表的な手法として、 k 近傍法が挙げられる。この手法は、対象とする地点のデータと前後の地点のデータの距離、差分を計算し異常検知を実現する手法である。変化点検知を行う上での代表的な手法として、ホテリング理論が挙げられる。この手法は、統計モデルに基づいてデータの判別を行い、異常検知を実現する手法である。この手法では、対象とするデータは正規分布の定義に従うと仮定する。仮定した正規分布から、分散・平均を推定し、異常度・閾値を設定することで急激な変化がみられる地点を検知する。課題として、正規分布に近い形になりにくい時間依存性の高い時系列データに対して、精度が高くなりにくいという点がある。

2.3 センサーを用いた異常検知に関する研究

2.3.1 IoT データを用いたリアルタイム異常検知システム

IoT センサーデータに対して、異常検知システムの構築を行った研究がある [16]。IoT データのリアルタイム異常検知システムの構築にディープニューラルネットワーク (以下、DNN) は有効である。一方で、IoT デバイスには、複雑な DNN モデルを導入する余裕はないため、異常検知の負荷を軽減することが期待されている。この背景には、IoT デバイスの計算能力がある。また、クラウドにおける遅延の発生も課題となっている。この論文では、分散型階層エッジコンピューティングの異常検出アプローチのデモの作成を行った。対象とする IoT データは単変量または多変量である。そして、オートエンコーダーと LSTM を用いて異常検出 DNN モデルの作成を行う。モデル選択には、入力データのコンテキスト情報に基づいて行う、マルコフ決定過程を用いる。ローカル分析において、データのやり取り回数も重要な評価要素である。帯域幅の消費による遅延とやり取りの回数は相関関係が存在する。

IoT 環境における異常検知に関する研究として、[17] や [18] や [19] がある。また、エッジコンピューティングを用いた異常検知に関する研究として、[20] や [21] がある。また、モデル圧縮によって、IoT デバイスの計算能力不足の課題解決を図った研究として、[22] がある。

2.3.2 非線形グラフィルフィルターを用いた異常検出システム

非線形グラフィルフィルターを用いた異常検知システムに関する研究がある [23]。スマートセンシング技術は、スマートシティ、自動走行、インフラ自動監視のような生活の質を向上させる技術と関連している。IoT センサーネットワークの異常なデータの検出の課題として、数値上の逸脱がわずかな場合でも異常を検出することが困難なことである。誤差を平均化して異常検知を行うような手法の場合、データが増加する程、深刻化してしまう課題である。この論文では、センサーデータとネットワーク構造情報について研究し、検出率の高い異常検出アルゴリズムの開発を行った。具体的には、センサーデータを再構成する際には非線形多項式グラフィルフィルターを用いた。この際に、ネットワーク構造情報を活用して学習を行った。従来型のセンサーの異常検出では個々のセンサーの時間的な行動を分析していた。そこで、データの密度を推定し、密度推定を閾値処理することによって異常検知を実現している。具体的には、異なるセンサー間の関係を時間的・データの観点から結び付けることによって検出率の向上を目指している。測定における偏差データが大きいほど、テスト時におけるエラーの数が多くなるという結果が得られた。これは、高い検出率を実現できたということを示す結果である。

2.3.3 温度センサーシステムの異常検知に Prophet を使用した時間予測駆動型アプローチ

温度センサーシステムの異常検知システムに関する研究がある [24]。下水管の腐食状況を把握する上で、スマートセンサーシステムは有効である。ランダムな異常データ、異常な連続ストリームデータの把握が課題となっていた。時系列予測モデルは、異常検知アプローチを開発するために用いられる。関連モデルとして、自己回帰和分移動平均 (ARIMA) モデルが挙げられる。下水管の監視システムには高い信頼性が求められている。評価指標として MAE と RMSE を採用している。予測日数が増えると予測精度が低下する傾向である。Prophet は、短期間のセンサーデータの予測には有効である。具体的には、加法モデリングアプローチに基づく予測である。1 日だけ先に予測し、そのデータを用いて予測を行うと予測パフォーマンスが上昇した。今後の課題として、温度センサーだけでなく、表面温度センサーと表面水分センサーを含めたマルチセンサーからの異常ストリーミングを検出するアルゴリズムが求められる。関連技術として、地中レーダー、電気抵抗率ベースのセンサー測定、静電容量センサーが挙げられる。Prophet モデルと比較したモデルは以下の通りである。Auto-ARIMA モデル、Exponential Smoothing State Space (ETS) モデル、Bagged モデル、TBATS モデルの 4 つである。ARIMA モデルは自己回帰モデルと移動平均モデルを統合することで構成される。

ARIMA モデルに関する研究として、[25] や [26] がある。また、ARIMA モデルの構成についての研究として、[27] がある。Auto-ARIMA モデルに関する研究として、[28] や [29] がある。ETS モデルに関する研究として、[30] がある。Bagged モデルに関する研究として、[31] がある。TBATS モデルに関する研究として [32] がある。

2.3.4 エントロピー法を用いた異常検知システム

エントロピー法を用いて、センサーの異常検知を行った研究がある [33]。使用したデータ Intel Berkeley Research laboratory の分散センサーネットワークから生成されたものを用いた。比較対象は、エリプティカル法とエントロピー法である。従来の方法では、2 次元データセットの分析で

は、センサー間の相関を必要としていた。エントロピー法は各センサーを独立して扱うため、センサーが相互に創刊していないケースでは有効である。この研究の背景には、データ空間全体に散在する以上の検出の重要性がある。楕円法はデータ領域を区切ることで正常データと異常データを区別している。楕円法では、領域の線引きが大きな課題となっていた。エントロピー法では、領域の線引きではなく、一つ一つのデータを識別することによってデータ空間の領域を考えることなく、異常データを検出できるようになっている。システム内のセンサーノード・制御ノードが正常に機能していることを確保することが、従来型の手法の課題であった。そこで、早期異常検出システム (EADS) が開発された。そして、EADS の評価指標として、システム全閉への負荷が挙げられる。IoT ネットワークの場合、統計的アプローチの適用は計算能力の観点から適切ではない。エントロピー法は求められる計算能力が比較的低く、異常検知速度も十分である。

楕円法に関する研究として、[34] がある。

2.3.5 ウェアラブルセンサーネットワークにおけるストリーミングデータに対する異常検知システム

ストリーミングセンサーデータに対する異常検知システムの構築に関する研究がある [35]。ウェアラブルヘルスデバイスとは、センサーを用いて、装着者の生物学的・生理的パラメータを収集・監視するシステムに用いられるデバイスである。具体的には、心拍数、脈拍数、呼吸数、体温、血圧、血糖値、血中酸素濃度 (SpO2) が挙げられる。この研究の目的は、医療用ワイヤレスセンサーから収集されたデータの異常検知アルゴリズムの開発である。評価指標として、アラーム制度と再現率が挙げられる。疾患の早期発見や予測には、自由度と機動性が求められる。そこで、ワイヤレスセンサーネットワークを用いた、患者情報をリアルタイムに監視する方法が提案された。

障害センサーを発見するために外れ値検出を実施した研究として、[36] がある。また、IoT デバイスに対する攻撃である False data injection attack に関する研究がある [37]。

2.4 IoT ベースの環境測定システムに関する研究

小型かつ低コストな IoT ベース室内環境モニター装置の開発を行ったものがある [38] [4]. [38] の研究では、収集したデータを用いて、室内滞在時間と CO₂ や NO₂ といった汚染物質の濃度との関係について考察している. 用いたデータは相対湿度と温度の 2 つのパラメータである. この研究によって、個人が特定できるカメラやサーモグラフィを使用せずに滞在時間のようなセンサー単体で直接計測・推定することが難しい情報の周辺情報をセンサーを用いて集めた情報を元に予測を行うことが可能であることが示された.

また, [4] の研究では、屋内環境モニタリングシステムの開発を行った. エネルギーの消費削減と教育環境向上の問題を同時に扱った研究である. この研究で開発されたシステムは、インターネットを介して、遠隔操作でデータの収集・制御を行うことが可能である. また、室内環境のエネルギー消費削減に活用されるデータセットを作成し、提供も行っている. 図 2.3 は、センサーと Raspberry Pi とサーバーと端末を用いて構成される環境モニタリングシステムの構造である.

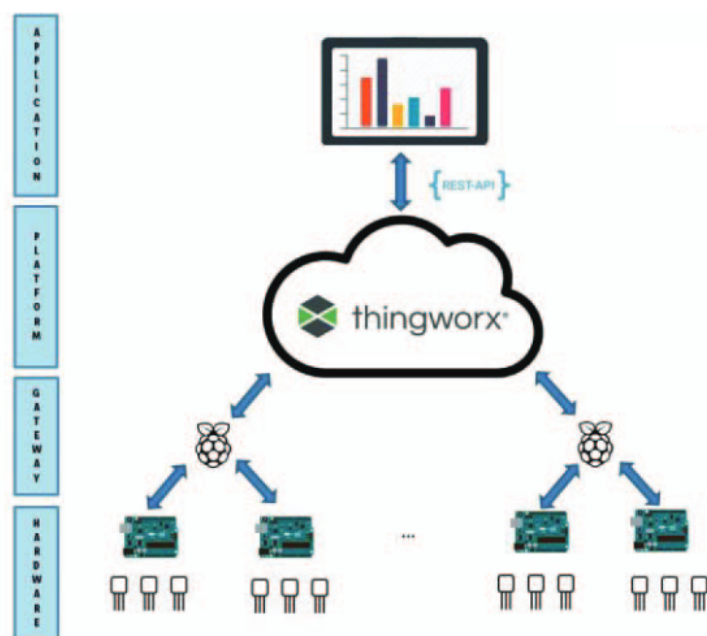


図 2.3: Environment Monitoring System Architecture [4]

異常検知技術を用いて、センサーデータを活用した居住環境の制御を実現した研究がある [39] [5]. [39] の研究では、居住環境におけるエネルギーの制御に関する研究である. CO₂ 濃度を制御しつつ、エネルギー効率の最適化を実現することが目的である. 具体的には、エネルギー効率を改善しつつ、CO₂ 濃度レベルを環境基準 (新 ASHRAE) 以下に下げることが目標としている. まず、CO₂ 濃度と温度や湿度との相関を求めた. 機械学習アルゴリズムの一種であり、非線形性を学習する能力に優れた多層パーセプトロンネットワークは選択したアルゴリズムの中で最も良い結果を得られた. 要求するコンピュータのスペックや計算コストを下げるために、制御装置の構築コスト、構築しやすさ、設計の単純さも考慮されている. [5] の研究では、環境センサーデータを用いて部屋の占有率予測を行なった研究である. 具体的には、照度・温度・湿度・CO₂ 濃度である. カメラを用いないで居住者の検出ができるシステムはプライバシーの観点からも有用である. センサーと raspberry pi は全て有線接続を行なっている. 正解データの作成にはカメラを用

いている。居住者推定の際に使用した機械学習アルゴリズムはロジスティック回帰、ガウシアンナイーブベイズ、k 近傍法、決定木、勾配ブースティングマシン、ランダムフォレスト、サポートベクターマシン、CART モデルである。図 2.4 は、CART モデルの実例である。予測精度が高く、解釈が簡単であるという利点がある。予測に使用する因子の数や選択するモデルによって予測精度が 32% から 97% まで変わる。また、時間情報を加えることによって、精度が上昇することも確認されている。確率的モデリングを用いることで、精度を向上させることが可能である。具体的には、k-means 法が挙げられる。

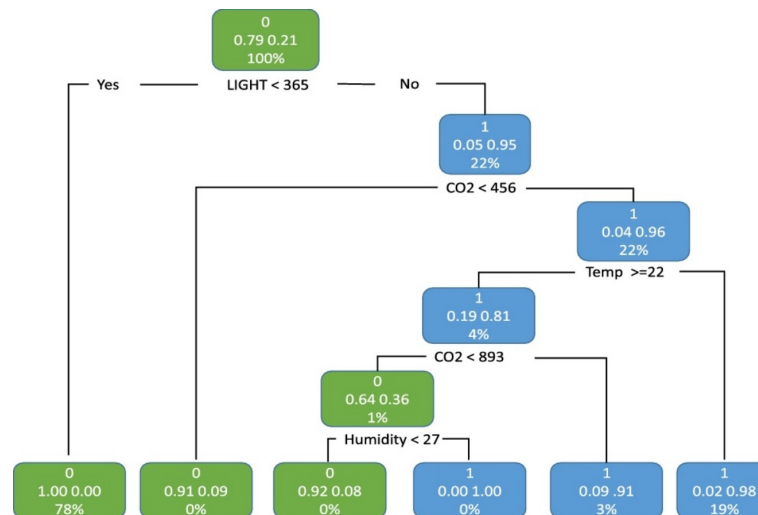


図 2.4: Environment Monitoring System Architecture [5]

2.5 機械学習

機械学習とは、一定のアルゴリズムに基づいて計算機上で学習を行い、対象とするデータ内に存在するパターンを発見し、抽出する手法である。本研究では、機械学習を活用して、目的とするデータの予測を行う。図 2.5 は、機械学習ライブラリである scikit-learn が提供しているモデル選択を行う際に用いる参考シートである。予測したい対象や取得できるデータセットの大きさによって、選択すべきモデルが変わる。データセットを作成する際には、量だけでなく質についても注意する必要がある。具体的には、センサーの観測精度に注意しなければならない。

以下で、実験で用いる機械学習アルゴリズムについて、説明する。

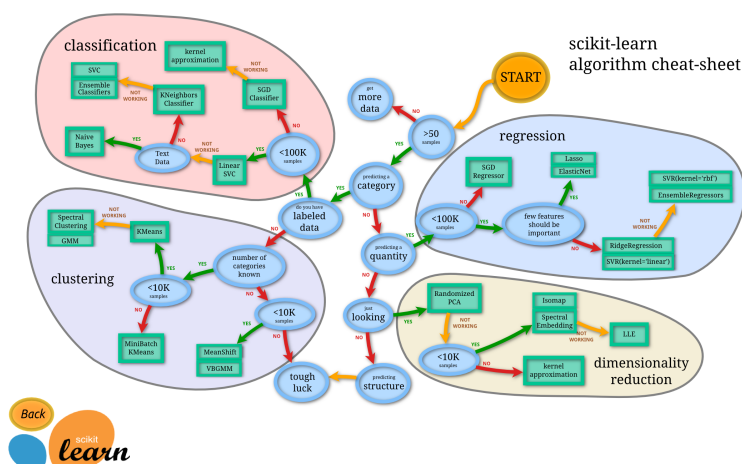


図 2.5: scikit-learn algorithm cheat-sheet [6]

ロジスティック回帰とは、教師あり機械学習アルゴリズムの一種である。具体的には、統計的回帰モデルの一種である。統計的回帰モデルとは、目的とする事象に関する予測や分類を行うことが可能なモデルである。実験では主に回帰を行うことを目的に使用する。回帰とは、数値の予測である。実験では、実験環境における人数の予測に用いる。また、部屋における人がいるかどうかという分類を行う際にも用いる。予測を行う際には、説明変数を用いる。実験では、CO₂ 濃度、照度、温度が該当する。また、ロジスティック回帰はベルヌーイ分布に従うと仮定されている。ベルヌーイ分布とは、李さん確率分布の一種である。

ガウシアンナイーブベイズとは、教師あり機械学習アルゴリズムの一種である。ナイーブベイズは、ベイズの定理を活用した分類手法である。ガウス分布に従うと仮定されている。そして、分類を行う際に求めるパラメータを尤度と呼ぶ。この尤度の分布がガウス分布に従うと仮定し、ナイーブベイズを適用しデータの処理を行うアルゴリズムがガウシアンナイーブベイズである。

k 近傍法とは、教師あり機械学習アルゴリズムの一種である。データの分類を行う際に用いるアルゴリズムである。具体的には、予測するデータと訓練データの距離を計算し、距離に応じて分類を行う。学習や予測に用いるデータはベクトルで表現される。課題として、大規模なデータの予測を行う場合に、計算量が膨大になるケースが存在する点が挙げられる。

決定木とは、教師あり機械学習アルゴリズムの一種である。実験では回帰木で用いている。具体的な動作としては、訓練データを用いて、分類器を作成し、予測対象のデータについてパラメータの数値で複数の分岐を行い、予測結果を決定する手法である。

ランダムフォレストとは、教師あり機械学習アルゴリズムの一種である。具体的には、決定木

を複数組み合わせることで、計算の並列処理を可能にしたアルゴリズムである。強みとして、計算が高速という点が挙げられる。

勾配ブースティングマシンとは、教師あり機械学習アルゴリズムの一種である。また、ランダムフォレストと同じ、決定木の組み合わせによって構成されたアルゴリズムである。違いとしては、勾配ブースティングマシンでは勾配降下法を用いて分類器の作成を行っている。勾配降下法とは、訓練データを用いて、予測を行っていく際に、パラメータを調整し、損失関数を最小にする手法である。本研究では、損失関数の 1 次導関数が 0 となるようにパラメータの調整を行う最急降下法を用いる。

サポートベクターマシンとは、教師あり機械学習アルゴリズムの一種である。仕組みとしては、マージン最大化とカーネルトリックを用いて分類を行っている。マージン最大化とは、分類する際に用いる境界線と境界線に最も近いデータとの距離を最大化する手法である。また、カーネルトリックとは、分類をするデータの構造が複雑な場合に、線形分類を可能にするためにデータを高次元空間に写像する手法である。データの次元数を大きくしても高い識別精度を維持できる反面、学習データが増えると計算量が膨大になり予測に時間がかかるという弱点も存在する。

第3章 IoTセンサーネットワークの構築とデータの可視化

本章では，研究全体の概要と本研究の第1段階にあたるIoTセンサーネットワークの構築と第2段階にあたるデータの可視化について説明する．

3.1 研究の段階分け

研究全体の概要について，説明する．本研究の目的である低い計算コストで高速かつ高精度に在室状況推定を実現する上で，課題となるのはセンサーデータのパラメータの最適化である．課題解決のために，複数のセンサーを用いて収集したデータに対して，前処理として実空間座標情報を考慮した前処理を施す．

本研究は6つの段階で構成される．第1段階は，センサーによるデータの収集を行うためのIoTセンサーネットワークの構築である．このIoTセンサーネットワークは在室状況把握の実現のための実験プラットフォームとしても用いる．第2段階は，環境センサーから収集したデータの可視化である．第3段階は，第1段階で作成したIoTネットワークを活用してのCandanedoの方法の検証・評価の実施である．第4段階は，実空間座標情報を考慮したデータの前処理を含めた，データセットの作成である．第5段階は，IoTセンサーネットワークと実空間座標情報を用いた在室状況推定の実験である．

3.2 IoTセンサーネットワークの構築

本研究の第1段階にあたるIoTセンサーネットワークの構築について，説明する．使用機材は，raspberrypi 4 model B，CO₂センサー，温湿度センサー8台，Webサーバー (www.sato.cnl.t.u-tokyo.ac.jp)，PC，スマートフォンである．ネットワークの構造は，6段階である．図3.1は各段階のデータのやり取りを図に示したものである．図中の番号はネットワークの構造の段階に対応している．CO₂センサーとRaspberry Pi 4は有線接続である．また，温度・湿度センサーとRaspberry Pi 4は，無線接続である．Raspberry Pi 4とWebサーバは有線接続である．

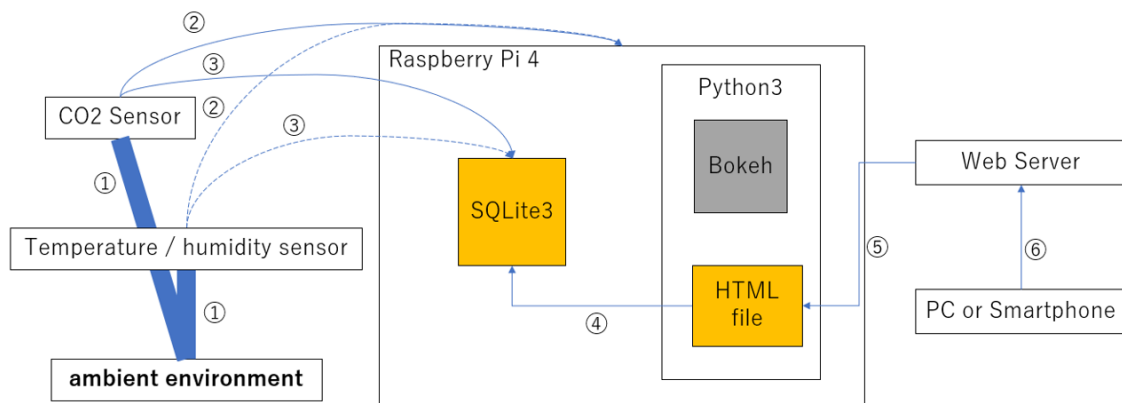


図 3.1: System configuration diagram

3.3 データの可視化

本研究の第 2 段階にあたる構築した IoT センサーネットワークを用いたセンサーデータの可視化について、説明する。図 3.1 における各段階の詳細について説明する。

図 3.1 の①は、環境センサーを用いて、環境データをセンサー内に取得する。具体的には、CO2 センサーを用いて、CO2 濃度を取得する。また、温度・湿度センサーを用いて、温度と湿度のデータを取得する。

図 3.1 の②は、Raspberry Pi 4 上で環境データを取得する。この時、CO2 センサーと Raspberry Pi 4 はジャンパー線で接続されており、シリアル通信によって CO2 濃度の数値を取得している。温度・湿度センサーとは、BLE 通信によって温度と湿度の数値を取得している。図 3.2 は、実際に CO2 センサーと Raspberry Pi 4 を接続した状態を示した図である。青い丸で囲まれた装置が CO2 センサーである。プログラム作成時に、使用した言語は Python である。温湿度センサーの場合は、適宜 BLE 通信を用いてセンサーと Raspberry Pi 4 を接続し、データの取得を行った。データを得るタイミングにその都度 1 対 1 で通信を行っている。温度・湿度センサーからデータを取得する際に、通信切断が頻繁に発生した。この問題に対応するために、通信試行回数を 1 回から 10 回に増やすことで対応した。通信試行回数を増やすことによってプログラムの異常終了が発生しなくなり、対策として機能している。また、cron タグを使用して、プログラムを定期実行することで、データの取得を定期的に実行している。

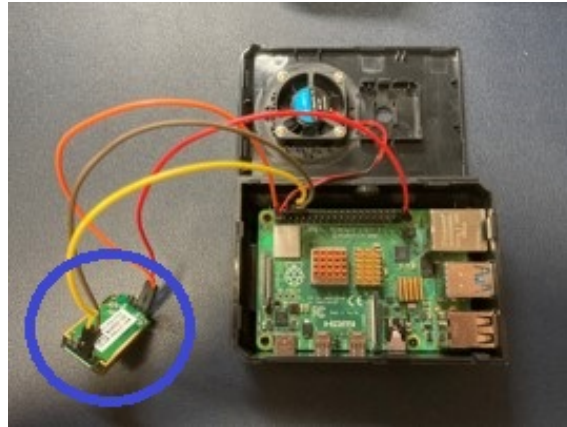


図 3.2: Connection environment

図 3.1 の③は、第 2 段階で取得した環境センサーデータの蓄積を行う。具体的には、SQLite3 を用いてデータベースを作成する。作成するデータベースの内部構造の 1 列目は、計測を行なった日付を示している。2 列目は、計測を行なった時間を示している。3 列目は計測された CO2 濃度情報 (ppm) を示している。本段階で作成するデータベースファイルは一日毎に新規で作成する。プログラム作成時に使用した言語は Python と SQL である。大規模なデータ収集を行う点とプログラムを用いた検索の利便性を考慮した結果、SQL を採用した。プログラム実行時、3 秒毎に CO2 濃度の計測を行なった結果を保存している。温湿度センサーの場合も同様の処理を行って、データの取得を行った。温度・湿度センサーの場合、データを取得するタイミングで接続し取得後に切断し、1 対 1 でセンサー毎に通信するため、環境データ (温度・湿度) の取得にかかる時間は設置したセンサー全てからデータを一通り取得するのにかかる時間は 90 秒程度となっている。

図 3.1 の④は、Raspberry Pi 4 上で、蓄積した CO2 データを元にグラフ・Web ページの生成を行う。SQLite3 を用いて作成されたデータベースファイルと Flask と Bokeh を用いて、HTML を作成する。この際に用いたプログラミング言語は Python と SQL である。この HTML ファイルの更新はデータベースファイルが更新される度に実行される。

図 3.1 の⑤は、作成した HTML ファイルを Web サーバー上に配置する。Raspberry Pi 4 と Web サーバを有線接続し、Raspberry Pi 4 内部のファイルを Web サーバに転送する。具体的には、cron タグを使用して、プログラムを定期実行することで、ファイルの配置を定期的に行っている。現在は、CO2 濃度の情報を 2 分毎に最新情報に更新している。

図 3.1 の⑥は、PC やスマホから Web サーバに保存された HTML ファイルにアクセスし、グラフを表示する。図 3.3 は、PC から作成した Web ページを表示した場合の画面の一例である。具体的には、CO2 センサーが取得した CO2 濃度 (単位: ppm) を縦軸に、観測した時間を横軸にし、時系列データをグラフ化したものである。グラフの Web ページ上への配置は一連の流れは、リアルタイムで処理されているため、その時の最新の状態が表示される。最新のタイミングから 20 分前までの情報を閲覧することができる。図 3.4 は、温湿度センサーが取得した温度 (単位: °C) や湿度 (単位: %) を縦軸に、観測した時間を横軸にし、時系列データをグラフ化したものである。図 3.5 は、観測した温度についてのヒートマップである。研究室のセンサーの配置に対応して、温度を色の濃淡を用いて表現している。具体的には、青が濃いほど温度が低く、赤が濃いほど温度が高くなる。

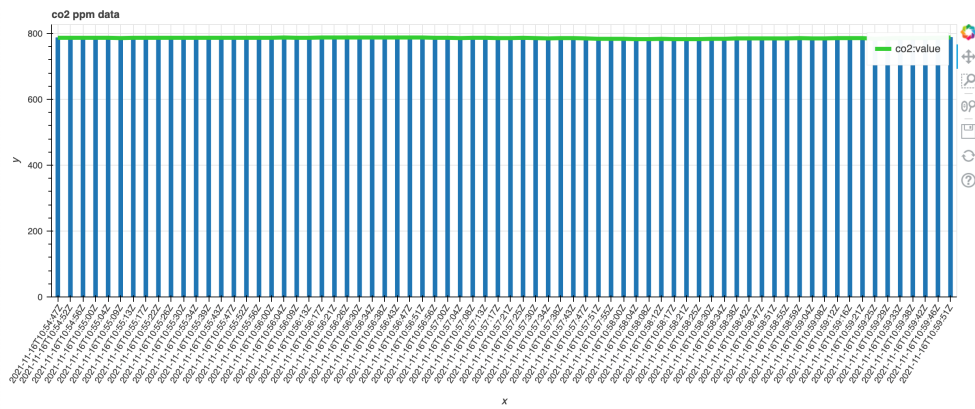


図 3.3: Graph on PC about CO2 sensor

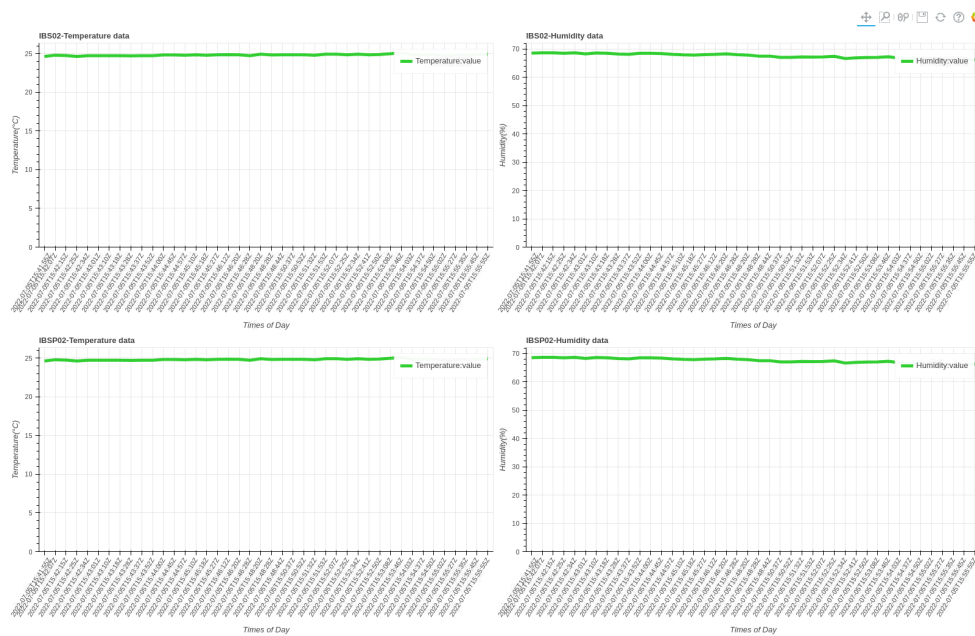


図 3.4: Graph on PC about Temperature / humidity sensor

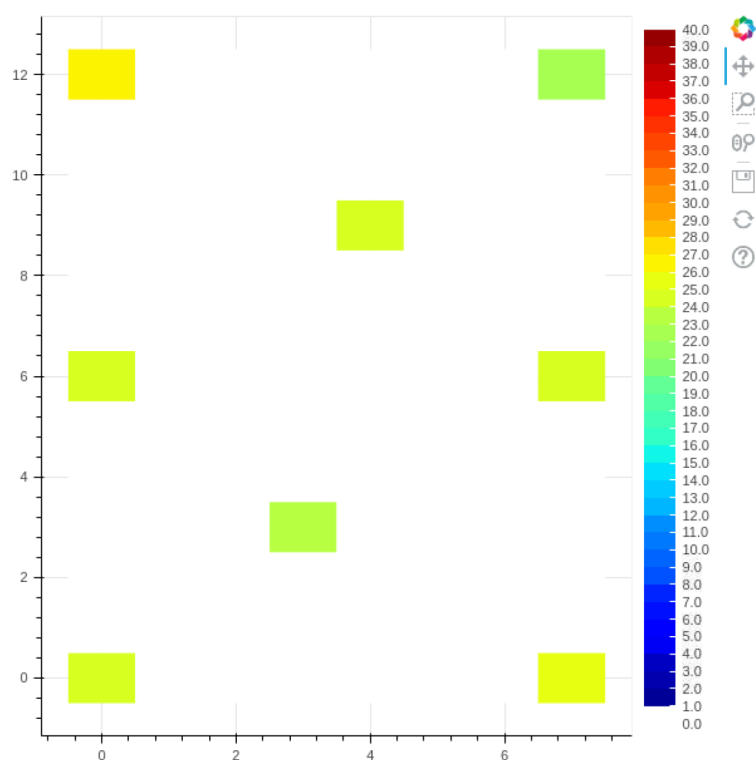


図 3.5: Heat map on PC about Temperature / humidity sensor

第4章 Candanedoの方法

本章では，本研究の第3段階にあたる Candanedo の方法の検証・評価について説明する．

4.1 予備実験の目的と概要

Candanedo らの研究 [5] では，本研究での実験環境と比較して狭い部屋において機械学習アルゴリズムを用いて，在室状況推定を行った．先行研究で実験を行っている部屋の広さは $5.85\text{m} \times 3.5\text{m}$ (W \times D) である．また，データセットの作成に必要な環境センサーデータの収集にかけた期間は17日間である．対象とする実験室における最大収容人数は2名である．行った予備実験は先行研究より広い部屋において同種の環境センサーデータ・機械学習アルゴリズムを用いて，在室状況推定を行うものである．予備実験を行った部屋の広さは $7.2\text{m} \times 12.8\text{m}$ (W \times D) である．

目的としては，部屋に人がいるかどうかについての推定精度は部屋の広さと関係があるかの確認と Candanedo の方法を用いた場合の推定精度の評価である．Candanedo の方法の具体的な手法は，実験方法にて説明する．

4.2 実験環境

予備実験の実験環境について，説明する．使用した実験機材は，Raspberry Pi1 台，温度・湿度センサーが1台，照度計1台，気圧計1台，CO2センサー1台である．図 4.1 は，実際に Raspberry Pi 4 と温度・湿度センサー，照度計，気圧計，CO2 センサーを接続した状態を示した図である．表 4.1 は，Raspberry Pi 4 の詳細な性能に関する情報をまとめた表である．表 4.2 は，Raspberry Pi 4 と有線接続する CO2 センサー，温度・湿度センサー，照度計，気圧計のスペック一覧である．

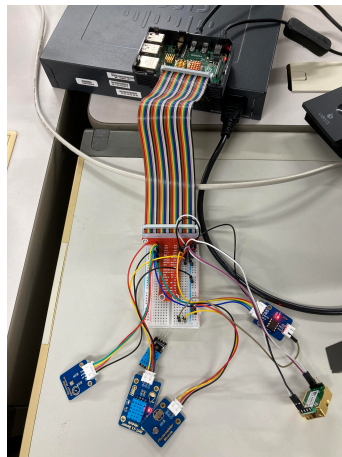


図 4.1: Connection environment for preliminary experiment

表 4.1: Computing environment

Model	Raspberry Pi 4 Model B
OS	Raspbian GNU/Linux 10 (buster)
CPU	Cortex-A72 4 コア/1.5[GHz]
Memory	4[GB]
GPU	512[MB]

表 4.2: Sensor equipment

Sensor	Prameter	Resolution	Range
MH-Z19C	CO2	1ppm	400~5000ppm
DHT22	Temperature	0.1 °C	-40~80 °C
DHT22	Humidity	1 %	0~100 %
CDS-55	Light(Photoresistor)	1mV	0~512mV
BMP180	Atmosphere pressure	0.01hPa	300~1100hPa

4.3 実験方法

予備実験の実験方法について、説明する。まず、収集した環境センサーデータを、訓練用とテスト用に分けることでデータセットを作成する。データを収集した期間は、2022 年 8 月 25 日から 2022 年 9 月 11 日までの 18 日間である。一つ目のテスト用データセットは人がいない日の情報のみで構成されている。表 4.3 は、分割後の訓練用データセットとテスト用データセットにおけるセンサーデータ数の一覧である。データセットを構成する要素は、日時、CO2 濃度、温度、湿度、気圧、照度、その時に部屋にいた人数の 7 種類である。その時に部屋にいた人数の計測は、実験環境内に定点カメラを設置し、5 分おきに撮影を行い、得られた写真からデータを入力した。データセット作成後に、データの相関係数を求める。相関係数が最も高いパラメーターのペアとその半分の数値のパラメーターのペアを選定する。選択したパラメーターのペアと機械学習アルゴリズムを用いて、部屋に人がいるかどうかの予測を行う。用いる機械学習アルゴリズムは、ロジスティック回帰、ガウシアンナイーブベイズ、k 近傍法、決定木、ランダムフォレスト、勾配ブースティングマシン、サポートベクターマシンの 7 種類である。予測を行う際に、対象とするデータセットと予測結果の比較を行うことで、評価を行う。用いる評価指標は正解率 Accuracy である。正解率とは、予測した数値が一致した割合である。予備実験では、部屋に人がいる状態とない状態の識別を行うため、人がいる状態を人がいると判断できたパターン数と人がいない状態をいないと判断できたパターン数の合計を全パターン数で割ることで正解率を求める。

表 4.3: Data set description

Data Set	Number of observations
Training	18418 of 7 variables
Test 1	8455 of 7 variables
Test 2	9752 of 7 variables

4.4 実験結果

予備実験の実験結果について、説明する．まず，センサーデータのプロットを行った．図 4.2 は，2022 年 9 月 5 日から 2022 年 9 月 11 日までの一週間分のセンサーデータの推移を示した図である．次に，収集したセンサーデータの各パラメーターについての相関係数を調べた．表 4.4 は，測定した各パラメーター同士の相関係数をまとめた表である．この表より，CO₂ 濃度と温度の相関係数が最も高いことが確認できる．また，その半分の数値の相関係数が確認できるパラメータの組み合わせとして，CO₂ 濃度と照度が挙げられる．図 4.3 は，測定したパラメーターの要素をグラフ上にプロットした散布図である．CO₂ 濃度と温度のペアと CO₂ 濃度と照度のペアの 2 つの組みと機械学習アルゴリズムを用いて，推定を行った．表 4.5 は，各モデルと各パラメーターを用いて，訓練データ，2 つのテストデータに対して推定を行った場合における正解率の数値の一覧である．

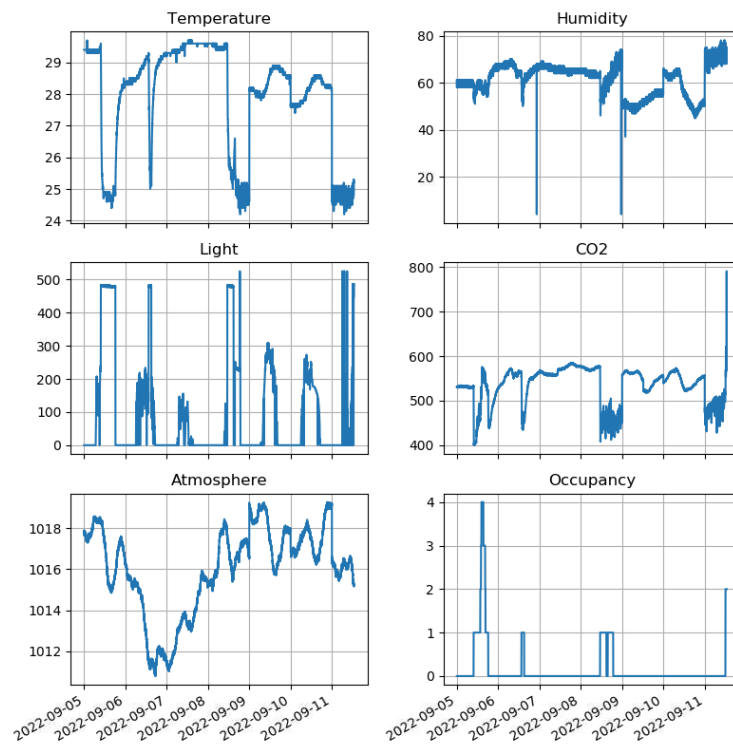


図 4.2: Measurement profiles for 1 week

表 4.4: Correlation matrix

	Temperature	Humidity	Light	CO ₂	Atmosphere	Occupancy
Temperature	1.000000	-0.158679	-0.436084	0.741529	-0.166572	-0.484403
Humidity	-0.158679	1.000000	-0.238436	-0.037237	-0.470357	-0.117586
Light	-0.436084	-0.238436	1.000000	-0.371051	0.004106	0.721287
CO ₂	0.741529	-0.037237	-0.371051	1.000000	-0.129283	-0.243428
Atmosphere	-0.166572	-0.470357	0.004106	-0.129283	1.000000	-0.048984
Occupancy	-0.484403	-0.117586	0.721287	-0.243428	-0.048984	1.000000

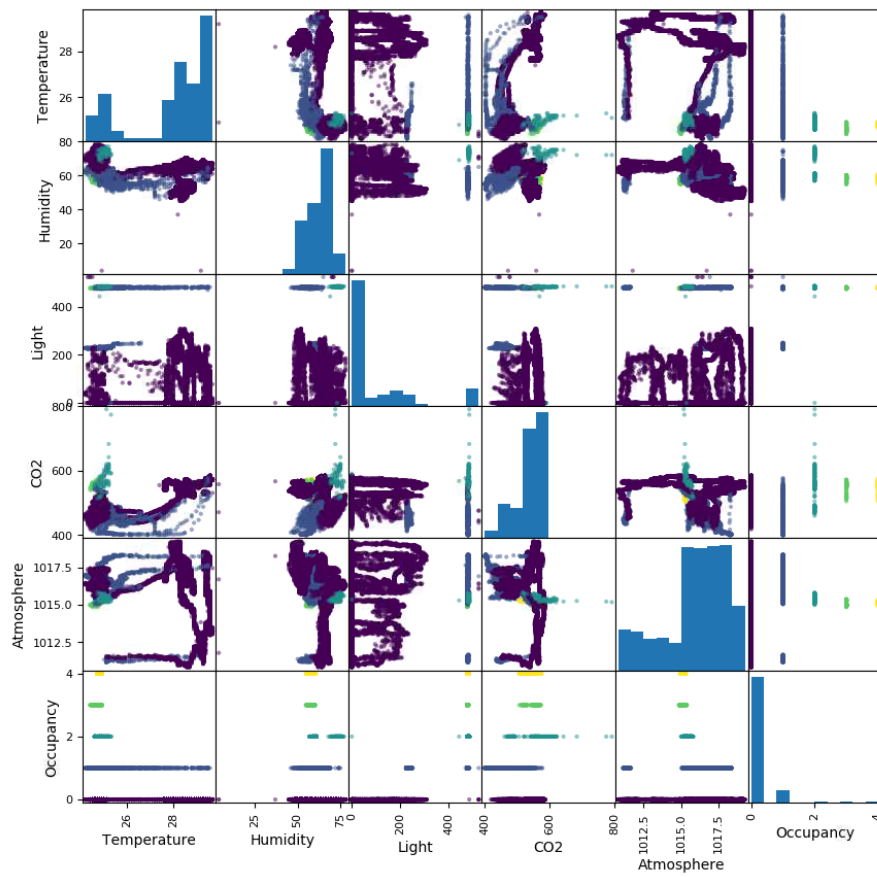


図 4.3: Pairs plot

表 4.5: Accuracy at each parameter for each model application

Model	Parameters	Training[%]	Test1[%]	Test2[%]
LR	CO2,Light	97.28	99.91	94.18
LR	CO2,T	90.74	100	91.64
GNB	CO2,Light	96.59	99.99	94.30
GNB	CO2,T	84.98	100	93.28
KN	CO2,Light	98.77	99.82	95.27
KN	CO2,T	94.09	100	86.17
DT	CO2,Light	98.97	99.13	95.83
DT	CO2,T	95.12	95.67	54.37
RF	CO2,Light	99.08	99.79	95.63
RF	CO2,T	96.48	87.09	47.77
GBC	CO2,Light	98.73	99.13	95.84
GBC	CO2,T	95.33	100	64.68
SVC	CO2,Light	99.08	99.98	95.18
SVC	CO2,T	96.27	87.89	71.97

4.5 評価

Candanedo の方法についての実験の評価を行う．一つ目のテスト用のデータセットの正解率から，人がいないことの識別の精度は部屋の広さに関係なく 95 %以上の精度を実現できることが確認できた．一つ目のテスト用のデータセットは人が一人も入室・退室していない日のデータで構成されているため，このように考察した．この結果から，人がいるかどうかの推定は上手くいっていると考えた．二つ目のテストデータの正解率から，決定木，ランダムフォレスト，勾配ブースティングマシン，サポートベクターマシンの 4 種類の機械学習アルゴリズムを適用した場合に正解率が低下してしまっていることが確認できた．また，実験の結果から，CO₂ 濃度と照度を組み合わせで用いた場合と比較して，CO₂ 濃度と温度データを組み合わせで用いた場合の精度は低くなっており，改善の余地がある．この結果について，部屋の広さが広がったにも関わらず，センサーの設置箇所を一つとしていたことも原因の一つであると考察した．

照度センサーを用いた場合の推定精度は温度センサーを用いた場合よりも高いという結果であったが，室内の照明をつけることなく人が実験環境に入った場合に推定することができないという可能性があると考えた，そこで，温度センサーであれば外的要因に左右されにくいパラメータであるため，6 章の実験においてメインで使用するパラメータとする．

在室状況の把握としては，人がいるかどうかだけでは不十分である．具体的には，在室している人数の把握をすることが必要である．これは，予測結果を空調制御に応用する場合に活用できる．次の実験では，在室している人数の推定を行う．それに合わせて，評価指標を変える必要があると考察した．この背景には，予備実験の結果は 2 値分類によって得た結果のため正解率という指標が評価する場合に正しく機能していたことが挙げられる．在室している人数を評価する場合に，この指標をそのまま使用してしまうと，一人でも人数がずれていた場合に不正解として処理してしまい，推定精度の評価として不適切である．

第5章 提案手法

本章では，提案手法について、説明する．提案手法の目的は，機械学習を行う際の計算量を削減し，計算コストを低減し，高速に推定を行うために，センサーデータの次元数を削減することである．

5.1 環境センサーデータを用いたデータセットの作成

まず，実験環境となる部屋の中に温度センサー，CO2 濃度センサーを設置する．この時に，温度センサーの数は任意である．設置した温度センサーの座標を部屋の広さを用いて，設定する．具体的には，図 5.1 のように部屋の横方向を x 軸，部屋の縦方向を y 軸として，センサーの座標を設定する．座標を縦方向または横方向に 1 移動すると，実際の空間では 1cm 移動するように座標空間を設定する．そして，環境センサー類と RaspberryPi を接続し，環境情報の収集を行う．以下では，複数のセンサーから収集・蓄積された温度情報を用いた，実験環境における温度の設定を行う時の説明をする．

一つ目の複数の温度情報を組み合わせて作成する実験環境における温度の設定は伝熱工学におけるフーリエの法則を参考に行う．フーリエの法則は以下の式で表される [40]．(T: 温度 [K], k : 熱伝導率 [W/(m·K)], q : 熱流束 [W/m²], x : 距離 [m])

$$q = -k \frac{\partial T}{\partial x}$$

フーリエの法則では，ある地点が受け取る熱量は距離に反比例する．この性質を参考に，以下の手順で実験環境における温度を複数のセンサーから収集したデータを用いて設定した．図 5.1 は，2 台の温度センサーの地点 (点 A, 点 B) と焦点として設定した地点 (点 F) の一例である．

以下の説明では，図 5.1 に沿って温度センサーが 2 台の場合について説明する．点 A の地点で計測した温度を T_a ，点 B の地点で計測した温度を T_b ，焦点 (点 F) の温度を T とする．部屋の温度として，求めたい温度は焦点の温度 T である．また，点 A の座標を (v_a, h_a) ，点 B の座標を (v_b, h_b) ，点 F の座標を (v_f, h_f) とする．

1. まず，焦点とセンサーの距離を求める．図 5.1 の場合は，点 A と点 F の距離 D_a ，点 B と点 F の距離 D_b を求める．

$$D_a = \sqrt{(v_f - v_a)^2 + (h_f - h_a)^2}$$

$$D_b = \sqrt{(v_f - v_b)^2 + (h_f - h_b)^2}$$

2. 次に，1. で求めた D_a と D_b が， $D_a < D_b$ の場合，以下のようにして S_a と S_b を求める．

$$S_a = \frac{D_b}{D_a}$$

$$S_b = \frac{D_a}{D_b}$$

また, 1. で求めた D_a と D_b が, $D_b < D_a$ の場合, 以下のようにして S_a と S_b を求める.

$$S_a = \frac{D_b}{D_a}$$

$$S_b = \frac{D_b}{D_b}$$

3. 2. で求めた S_a と S_b と点 A で計測した温度 T_a と点 B で計測した温度 T_b を用いて, 点 F の温度 T を求める.

$$T = t_a \times \frac{S_a}{(S_a + S_b)} + t_b \times \frac{S_b}{(S_a + S_b)}$$

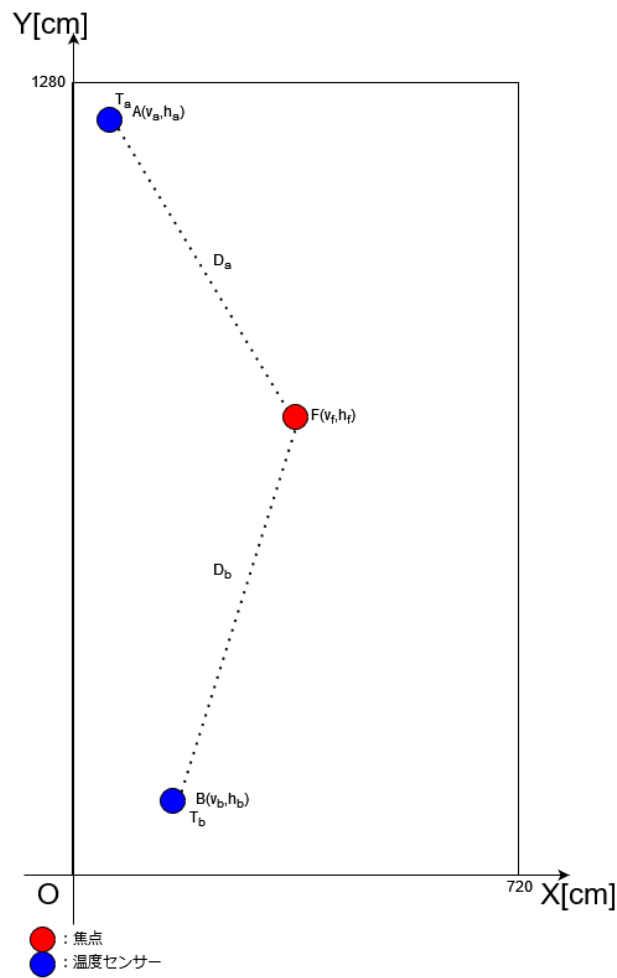


図 5.1: An example of focus and sensor positions in the experimental environment

S_a と S_b の大きさの比率によって、点 A で計測した温度 T_a と点 B で計測した温度 T_b に重みを付けて点 F の温度 T を求めている。伝わる熱量は距離に反比例するというフーリエの法則の性質を参考に、目的の温度を求める際のそれぞれの計測した温度への重みは距離に反比例する。

二つ目の複数の温度情報を組み合わせて作成する実験環境における温度の設定は伝熱工学におけるラプラスの式と差分法を参考に行う。ラプラスの式は以下の式で表される。(T: 温度 [K], x: 水平距離 [m], y: 垂直距離 [m])

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0$$

ラプラスの式に差分法を適用した場合を参考にする。焦点として設定した点の温度 T は周囲の 4 点で計測された温度の平均値として求めある。焦点の温度を T 、周囲の 4 点で計測された温度を T_a, T_b, T_c, T_d とすると、以下のようして T を求めることができる。

$$T = \frac{T_a + T_b + T_c + T_d}{4}$$

以上の手順で求めた温度情報と CO2 濃度の情報を組み合わせて、環境情報データセットを作成する。

5.2 在室推定で用いる機械学習アルゴリズムと評価指標

作成した環境情報データセットから訓練用データセットとテスト用データセットを作成する。作成したデータセットを用いて、選択した機械学習アルゴリズムを適用し、在室人数を推定する。具体的には、用いる機械学習アルゴリズムは、ロジスティック回帰、ガウシアンナイーブベイズ、k 近傍法、決定木、ランダムフォレスト、勾配ブースティングマシン、サポートベクターマシンの 7 種類である。予測を行う際に、対象とするデータセットと予測結果の比較を行うことで、評価を行う。

使用する評価指標について説明する。機械学習アルゴリズム適用後に、正しく予測を行えているかの評価指標として、RMSE と MAE を用いる。RMSE とは、二重平均平方根誤差とも呼ばれる。具体的には、予測した値と正解データの値の二乗誤差の平均値を求め、その値の平方根を計算すると求めることができる。平方根を計算しない場合の数値は MSE と呼ばれる。RMSE は予測値と実際の値の誤差の大きさが小さい程、小さな値になるため、値が小さい程良い精度の予測ができていると評価できる。RMSE は、下記のような式で計算される。(n は比較する要素の数、 f_i は予測した数値、 y_i は正解データの数値とする。)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - y_i)^2}$$

MAE とは、平均絶対誤差とも呼ばれる。具体的には、予測した値と正解データの値の誤差の絶対値から算出することで求める。MAE は予測値と実際の値の誤差の大きさが小さい程、小さな値になるため、値が小さい程良い精度の予測ができていると評価できる。MAE は、下記のような式で計算される。(n は比較する要素の数、 f_i は予測した数値、 y_i は正解データの数値とする。)

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i|$$

MAE と RMSE の違いとしては以下のような点が挙げられる。MAE は誤差を平均的に評価しているため、局所的な誤差、外れ値の影響を受けにくい。RMSE は局所的な誤差も考慮するため外れ値の影響を受けやすい。MAE・RMSE 両方の指標を用いての評価を行う。

第6章 IoTセンサーネットワークと実空間座標情報を用いた在室人数推定

本章では、本研究の第4段階にあたる実空間座標情報を考慮したデータの前処理を含めた、データセットの作成と第5段階にあたる、IoTセンサーネットワークと実空間座標情報を用いた在室状況推定の実験について、説明する。

6.1 実験環境

実験環境について、説明する。使用した実験機材は、Raspberry Pi1 台、温度・湿度センサーが10 台、照度計1 台、気圧計1 台、CO2 センサー1 台である。表 6.1 は Raspberry Pi 4 と接続するCO2 センサー、温度・湿度センサー、照度計、気圧計のスペック一覧である。図 6.1 は、実験環境の中におけるセンサーの配置を表した図である。OmB02, OmB03 は表 6.1 の 2JCIE-BL01 に対応している。IBS01, IBS02, IBS03 は表 6.1 の IBS-TH2 に対応している。IBSP02, IBSP03 は表 6.1 の IBS-TH2PLUS に対応している。SwTH02, SwTH03 は表 6.1 の SWITCHBOTMETER-GH に対応している。図 6.2 は実験環境内における各センサーの座標を示した図である。実験環境の部屋の広さは 7.2m × 12.8m(W × D) である。この部屋の広さを cm に変換し、部屋の座標として用いている。

表 6.1: IoT sensor equipment

Sensor	Prameter	Resolution	Range
MH-Z19C	CO2	1ppm	400～5000ppm
DHT22	Temperature	0.1 °C	-40～80 °C
DHT22	Humidity	1 %	0～100 %
CDS-55	Light(Photoresistor)	1mV	0～512mV
BMP180	Atmosphere pressure	0.01hPa	300～1100hPa
2JCIE-BL01	Temperature	0.1 °C	-10～60 °C
2JCIE-BL01	Humidity	0.1 %	30～85 %
IBS-TH2	Temperature	0.01 °C	-40～60 °C
IBS-TH2	Humidity	0.01 %	0～99 %
IBS-TH2 PLUS	Temperature	0.01 °C	-40～60 °C
IBS-TH2 PLUS	Humidity	0.01 %	0～99 %
SWITCHBOTMETER-GH	Temperature	0.1 °C	-20～80 °C
SWITCHBOTMETER-GH	Humidity	1 %	0～99 %

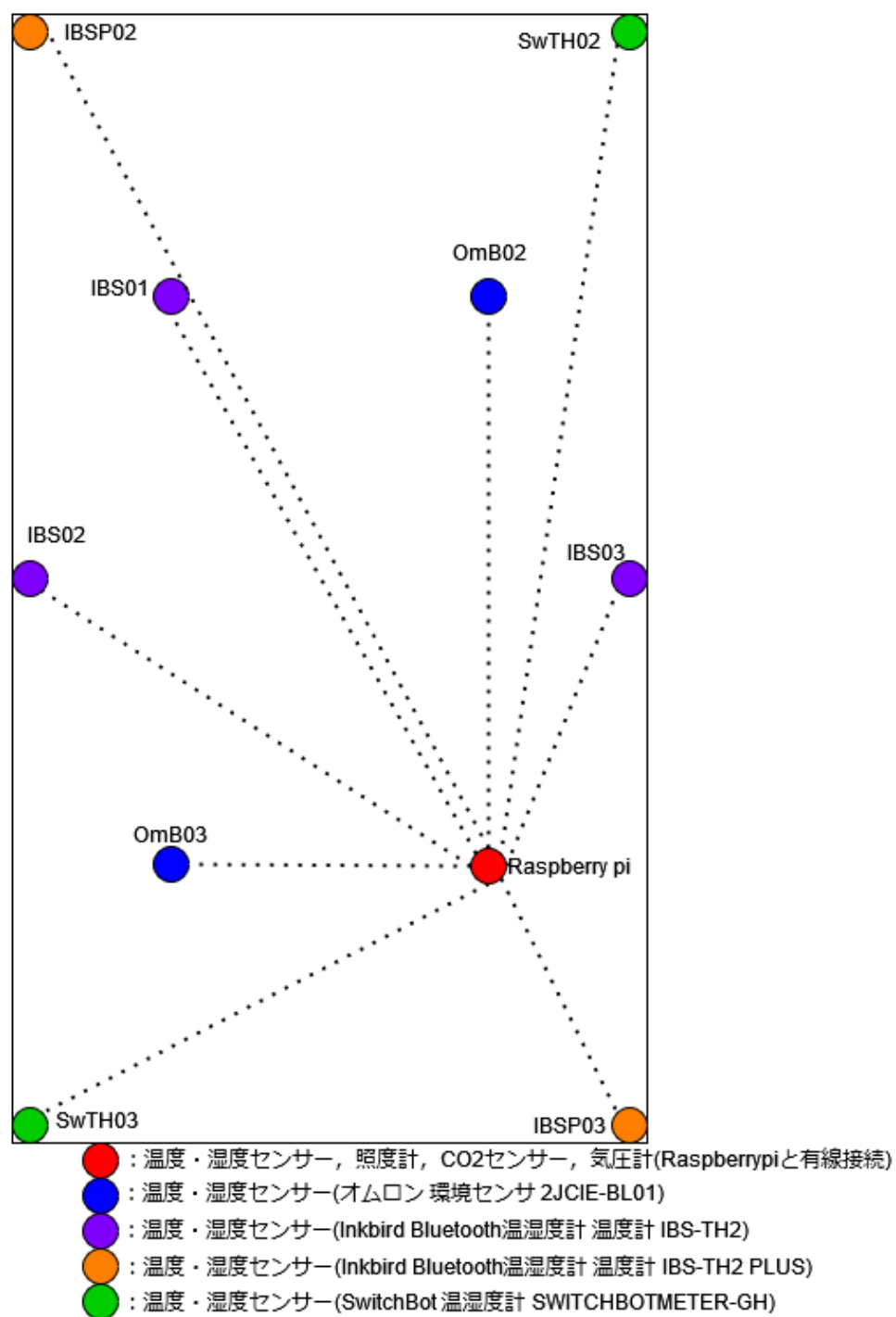


図 6.1: Sensor placement

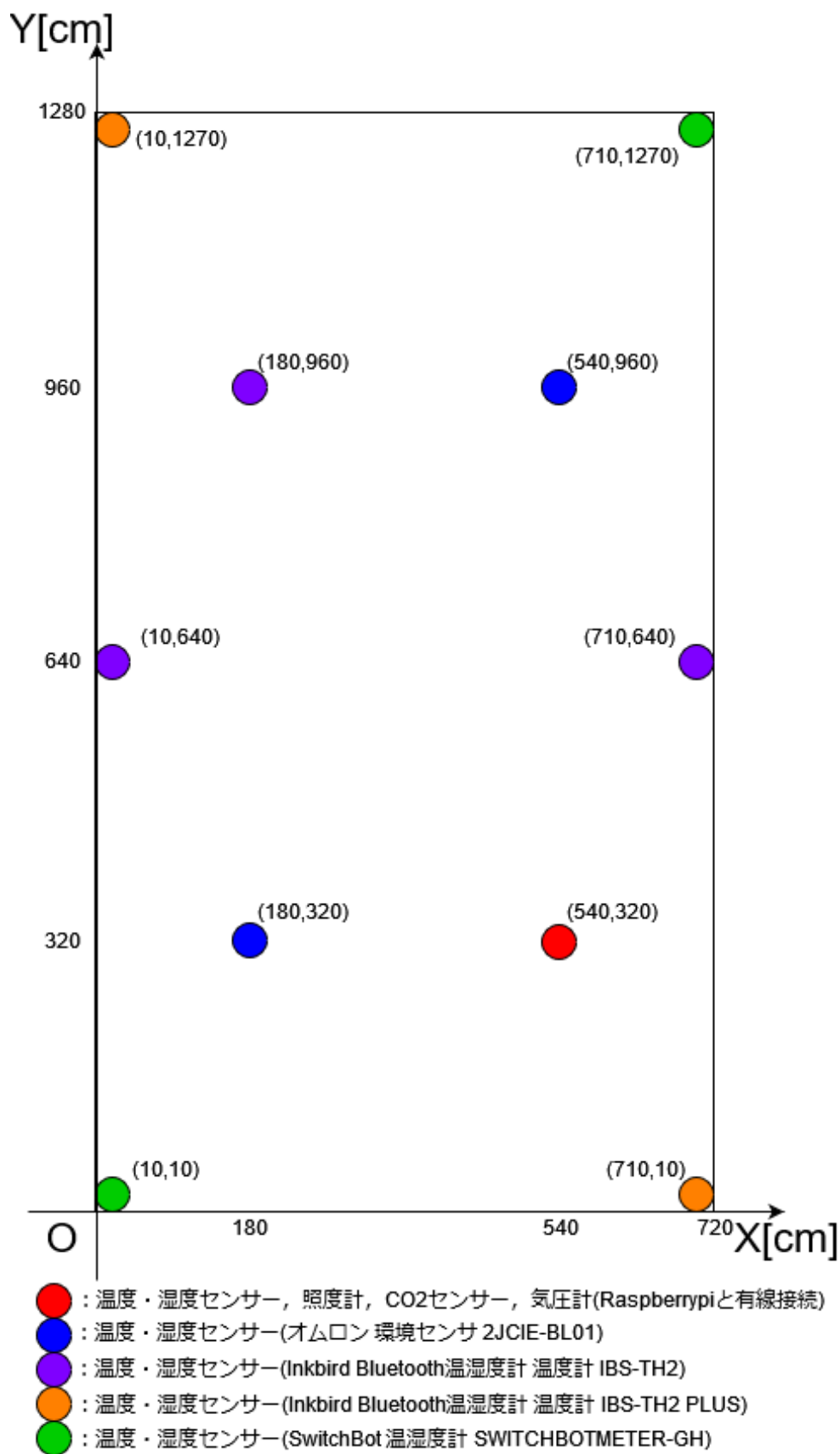


図 6.2: Sensor location

6.2 実験方法

本実験における、実験方法について、説明する．まず、収集した環境センサーデータを、訓練用とテスト用に分けることでデータセットを作成する．データを収集した期間は、2022 年 11 月 18 日から 2022 年 12 月 12 日の 25 日間である．表 6.2 は、分割後の訓練用データセットとテスト用データセットにおけるセンサーデータ数の一覧である．その時に部屋にいた人数の計測は、実験環境内に定点カメラを設置し、5 分おきに撮影を行い、得られた写真からデータを入力した．収集したデータに提案手法を適用し、データセットを作成する．作成したデータセットと機械学習アルゴリズムを用いて、部屋の人数を推定する．用いる機械学習アルゴリズムは、ロジスティック回帰、ガウシアンナイーブベイズ、k 近傍法、決定木、ランダムフォレスト、勾配ブースティングマシン、サポートベクターマシンの 7 種類である．予測を行う際に、対象とするデータセットと予測結果の比較を行うことで、評価を行う．使用する評価指標は RMSE と MAE の 2 種類である．参考までに MSE の値も求める．

表 6.2: Data set description

Data Set	Number of observations
Training	10212 of 25 variables
Test 1	4583 of 25 variables
Test 2	5061 of 25 variables

6.3 実験結果

本実験の実験結果について、説明する．表で用いている $t1_1$, $t1_2$, $t1_3$, $t2$, $t3$ は提案手法を用いて求めた温度である． $t1_1$, $t1_2$, $t1_3$ は一つ目の提案手法を用いて求めている．具体的には、 $t1_1$ の場合は焦点の座標を (360,640), $t1_2$ の場合は焦点の座標を (360,960), $t1_3$ の場合は焦点の座標を (360,1280) としている． $t2$ と $t3$ は二つ目の提案手法を用いて求めている． $t2$ の場合は、(180,960), (540,960), (180,320), (540,320) に設置した温度センサーの値を用いて求める． $t3$ の場合は、(10,1270), (710,1270), (10,10), (710,10) に設置した温度センサーの値を用いて求める．表 6.3 は、各モデルと各パラメーターを用いて、訓練データと 2 つのテストデータに対して推定を行った場合における MAE の数値の一覧である．表 6.4 は、各モデルと各パラメーターを用いて、訓練データと 2 つのテストデータに対して推定を行った場合における MSE の数値の一覧を参考までに示す．表 6.5 は、各モデルと各パラメーターを用いて、訓練データと 2 つのテストデータに対して推定を行った場合における RMSE の数値の一覧である．

表 6.3: MAE at each parameter for each model application

Model	Parameter	Training	Test1	Test2
LR	CO2,T	0.519	0.655	0.402
LR	CO2,t1_1	0.529	0.661	0.406
LR	CO2,t1_2	0.523	0.669	0.406
LR	CO2,t1_3	0.515	0.672	0.404
LR	CO2,t2	0.509	0.561	0.372
LR	CO2,t3	0.514	0.667	0.412
GNB	CO2,T	0.452	0.647	0.445
GNB	CO2,t1_1	0.472	0.696	0.441
GNB	CO2,t1_2	0.479	0.697	0.437
GNB	CO2,t1_3	0.482	0.690	0.442
GNB	CO2,t2	0.483	0.667	0.448
GNB	CO2,t3	0.489	0.572	0.433
KN	CO2,T	0.403	0.660	0.704
KN	CO2,t1_1	0.408	0.679	0.644
KN	CO2,t1_2	0.406	0.666	0.676
KN	CO2,t1_3	0.399	0.671	0.712
KN	CO2,t2	0.415	0.664	0.659
KN	CO2,t3	0.408	0.675	0.657
DT	CO2,T	0.499	0.929	0.396
DT	CO2,t1_1	0.477	0.882	0.368
DT	CO2,t1_2	0.481	0.887	0.367
DT	CO2,t1_3	0.504	1.163	0.418
DT	CO2,t2	0.506	0.656	0.379
DT	CO2,t3	0.501	0.924	0.369
RF	CO2,T	0.182	1.077	0.821
RF	CO2,t1_1	0.180	1.275	0.816
RF	CO2,t1_2	0.185	1.164	0.830
RF	CO2,t1_3	0.027	1.193	0.884
RF	CO2,t2	0.212	1.157	0.772
RF	CO2,t3	0.021	1.269	0.966
GBC	CO2,T	0.245	1.095	0.874
GBC	CO2,t1_1	0.265	1.142	0.812
GBC	CO2,t1_2	0.269	1.209	0.834
GBC	CO2,t1_3	0.253	1.289	0.906
GBC	CO2,t2	0.284	1.033	0.809
GBC	CO2,t3	0.195	1.249	1.009
SVC	CO2,T	0.179	1.235	0.965
SVC	CO2,t1_1	0.187	1.306	0.923
SVC	CO2,t1_2	0.191	1.309	0.922
SVC	CO2,t1_3	0.181	1.317	0.960
SVC	CO2,t2	0.214	1.285	0.899
SVC	CO2,t3	0.141	1.124	0.865

表 6.4: MSE at each parameter for each model application

Model	Parameter	Training	Test1	Test2
LR	CO2,T	1.059	1.202	0.501
LR	CO2,t1_1	1.084	1.221	0.507
LR	CO2,t1_2	1.069	1.227	0.506
LR	CO2,t1_3	1.044	1.227	0.503
LR	CO2,t2	1.008	0.918	0.474
LR	CO2,t3	1.044	1.221	0.512
GNB	CO2,T	0.815	1.376	0.587
GNB	CO2,t1_1	0.852	1.707	0.589
GNB	CO2,t1_2	0.887	1.690	0.603
GNB	CO2,t1_3	0.896	1.630	0.629
GNB	CO2,t2	0.912	1.500	0.638
GNB	CO2,t3	0.934	1.063	0.561
KN	CO2,T	0.807	1.245	1.308
KN	CO2,t1_1	0.833	1.310	1.124
KN	CO2,t1_2	0.826	1.275	1.218
KN	CO2,t1_3	0.808	1.284	1.301
KN	CO2,t2	0.839	1.268	1.161
KN	CO2,t3	0.847	1.289	1.186
DT	CO2,T	0.998	1.474	0.484
DT	CO2,t1_1	0.975	1.426	0.436
DT	CO2,t1_2	0.979	1.431	0.432
DT	CO2,t1_3	1.198	2.605	0.604
DT	CO2,t2	1.044	1.284	0.497
DT	CO2,t3	0.998	1.477	0.443
RF	CO2,T	0.282	2.194	1.667
RF	CO2,t1_1	0.306	2.710	1.684
RF	CO2,t1_2	0.320	2.448	1.682
RF	CO2,t1_3	0.046	2.541	1.858
RF	CO2,t2	0.375	2.396	1.548
RF	CO2,t3	0.042	2.638	2.208
GBC	CO2,T	0.374	2.324	1.837
GBC	CO2,t1_1	0.432	2.731	1.704
GBC	CO2,t1_2	0.444	2.742	1.778
GBC	CO2,t1_3	0.412	2.929	2.025
GBC	CO2,t2	0.475	2.278	1.689
GBC	CO2,t3	0.306	2.840	2.425
SVC	CO2,T	0.284	2.389	1.817
SVC	CO2,t1_1	0.309	2.763	1.709
SVC	CO2,t1_2	0.312	2.764	1.689
SVC	CO2,t1_3	0.306	2.717	1.814
SVC	CO2,t2	0.363	2.628	1.619
SVC	CO2,t3	0.228	2.874	1.793

表 6.5: RMSE at each parameter for each model application

Model	Parameter	Training	Test1	Test2
LR	CO2,T	1.029	1.096	0.708
LR	CO2,t1_1	1.041	1.105	0.712
LR	CO2,t1_2	1.034	1.108	0.711
LR	CO2,t1_3	1.021	1.108	0.709
LR	CO2,t2	1.004	0.958	0.689
LR	CO2,t3	1.022	1.105	0.716
GNB	CO2,T	0.903	1.173	0.766
GNB	CO2,t1_1	0.923	1.307	0.767
GNB	CO2,t1_2	0.942	1.300	0.776
GNB	CO2,t1_3	0.947	1.276	0.793
GNB	CO2,t2	0.955	1.225	0.799
GNB	CO2,t3	0.966	1.031	0.749
KN	CO2,T	0.898	1.116	1.144
KN	CO2,t1_1	0.913	1.145	1.060
KN	CO2,t1_2	0.909	1.129	1.104
KN	CO2,t1_3	0.899	1.133	1.144
KN	CO2,t2	0.916	1.126	1.078
KN	CO2,t3	0.920	1.136	1.089
DT	CO2,T	0.999	1.214	0.696
DT	CO2,t1_1	0.987	1.194	0.660
DT	CO2,t1_2	0.989	1.196	0.657
DT	CO2,t1_3	1.094	1.614	0.777
DT	CO2,t2	1.022	1.133	0.705
DT	CO2,t3	0.999	1.211	0.666
RF	CO2,T	0.530	2.194	1.291
RF	CO2,t1_1	0.553	2.710	1.298
RF	CO2,t1_2	0.566	2.448	1.297
RF	CO2,t1_3	0.216	2.541	1.363
RF	CO2,t2	0.612	2.396	1.244
RF	CO2,t3	0.206	2.638	1.486
GBC	CO2,T	0.611	1.525	1.356
GBC	CO2,t1_1	0.657	1.652	1.305
GBC	CO2,t1_2	0.666	1.656	1.333
GBC	CO2,t1_3	0.642	1.709	1.423
GBC	CO2,t2	0.689	1.509	1.299
GBC	CO2,t3	0.553	1.685	1.557
SVC	CO2,T	0.533	1.546	1.348
SVC	CO2,t1_1	0.556	1.662	1.307
SVC	CO2,t1_2	0.559	1.663	1.299
SVC	CO2,t1_3	0.553	1.648	1.347
SVC	CO2,t2	0.603	1.621	1.272
SVC	CO2,t3	0.477	1.695	1.339

6.4 対照実験

本実験の比較対象として行った対照実験について、説明する。表 6.2 のデータセットの CO2 濃度のデータと外れ値検知アルゴリズム LOF(Local Outlier Factor) を用いて異常検知、在室状況の変化を調べる。具体的には、CO2 濃度データの一次差分を求めて、そのデータに LOF を適用して異常検知を行う。具体的な一次差分の求め方としては、一次差分を取得したいタイミングの CO2 濃度と一つ前のタイミングで計測された CO2 濃度の比較を行うことで得る。

6.4.1 LOF(Local Outlier Factor) とは

LOF とは、Breunig らの研究 [41] で開発・発表されたアルゴリズムである。LOF は外れ値検知で用いるアルゴリズムの一種である。対象とするタイミングのデータと周辺のデータを用いて、データの密度を計算し、局所的な密度を求める。具体的には、数値の差が小さいほど、密度が大きくなる。求めたデータの密度を用いて、外れ値かどうかを判別する。この手法の強みは、データの分布に依存することなく外れ値検知を行うことができる点である。IoT ネットワークにおけるセキュリティと信頼性の向上を目的として、LOF を用いた研究がある [42]。

6.4.2 対照実験の結果

対照実験の結果について、説明する。Fig6.3 は、表 6.2 の Training データセットにおける、時系列に沿った CO2 濃度の分布を表した図である。Fig6.4 は、表 6.2 の Training データセットにおける、時系列に沿った CO2 濃度の一次差分の分布を表した図である。Fig6.5 は、Fig6.4 の CO2 濃度の一次差分についてのデータに LOF を適用した結果、異常と判断されたタイミングを紫の丸で表現し、Fig6.4 のグラフに重ねた図である。Fig6.6 は、表 6.2 の Test1 データセットにおける、時系列に沿った CO2 濃度の分布を表した図である。Fig6.7 は、表 6.2 の Test1 データセットにおける、時系列に沿った CO2 濃度の一次差分の分布を表した図である。Fig6.8 は、Fig6.7 の CO2 濃度の一次差分についてのデータに LOF を適用した結果、異常と判断されたタイミングを紫の丸で表現し、Fig6.7 のグラフに重ねた図である。Fig6.9 は、表 6.2 の Test2 データセットにおける、時系列に沿った CO2 濃度の分布を表した図である。Fig6.10 は、表 6.2 の Test2 データセットにおける、時系列に沿った CO2 濃度の一次差分の分布を表した図である。Fig6.11 は、Fig6.10 の CO2 濃度の一次差分についてのデータに LOF を適用した結果、異常と判断されたタイミングを紫の丸で表現し、Fig6.10 のグラフに重ねた図である。

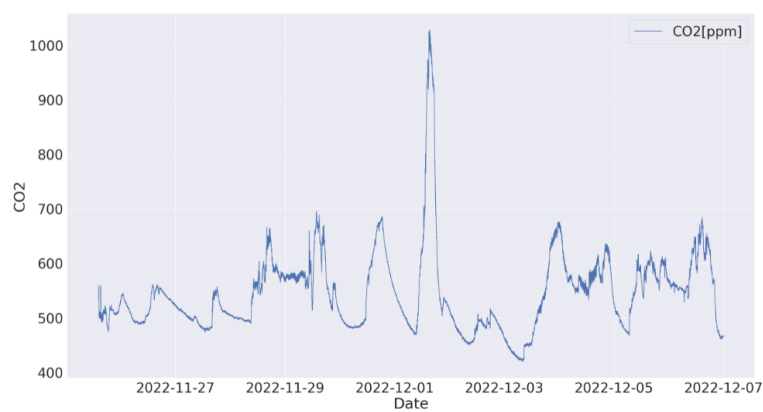


図 6.3: Distribution map of CO2 concentration in the training data

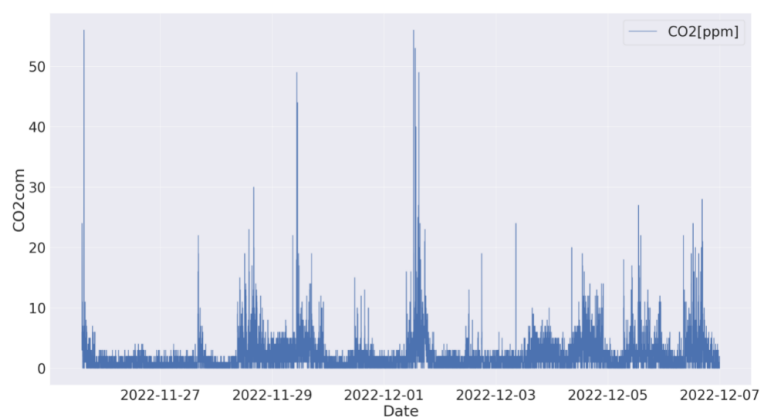


図 6.4: The first-order difference in the concentration of CO2 in the training data

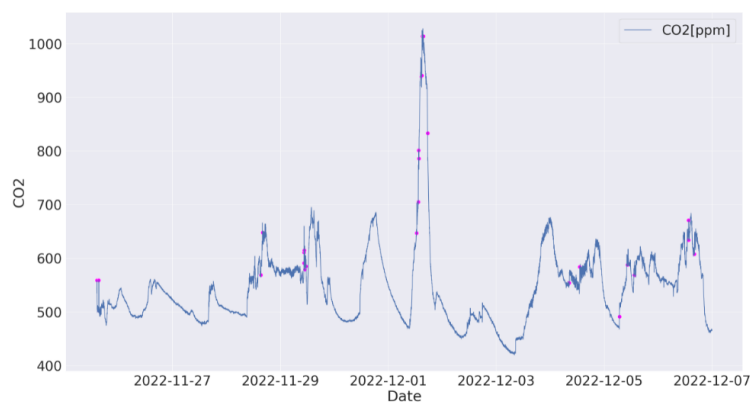


図 6.5: Outlier detection results in the training data

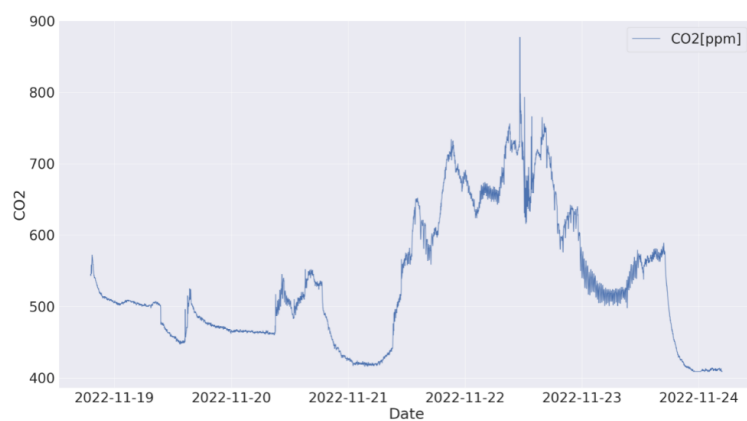


図 6.6: Distribution map of CO2 concentration in the test1 data

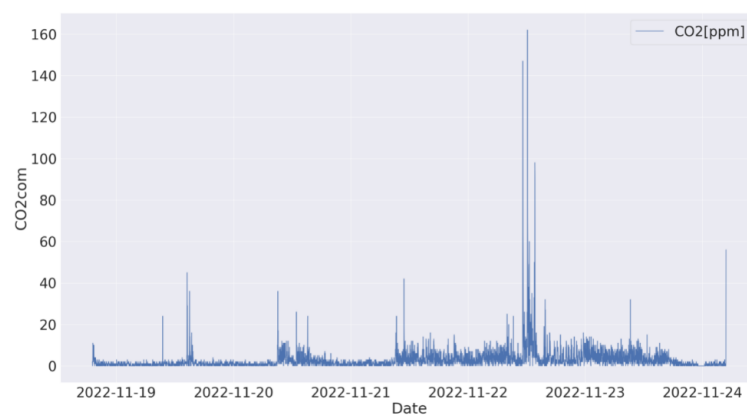


図 6.7: The first-order difference in the concentration of CO2 in the test1 data

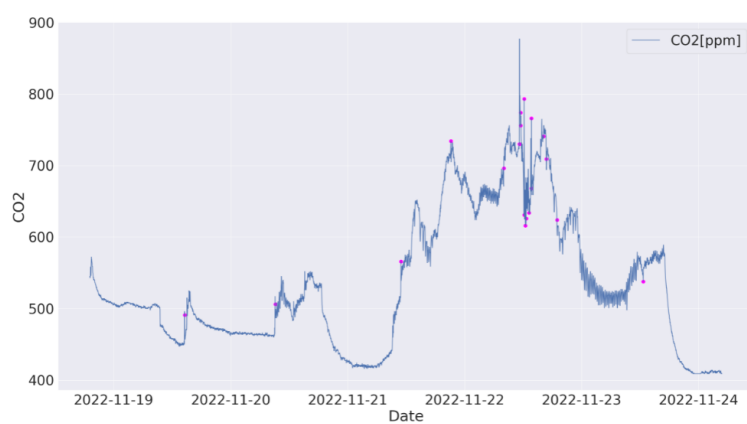


図 6.8: Outlier detection results in the test1 data

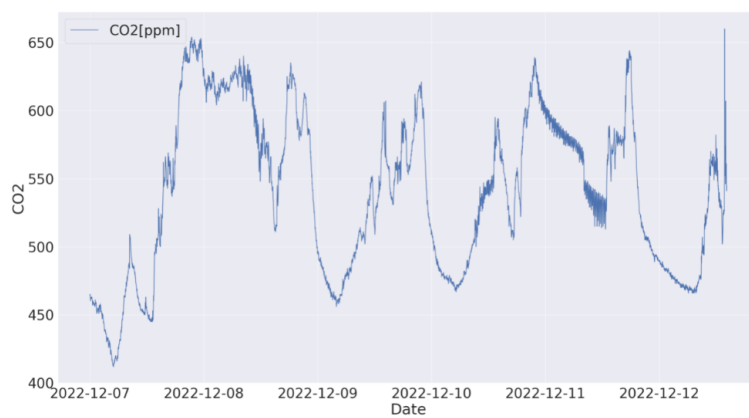


図 6.9: Distribution map of CO2 concentration in the test2 data

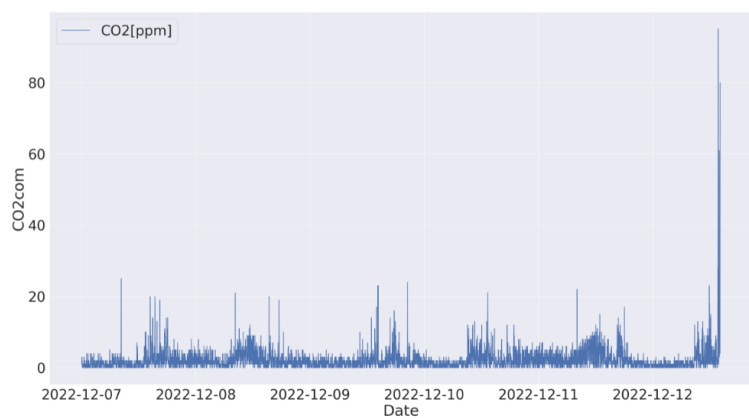


図 6.10: The first-order difference in the concentration of CO2 in the test2 data

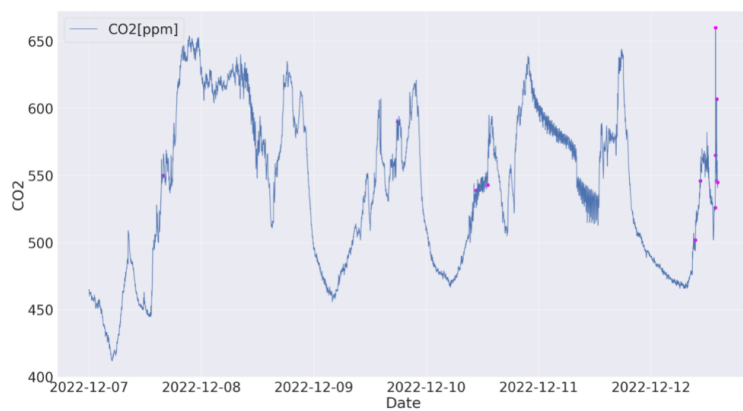


図 6.11: Outlier detection results in the test2 data

6.5 評価

対照実験との比較を含めた、本実験についての評価を行う。まず、MAE での評価について考察する。本実験の実験結果より、ロジスティック回帰と t_2 を用いた予測結果の精度が全モデルの中で最も高い。一方で、ランダムフォレスト、勾配ブースティングマシン、サポートベクターマシンを適用した場合には、MAE の数値が 1 以上、すなわち、実際に部屋にいた人数と予測結果の誤差が 1 人以上となってしまっている。テスト用データでの MAE の値と比較して、訓練用データでの MAE の値の方が低い、すなわち推定精度が高くなっている。これは過学習が発生してしまったと考察した。評価指標については、発生している誤差が最も精度の高いアルゴリズムであるロジスティック回帰の場合で 0.4~0.6 人の範囲の誤差が発生していることから、正解率を採用した場合は Candanedo の方法を用いた場合と比較して非常に低い数値になる可能性があると考えた。また、正解率を採用した場合、誤差が 1 人の場合と誤差が 5 人の場合を一律で不正解データとして処理してしまうため、精度の評価には不適切であると考えた。

対照実験の結果より、CO₂ 濃度が大きく変化する場面を除いて、CO₂ 濃度データと LOF を用いた外れ値検知の精度は高くないことが確認できた。例として、図 6.8 より、2022 年 12 月 1 日のデータでは外れ値が検知されていないことが確認できる。実際には実験環境への人の出入りは発生しており、これは正しく検知できていないと考察した。また、外れ値を検知したタイミングが人が入室したのか、退室したのか、その人数は何名であるのかといった状況の推定をすることが難しいため、外れ値検知を用いる場合、人が部屋にいるかどうかの検知には課題があると考えた。以上の事から、推定速度については、学習をする必要がないため、外れ値検知の方が早いですが、実際の状況を把握するという面では課題があると考えた。実際の人数を把握する場面においては、本実験で行った機械学習アルゴリズムと環境センサーデータを用いて、そのタイミングでの人数の予測をする方が在室状況把握という面からは望ましいと考えた。

第7章 結論

7.1 まとめ

冒頭で述べたように、世界的な電力不足が発生している。また、電力不足の問題を解決するためには、室内における電力制御が重要な要素である。そこで、適切な電力制御を実現するために、部屋における在室状況把握を行った。具体的には、IoT センサーと raspberry pi を用いて構成した IoT センサーネットワークと実字君における位置情報を組み合わせて、在室状況把握、部屋の人数の予測を実現することを検討した。提案手法を適用して、訓練用データセット、テスト用データセットを作成し、機械学習アルゴリズムを適用することで、室内の人数の予測を行った。予測の精度としては、ロジスティック回帰が最も精度が高く、MAE において 0.4~0.6 人という誤差での予測を実現している。また、ラプラスの式と差分法を参考に提案した二つ目の提案手法を使用して設定した温度データと CO2 濃度データを組み合わせて推定を行った場合の精度が最も高かった。目標である誤差 ± 1 以下を達成することができた。しかし、Candanedo の方法を用いる場合は 2 値分類であったため、高い精度で部屋に人がいるかどうかの識別ができていたが、本実験の場合は多値分類となっており、推定の難易度は上昇しており、一部の機械学習アルゴリズムでは誤差が 2.7 人という結果になってしまっている場合もある。

以上より、適切に提案手法を適用した温度と CO2 濃度の環境センサーデータと機械学習アルゴリズムを選択することによって、低い計算コストで高速かつ高精度に在室状況の推定を実現することができた。提案手法によって、入力データの次元数を減らすことによって高速に推定できたと考察した。

7.2 今後の展望

今後は、物理学的な観点からパラメータについて考慮しつつ、誤差の改善を実現する必要がある。また、誤差の改善の後に、位置の推定を行うことで更なる電力制御の参考情報を得ることができる。方法としては、BLE 通信可能な CO2 センサーを実験環境内に配置し、人の位置とセンサーからの距離によって CO2 濃度がどのように変化するのかを観察した上での、更なるデータセットの改良が挙げられる。また、スマートメータのデータを活用してデータセットの改良を行うことも可能であると考えた。

謝辞

本研究を進めるにあたり，日頃よりご指導を頂いた研究室の佐藤周行准教授に深く感謝いたします。新型コロナウイルスの影響下での研究生活を気遣い頂き，対面・リモート両面で柔軟に対応してくださりました。多くの時間を割いていただき，研究会やミーティングにて多くの助言をしていただきました。

また，佐藤周行研究室およびコンピュータネットワーク研究室の皆様に感謝いたします。合同練習会や合同ミーティングを企画し助言をくださった工藤知宏教授，関谷勇司教授，中山雅哉准教授，小川剛史准教授，宮本大輔准教授に厚く御礼申し上げます。研究室生活の相談に乗ってくださり生活を楽しくしてくださった先輩方や同期，後輩の皆様に感謝の意を表します。

最後にこれまで支えてくれた家族や友人に深く感謝いたします。本当にありがとうございました。

参考文献

- [1] 総務省, “第3章 第4節 本格的なデータ活用社会の到来 平成26年版情報通信白書,” 2014.
<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h26/pdf/n3400000.pdf>.
- [2] I. Alrashdi, A. Alqazzaz, E. Aloufi, R. Alharthi, M. Zohdy, and H. Ming, “Ad-iot: Anomaly detection of iot cyberattacks in smart city using machine learning,” in 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), pp. 0305–0310, 2019.
- [3] N. Ghosh, K. Maity, R. Paul, and S. Maity, “Outlier detection in sensor data using machine learning techniques for iot framework and wireless sensor networks: A brief study,” in 2019 International Conference on Applied Machine Learning (ICAML), pp. 187–190, 2019.
- [4] D. Assante and C. Fornaro, “An educational iot-based indoor environment monitoring system,” in 2019 IEEE Global Engineering Education Conference (EDUCON), pp. 1475–1479, 2019.
- [5] L. M. Candanedo and V. Feldheim, “Accurate occupancy detection of an office room from light, temperature, humidity and co2 measurements using statistical learning models,” Energy and Buildings, vol. 112, pp. 28–39, 2016.
- [6] scikit-learn 1.1.1 documentation.2022 06-10, “Choosing the right estimator–scikit-learn 1.1.1 documentation,” 2022. https://scikit-learn.org/stable/tutorial/machine_learning_map/.
- [7] 経済産業省資源エネルギー庁, “第1部 第3章 エネルギーをめぐる不確実性への対応エネルギー白書,” 2022. https://www.enecho.meti.go.jp/about/whitepaper/2022/pdf/1_3.pdf.
- [8] Y. Yau and S. Hasbi, “A review of climate change impacts on commercial buildings and their technical services in the tropics,” Renewable and Sustainable Energy Reviews, vol. 18, pp. 430–441, 2013.
- [9] 金. 研, 橋. 達也, and 廣. 純一, “Aiによる解析を利用した快適性を損なわない省エネルギー空調方式の提案と検証,” 空気調和・衛生工学会大会 学術講演論文集, vol. 2020, pp. 293–296, 2020.
- [10] 啓. 角田, 直. 荒井, and 和. 尾花, “時間窓可変型局所モデル構築に基づく co2 濃度からの滞在人流の推定,” in マルチメディア, 分散協調とモバイルシンポジウム 2058 論文集, vol. 2020, pp. 259–266, jun 2020.
- [11] S. N. Kalid, K.-H. Ng, G.-K. Tong, and K.-C. Khor, “A multiple classifiers system for anomaly detection in credit card data with unbalanced and overlapped classes,” IEEE Access, vol. 8, pp. 28210–28221, 2020.

- [12] L. Coppolino, S. D'Antonio, V. Formicola, C. Massei, and L. Romano, "Use of the Dempster-Shafer theory to detect account takeovers in mobile money transfer services," Journal of Ambient Intelligence and Humanized Computing, vol. 6, pp. 753–762, 2015.
- [13] M. Schlechtingen and I. Ferreira Santos, "Comparative analysis of neural network and regression based condition monitoring approaches for wind turbine fault detection," Mechanical Systems and Signal Processing, vol. 25, no. 5, pp. 1849–1875, 2011.
- [14] Y. Chen, R. S. Blum, and B. M. Sadler, "Optimal quickest change detection in sensor networks using ordered transmissions," in 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 1–5, 2020.
- [15] E. Nitzan, T. Halme, and V. Koivunen, "Bayesian methods for multiple change-point detection with reduced communication," IEEE Transactions on Signal Processing, vol. 68, pp. 4871–4886, 2020.
- [16] M. V. Ngo, T. Luo, H. Chaouchi, and T. Q. Quek, "Contextual-bandit anomaly detection for iot data in distributed hierarchical edge computing," in 2020 IEEE 40th International Conference on Distributed Computing Systems (ICDCS), pp. 1227–1230, 2020.
- [17] T. Luo and S. G. Nagarajan, "Distributed anomaly detection using autoencoder neural networks in wsn for iot," in 2018 IEEE International Conference on Communications (ICC), pp. 1–6, 2018.
- [18] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff, "Lstm-based encoder-decoder for multi-sensor anomaly detection," CoRR, vol. abs/1607.00148, 2016.
- [19] M. V. Ngo, H. Chaouchi, T. Luo, and T. Q. S. Quek, "Adaptive anomaly detection for iot data in hierarchical edge computing," 2020.
- [20] Q. D. La, M. V. Ngo, T. Q. Dinh, T. Q. Quek, and H. Shin, "Enabling intelligence in fog computing to achieve energy and latency reduction," Digital Communications and Networks, vol. 5, no. 1, pp. 3–9, 2019. *Artificial Intelligence for Future Wireless Communications and Networking*.
- [21] Z. Chen, W. Hu, J. Wang, S. Zhao, B. Amos, G. Wu, K. Ha, K. Elgazzar, P. Pillai, R. Klatzky, D. Siewiorek, and M. Satyanarayanan, "An empirical study of latency in an emerging class of edge computing applications for wearable cognitive assistance," in Proceedings of the Second ACM/IEEE Symposium on Edge Computing, SEC '17, (New York, NY, USA), Association for Computing Machinery, 2017.
- [22] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding," 2015.
- [23] Z. Li, Z. Xiao, and C. Lan, "Anomalous sensor detection based on nonlinear graph filter," in 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp. 1–4, 2019.

- [24] K. Thiagarajan, S. Kodagoda, N. Ulapane, and M. Prasad, "A temporal forecasting driven approach using facebook's prophet method for anomaly detection in sewer air temperature sensor system," in 2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 25–30, 2020.
- [25] G. E. P. Box and G. C. Tiao, "Intervention analysis with applications to economic and environmental problems," Journal of the American Statistical Association, vol. 70, no. 349, pp. 70–79, 1975.
- [26] M. Sanayha and P. Vateekul, "Fault detection for circulating water pump using time series forecasting and outlier detection," in 2017 9th International Conference on Knowledge and Smart Technology (KST), pp. 193–198, 2017.
- [27] M. A. Benjamin, R. A. Rigby, and D. M. Stasinopoulos, "Generalized autoregressive moving average models," Journal of the American Statistical Association, vol. 98, no. 461, pp. 214–223, 2003.
- [28] T.-M. Choi, Y. Yu, and K.-F. Au, "A hybrid sarima wavelet transform method for sales forecasting," Decision Support Systems, vol. 51, no. 1, pp. 130–140, 2011.
- [29] R. Hyndman and Y. Khandakar, "Automatic time series forecasting: The forecast package for r," Journal of Statistical Software, vol. 26, 07 2008.
- [30] R. J. Hyndman, A. B. Koehler, R. D. Snyder, and S. Grose, "A state space framework for automatic forecasting using exponential smoothing methods," International Journal of Forecasting, vol. 18, no. 3, pp. 439–454, 2002.
- [31] C. Bergmeir, R. Hyndman, and J. Benítez, "Bagging exponential smoothing methods using stl decomposition and box-cox transformation," International Journal of Forecasting, vol. 32, no. 2, pp. 303–312, 2016.
- [32] A. M. D. Livera, R. J. Hyndman, and R. D. Snyder, "Forecasting time series with complex seasonal patterns using exponential smoothing," Journal of the American Statistical Association, vol. 106, no. 496, pp. 1513–1527, 2011.
- [33] A. A. Waskita, H. Suhartanto, and L. T. Handoko, "A performance study of anomaly detection using entropy method," in 2016 International Conference on Computer, Control, Informatics and its Applications (IC3INA), pp. 137–140, 2016.
- [34] S. Rajasegarar, J. C. Bezdek, C. Leckie, and M. Palaniswami, "Elliptical anomalies in wireless sensor networks," ACM Trans. Sen. Netw., vol. 6, jan 2010.
- [35] P. Wang, Y. Han, J. Qin, B. Wang, and X. Yang, "Anomaly detection for streaming data from wearable sensor network," in 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), pp. 263–268, 2017.

- [36] T. Banerjee, B. Xie, and D. P. Agrawal, “Fault tolerant multiple event detection in a wireless sensor network,” Journal of Parallel and Distributed Computing, vol. 68, no. 9, pp. 1222–1234, 2008.
- [37] X. Wang, X. Luo, Y. Zhang, and X. Guan, “Detection and isolation of false data injection attacks in smart grids via nonlinear interval observer,” IEEE Internet of Things Journal, vol. 6, no. 4, pp. 6498–6512, 2019.
- [38] P. K. Sharma, T. De, and S. Saha, “Iot based indoor environment data modelling and prediction,” in 2018 10th International Conference on Communication Systems & Networks (COMSNETS), pp. 537–539, 2018.
- [39] S. Taheri and A. Razban, “Learning-based co2 concentration prediction: Application to indoor air quality control using demand-controlled ventilation,” Building and Environment, vol. 205, p. 108164, 2021.
- [40] 日本機械学会, JSME テキストシリーズ伝熱工学. 丸善出版, 2005.
- [41] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, “Lof: Identifying density-based local outliers,” in Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, SIGMOD '00, (New York, NY, USA), p. 93–104, Association for Computing Machinery, 2000.
- [42] L. Boukela, G. Zhang, M. Yacoub, S. Bouzefrane, S. Bagheri, and H. Jeldar, “A modified LOF based approach for outlier characterization in IoT,” Annals of Telecommunications - annales des télécommunications, vol. 76, pp. 145–153, Feb. 2021.

発表文献

- [1] 中村有佑&佐藤周行.IoT センサーネットワークと実空間における位置情報を用いた在室状況推定. 情報処理学会第 85 回全国大会,2023.(予定)