

Using Artificial Intelligence for Some Activity Recognition and Anomaly Identification Using a Multi-Sensor Based Smart Home System

by

Ashish Saragadam

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Public Health Sciences

Waterloo, Ontario, Canada, 2025

© Ashish Saragadam 2025

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Ambient Assisted Living (AAL) research frequently contends with limitations including reliance on supervised data, lack of personalization and interpretability, and evaluations in artificial laboratory settings. This study aimed to address these gaps by developing an unsupervised, personalized, and interpretable AAL system using low-cost sensors for long-term, real-world activity recognition and behavioural anomaly detection. A multi-modal sensor network (including contact, vibration, outlet, air quality sensors) was deployed in a single participant’s apartment for over 90 days. Primarily unsupervised machine learning techniques, augmented with interpretability methods (SHAP), were used to identify key activities (cooking, couch-sitting, showering) and detect personalized behavioural deviations. Minimally supervised approaches for showering detection were also accurately achieved using humidity data to address the shortcomings of unsupervised showering model. More importantly, the system effectively identified interpretable anomalies demonstrating the model’s capability to learn the individual’s normal behaviour in the home and identify anomalies, representing significant deviations from the participant’s established routines. In addition, the model was also able to be interpretable that allowed for the participant to understand why each anomaly occurred. This study confirms the feasibility of leveraging unsupervised, interpretable methods with affordable sensors for personalized, ecologically valid AAL, significantly reducing labelling dependence and enhancing system trustworthiness for scalable, unobtrusive health monitoring.

Acknowledgements

I would like to express sincere gratitude to my supervisor Dr. Plinio Pelegrini Morita, as well as the entire Ubiquitous Health Technologies Lab (Ubilab) at the University of Waterloo. I would also like to thank the entire research graduate programs coordinators for their invaluable work in supporting me and many other graduate students. I would also like to thank my family and friends for their invaluable support throughout it all.

In addition, I would like to acknowledge the financial support given to me by the University of Waterloo through various graduate research scholarships. I also want to thank the Canadian Institute for Health Research for selecting to be one of the recipients of their financial scholarship, the Canadian Graduate Scholarship - Master's.

Table of Contents

Author’s Declaration	ii
Abstract	iii
Acknowledgements	iv
1 Background	1
1.1 AAL Systems	1
1.1.1 Ambient Sensor Networks for Behavioural Monitoring	1
1.2 Significance and Implications in Aging Societies	2
1.3 Significance and Implications of AAL beyond Older Adults	4
1.3.1 Individuals with Disabilities or Injuries	5
1.3.2 Individuals with Chronic Health Conditions	5
1.3.2.1 Mental Health Monitoring and Support	5
1.3.2.2 Benefits for Healthy Adults and General Households	6
1.4 Current Research Gaps	6
1.4.1 Limitations with Supervised Activity Recognition in AAL	6
1.4.2 Personalized Behaviour Baselines for Anomaly Detection	8
1.4.3 Need for Adaptive and Evolving Models for Activity Recognition and Anomaly Detection	9
1.4.4 Research Done in Test Environments	9

1.4.4.1	Ecological Validity in Test Environments	9
1.4.4.2	Behavioural Distortions in Artificial Settings	10
1.4.4.3	Longitudinal Real-World Data and Limitations of Static Models	11
1.4.4.4	Financial and Practical Barriers	12
1.4.5	Challenged with Interpretability with Current Anomaly Detection in AAL	12
1.5	Purpose of the Present Study	13
1.5.1	Rationale	14
1.5.2	Hypotheses	15
1.5.3	Specific Objectives	15
2	Methods	16
2.1	Study Design	16
2.1.1	Participant Selection	16
2.1.1.1	Inclusion Criteria:	16
2.1.1.2	Exclusion Criteria:	17
2.1.1.3	Justification for Exclusion:	17
2.1.2	Participant Recruited	18
2.1.3	Sensor Selection and Placement	18
2.2	Data Collection Period	21
2.3	System Configuration and Data Acquisition	21
2.3.1	Raspberry Pi Setup	21
2.3.2	Domoticz Installation and Sensor Pairing	21
2.3.3	Time-Series Data Pipeline	22
2.3.4	Secure Upload to UWaterloo HealthCloud	22
2.4	Data Preprocessing and Cohesive Time-Series Construction	22
2.4.1	File Organization and Chronological Merging	22

2.4.2	Second-Level Time Index and Consistent Filling	23
2.5	Feature Engineering: Sensor Groupings and Semantic Labels	23
2.5.1	Door Grouping and Derived Variables	23
2.5.2	Column Renaming	24
2.6	Occupant Presence, Multi-Resident Intervals, and Room-Level Feature An- notation	24
2.6.1	Sorting by Time and Computing <code>row_dt_s</code>	24
2.6.2	Main-Door Cycles and Occupant Presence	24
2.6.3	Incorporating Occupant Presence	25
2.6.4	Generation of Binary ON/OFF Features from Illuminance and Outlet- Current Data	25
2.6.4.1	Illuminance Data	25
2.6.4.2	Outlet Data	26
2.6.5	Occupant Presence and Room-Level Location Inference Using Multi- Sensor Fusion	26
2.6.6	Data Cleaning, Multi-Presence Filtering, and Occupancy Session Identification	28
2.6.7	Unsupervised Bathroom Modelling	30
2.6.7.1	Humidity-Based Features	30
2.6.7.2	Environmental Average Features	31
2.6.7.3	Duration and Temporal Features	31
2.6.7.4	Unsupervised Shower Model	31
2.6.7.5	Unsupervised Bathroom Behavioural Anomaly Model	32
2.6.8	Unsupervised Living Room Modelling	34
2.6.8.1	Extraction of Living Room Activity Features	34
2.6.8.2	Unsupervised Living-Room Behavioural Anomaly Model	35
2.6.9	Unsupervised Bedroom Modelling	36
2.6.9.1	Extraction of Bedroom Activity Features	36

2.6.9.2	Bedroom Behavioural Anomaly Model	37
2.6.10	Minimally Supervised Showering Model	38
2.6.10.1	Innovative Approach for Supervised Showering Labels	38
2.6.10.2	Supervised Shower Model	39
3	Results	40
3.1	Dataset	40
3.2	Unsupervised Appliance On/Off Detection	41
3.3	Machine Learning for Showering Detection	42
3.3.1	Minimally Supervised Model for Shower Detection	42
3.3.1.1	Performance of Models at Various Labelling Levels	43
3.3.1.2	Model Interpretability	44
3.3.2	Unsupervised Machine Learning Model for Shower Detection	46
3.3.3	Comparative Analysis of Misclassifications and Achieving Full Accuracy	48
3.4	Anomaly Detection	49
4	Discussion	59
4.1	Objective 1: System Implementation and Data Collection	59
4.2	Objective 2a: Unsupervised Activity and Feature Identification	60
4.2.1	Outlet-Based Activity Recognition and Additional Features	60
4.2.1.1	Cooking Activity Recognition	61
4.2.1.2	Television ON Detection	61
4.2.1.3	Contribution of this Research to Outlet Modelling	61
4.2.2	Couch Sitting Detection Findings	62
4.2.2.1	Prior Work and Specific Implications	62
4.2.3	Showering Detection Findings	63
4.2.3.1	Prior Work and Specific Implications	64

4.3	Objective 2b: Personalized and Interpretable Unsupervised Anomaly De- tection	65
4.3.1	Bathroom Anomaly Model Findings	65
4.3.2	Living Room Anomaly Model Findings	66
4.3.3	Bedroom Anomaly Model Findings	67
4.3.4	Kitchen Anomaly Model Findings	67
4.3.5	Comparison with Prior Work	67
4.4	Key Takeaways & Contributions	68
4.5	Implications of the Research	69
4.6	Limitations and Future Work	70
5	Conclusion	72
	References	73
	APPENDICES	79

Chapter 1

Background

1.1 AAL Systems

With the advent of an aging population worldwide, one key area of research has been technologies that can assist older adults as they age.¹⁻⁴ Active Assisted Living (AAL) systems are a class of intelligent solutions that leverage modern information and communication technologies to support the health, safety, and independence of individuals in their daily lives.¹⁻⁴ In general terms, an AAL system integrates wearable and non-wearable sensors, devices, and software into the living environment of an individual to create a “smart” surrounding that can monitor the person’s status and assist as needed.¹⁻⁴ These technologies range from simple sensor networks in a home to more complex interactive systems, but all share the primary goal of enhancing the user’s quality of life and autonomy.¹⁻³ In essence, AAL systems provide multi-dimensional support to the individual and optionally their family, and their professional and informal caregivers, by continuously observing health and activity and by intervening or alerting when appropriate, thereby improving safety and well-being.¹⁻⁴

1.1.1 Ambient Sensor Networks for Behavioural Monitoring

AAL systems encompass a broad spectrum of technologies with one main type being sensor based monitoring systems designed for behavioural monitoring.¹⁻⁴ These systems utilize networks of sensors, wearable and non-wearable, to collect data about the resident’s activities. Sensors like motion detectors, pressure sensors on beds/chairs, door sensors, smart

appliances, continuously collect data which are then analyzed to recognize various daily activities and report on abnormalities in patterns of behaviour by identifying behaviour deviations using artificial intelligence (AI) techniques.^{1-3;5} For instance, a sensor-based AAL system might detect that an individual has not left their bed in the morning which is abnormal for them since they usually leave the bed at 8 am, and trigger an alert.^{1-3;5} One main subset of AAL systems designed for behavioural monitoring that is emerging recently are those that exclusively only use non-wearable sensors, specifically referred to as ambient sensors. By unobtrusively monitoring parameters like movement, sleep, and medication adherence, non-wearable sensor-based systems can provide early warnings of health issues that may occur when abnormal behaviour is detected without infringing on the privacy of the individual.^{1;3-7} Regarding privacy, a significant portion of past AAL solutions depend on wearable devices or cameras to track the user’s activities; however, wearables require the individual’s ability to consistently keep it on, which often times results is just not feasible or preferred. In addition, remembering to place wearables on can be especially inconsistent for those with cognitive or memory issues. In addition, many research studies focus on developing systems that also rely on camera-based sensors raises serious privacy and ethical concerns with participants, that make many individuals feel uncomfortable, and ultimately results in low adoption.^{1-3;5;7} By contrast, non-wearable, non-camera sensors, such as motion detectors, door contacts, bed pressure sensors, smart plugs, operate passively in the background and are generally more acceptable in private homes.^{1-3;5;7} These sensors can be strategically placed throughout the residence, for example, motion sensors in each room, contacts on entry doors and frequently used appliances, to provide coverage of the occupant’s routine activities.⁷ A representative AAL sensor layout could monitor events like room entry/exit, appliance use (e.g. microwave, TV), toilet flushes, temperature/humidity, and even ambient luminance.^{4;7} In practice, most real-world smart homes deploy around 16–20 ambient sensors to cover the key rooms of a typical single-family dwelling.⁷ This instrumentation enables continuous logging of what is happening (sensor event on/off) and where, by mapping each sensor to a location or object.⁷

1.2 Significance and Implications in Aging Societies

Using AAL systems for health anomaly detection are mostly used for developing targeted solutions for older adults, this is because the global demographic shift towards an aging population presents multifaceted challenges for healthcare systems worldwide.^{4;8} According to the World Health Organization (WHO), by 2050, the world’s population of people aged 60 years and older will double to 2.1 billion, with the number of persons aged 80 years or

older expected to triple between 2020 and 2050 to reach 426 million.^{9;10} This aging trend is global, with approximately 80 percent of older adults projected to reside in low- and middle-income countries by 2050.^{9;10} Moreover, about 80 percent of older adults aged 65 and above live with at least one chronic disease, and 68 percent have multiple chronic conditions, complicating healthcare management and significantly increasing resource utilization.^{11;12} As the proportion of older adults increases, many issues surface that collectively exacerbate the demand for healthcare services.¹³ The natural aging process also leads to a decline in physical and cognitive functions, increasing the need for supportive care, rehabilitation services, and assistive technologies.¹³ Specific health conditions and comorbidities, such as osteoporosis or balance disorders (increasing fall risk), hypertension or atrial fibrillation (increasing stroke risk), and poorly managed diabetes (leading to sudden medical episodes), which are more likely to develop as people age, often place older adults living independently at an increased risk of acute health events like falls, strokes, or sudden medical episodes that demand immediate medical attention.^{6;13} Delayed detection of these events significantly worsens health outcomes, leading to prolonged recovery periods, decreased independence, and increased long-term healthcare costs.⁶ However, many older adults still continue to want to live alone because of comfort, the lack of availability in long term care centres or expensiveness of long term centres or home health care professionals.^{2;3} It's also very hard to know when adults have gotten to a point where they need support to avoid adverse health outcomes.⁶ Understanding and addressing these complexities are crucial for sustaining healthcare efficacy and ensuring accessibility for an aging society.¹³ In response to these challenges, monitoring and maintaining the well-being of older adults within their living environments emerges as a critical aspect of contemporary healthcare strategy.¹⁴ Whether that is to determine and track their health prior to more disruptive healthcare interventions including moving to long term care centres or employing home health nurses, or in supplementation with healthcare professional care to ensure and monitor their health, tracking older adult health allows for more informed timely intervention when required.^{2;3} Timely recognition of subtle behavioural changes or early symptoms indicating declining health becomes crucial for preventive intervention and care optimization.^{2;3} AAL systems support the aging population by enhancing their safety and supports aging-in-place while minimizing intrusiveness.^{2;3} These systems utilize a network of ambient sensors that are integrated into the living environment of seniors.^{2;3} These sensors can gather a wide range of data without requiring any active input from the user, including motion detection, usage of appliances, and other environmental variables.^{2;3} They continuously monitor the physical environment and its activities, thereby enabling the passive collection of data pertinent to an individual's behaviour in the home.^{2;3} This data is then processed using advanced algorithms to detect patterns and anomalies in behaviour, which can effectively monitor older adults confirming to their care team that the person is doing well, indicate changes

in health status such as unusual behaviour in the home which may indicate early signs of cognitive decline such as prolonged duration in any one room, repeated behaviour that is normally not done more than once in a day.^{2;3} Continuous monitoring and assistance mean that health deteriorations can be caught early and managed proactively, such as noticing changes in erratic behaviour at night, or detecting signs of frequent bathroom visits throughout the day, which can not only prevent minor issues from becoming severe, but also promote preventive care and healthy behaviours.^{2;3} AAL technologies emerge as a key solution to track and monitor care.^{1;2} In addition with fewer young people available to care for an increasing elderly population, technology that supports older adults' daily needs can offset labour shortages in elder care and provide relief to family members who act as caregivers.^{1;15}

1.3 Significance and Implications of AAL beyond Older Adults

Although the concept of AAL systems has predominantly been associated with older adults, the underlying technologies anomaly systems, possess broad applicability beyond the older adult population.

Modern definitions of AAL emphasize support for individuals with varying needs and abilities, not only seniors.^{3;16;17} Scholars conducting reviews of AAL consistently call out the importance of inclusive AAL system research.³ Research no longer views AAL as exclusively for older adults but rather as a subset of ubiquitous computing and assistive tech that can cater to “persons with different abilities and skills” across the lifespan.^{2;3;16;17} Likewise, in 2022, WHO and UNICEF released the Global Report on Assistive Technology, which revealed that over 2.5 billion people currently need one or more assistive products, a number expected to rise above 3.5 billion by 2050 as the global population grows and ages.¹⁸ Crucially, this massive figure encompasses not just older adults but people of all ages, from children with disabilities to working-age adults with chronic conditions and managing various mental health conditions, underlining that AAL solutions, is a worldwide need, not an age-restricted niche.¹⁸ As it may not be immediately clear as to how non-older adults may benefit, the section below highlights some work proven to show AAL systems focused on detecting behaviour anomalies can be effective among a non-older adult demographic.

1.3.1 Individuals with Disabilities or Injuries

All individuals, no matter their age, with physical, sensory, or cognitive disabilities can greatly benefit from sensor-based smart home adaptations.¹⁷ By equipping living environments with anomaly based systems, these technologies enhance insight should something wrong happen, which in turn reduces reliance on caregivers and empowers those with disabilities to perform tasks on their own with the confidence that should anything happen, the system is able to detect and send timely alerts to the support team and themselves.¹⁷

1.3.2 Individuals with Chronic Health Conditions

AAL and smart home solutions can also be tailored to people of any age who live with chronic illnesses to support health monitoring and timely interventions at home.¹⁹ Sensor technologies can continuously track relevant behaviour in the home that could become health indicators and detect critical events, providing a safety net for patients and peace of mind for families.¹⁹ For conditions like diabetes or cardiovascular and neurological conditions, home technology is enabling continuous, passive oversight of health metrics through smart home health devices that monitor eating patterns and activity levels, giving patients automatic feedback to improve daily management of diabetes.¹⁹ One study notes that smart home systems may help people with MS identify patterns and locations in the home associated with increased fall risk, enabling targeted fall-prevention measures.^{19;20} More generally, such technologies improve health outcomes by monitoring symptoms, medication adherence, physical activity, and ultimately improving quality of life for people with conditions like heart disease, COPD, or diabetes.¹⁹

1.3.2.1 Mental Health Monitoring and Support

Another frontier for AAL technologies is in supporting individuals with mental health needs, including conditions like depression, anxiety, or cognitive/emotional disorders, through passive monitoring and early intervention. Many behavioural and physiological signals of mental well-being can be unobtrusively captured by home sensors.²¹ Changes in daily routines or habits can serve as important markers of deteriorating mental health.²² For example, reduced household mobility, staying in bed longer, decreased room-to-room movement, neglecting daily activities may indicate emerging depression.²² By analyzing these patterns with machine learning, anomaly based AAL systems can detect these when an individual's behavioural rhythms deviate from their healthy baseline, effectively identifying early warning signs of mental distress. This information can trigger alerts or supportive interventions,

such as notifying a remote clinician, or even adjusting the environment (dimming lights, playing calming music) if high anxiety is detected.^{21;22}

1.3.2.2 Benefits for Healthy Adults and General Households

The utility of AAL systems also extends to healthy populations through proactive health monitoring and early detection capabilities.^{2;3;6} Sensor-based technologies can identify subtle deviations from established behavioural patterns, such as changes in sleep quality, decreased physical activity, or increased sedentary behaviour, providing critical insights before explicit health issues arise.^{2;3;6} Early identification of these deviations enables timely intervention—such as adjustments to sleep environments, encouragement toward healthier habits, or prompting medical consultation if necessary—thus promoting preventive health and sustained well-being to any individual.^{2;3;6}

1.4 Current Research Gaps

Although AAL technology holds immense potential and more research is being done in this area as of late as a result, there are still some key limitations with related work.^{1-3;23} This section critically examines these limitations, providing essential context to effectively introduce the subsequent section, which details how the present study aims to address and overcome these identified gaps.^{2;3;23}

1.4.1 Limitations with Supervised Activity Recognition in AAL

Accurately recognizing and modelling daily activities is a foundational component of effective AAL anomaly detection systems.^{3;24} In practice, AAL deployments rely on continuous, unobtrusive monitoring of individuals’ routine activities; which can then, using a variety of different data analysis techniques, may result in effective detection of these activities.^{3;24} In essence, without robust activity recognition, an AAL anomaly detection system cannot reliably detect when something is wrong in relation to that daily activity, and because that activity is done so commonly it has to be a key part of what the anomaly model uses and detects if there are anomalies in relation to those activities. Accurately modelling everyday behaviours thus underpins the ability of these anomaly detection systems’ effectiveness.^{3;24}

Current approaches to activity recognition in AAL largely rely on supervised techniques.^{3;24} Many early AAL systems employed supervised rule-based or threshold-based

techniques, for example, triggering an alert if a sensor detects no movement for a prolonged period or if a door opens at an unusual time.^{3;24} These rule-based methods can be effective for detecting straightforward events (such as an emergency when a threshold is crossed).²⁴ However, they are inherently limited: hand-crafted rules cannot easily accommodate the diversity and complexity of human routines.¹ Daily activities often vary greatly between individuals and contexts, so a fixed set of rules tends to be prone to false alarms or misses when behaviour deviates slightly, and difficult to generalize beyond the specific scenario for which they were written.¹ Supervised machine learning models, including modern deep learning approaches, have been introduced to recognize more complex activity patterns.²⁴ Yet, supervised learning in AAL also are limited in real-world applications. First, training robust supervised models requires large, representative datasets that are fully labelled covering the full spectrum of the target daily behaviours.^{2;3;7;24} AAL installations rarely have the luxury of long data collection phases with full labelling because it's just too impractical for the participant or a researcher to be able to collect those labels sustainably over time, especially if there are multiple participants.³ Supervised techniques heavily rely on someone providing labels every single time that activity is done, this means it may not even be feasible in studies that occur in real-world settings because participants may not be willing to or able to provide exact labels everytime they did a certain activity, especially for the long period of time required for a model to learn all the different ways an activity can be done "normally".^{2;3;25} In summary, while supervised techniques are primarily how activity recognition is done for the purposes of anomaly detection systems today, they are inherently constrained. They require intensive upfront knowledge, either in the form of labelled rules or labeled examples, and often lack the flexibility to cope with the variability found across individuals, cultures, and home environments in the real world. This motivates exploration of more adaptive or unsupervised techniques that can learn residents' routines without exhaustive labelling.¹

Another valuable implication of adopting unsupervised or semi-supervised methods in AAL anomaly detection studies is the potential increase in participant recruitment and retention.^{2;3;7} Conventional supervised techniques, which necessitate participants to actively engage in continuous annotation of their activities or behaviours, inherently require significant cognitive and physical effort from the participant, often causing reluctance or attrition among study cohorts.^{2;7} In contrast, unsupervised methods or semi-supervised greatly reduce this participation burden, as they eliminate the necessity of explicit behavioural labelling.^{2;3;7} This lowered requirement not only broadens the appeal of participation to a larger demographic, including those who might otherwise avoid studies involving meticulous documentation tasks, but also enhances participant compliance and satisfaction, ensuring more robust and representative datasets.^{25;26}

Furthermore, unsupervised or semi-supervised anomaly detection methodologies confer an important psychological and behavioural advantage by mitigating participants’ awareness of being continuously observed, thereby reducing observer-induced behavioural alterations, often referred to as the Hawthorne effect.²⁷ Supervised methods, frequently reliant on active participant engagement through constant labelling or direct observation by research personnel, either in person, remotely via cameras, or through regular lab visits, inadvertently remind participants of ongoing surveillance.³ This heightened awareness has been documented to cause subtle but significant deviations from participants’ natural, routine behaviours, potentially compromising the authenticity of the collected data and, subsequently, the reliability and validity of study outcomes.^{25;28} In stark contrast, unsupervised or semi-supervised techniques unobtrusively operate in the background for at least most of the study, facilitating the collection of genuine behavioural patterns and anomalies that more accurately reflect natural living conditions, and validating through some ground truth at the beginning or end.³ This not only preserves ecological validity but also significantly enhances the practical applicability and interpretative power of the resultant AAL systems.^{7;26}

1.4.2 Personalized Behaviour Baselines for Anomaly Detection

A consistent finding in recent AAL research is that normal behaviour is highly individual, necessitating personalized baselines learned over time.³ Rather than relying on generic profiles, systems must model each person’s routine through long-term observation.^{3;29} This is especially possible now than ever before due to recent advancements in technology that allow for large scale data collection and advancements in machine learning techniques and computing allow for the development of models designed and customized for each individual’s behaviour automatically.^{3;29} General “off-the-shelf” solutions are often inadequate – as one perspective article notes, given the idiosyncratic behaviours of residents and unique home settings, it is “not realistic” to assume a pre-trained model will work without personalization.^{3;29} In other words, what is normal for one individual may be abnormal for another; hence, systems must be individual-specific so that they can detect anomalies solely on that individual’s “normal” behaviour rather than a group which logically results in more false negatives and false positives due to a lack of customization of the model.^{3;29} Overall, the literature strongly supports that normality is subjective, AAL platforms should learn each user’s normal behaviour profile through long-term monitoring as a foundation for reliable anomaly detection.

1.4.3 Need for Adaptive and Evolving Models for Activity Recognition and Anomaly Detection

Another key theme is the requirement for adaptive systems that evolve their models over time.^{2;3;7;24;29} Human behaviour and activities are not static; changes due to aging, illness, or routine adjustments can render a once-accurate model obsolete if it cannot adapt.^{2;3;7;24;29} However a lot of research in the area focuses on one-off training models which do not have the capability of being built in a way that takes feedback to evolve its definition of "normal" and what is considered an anomaly.^{2;3;7;24;29} Instead, the system's analytics should accommodate an evolving baseline, updating what is considered normal as new data arrives. For example, one survey notes that data analysis methods must allow "dynamically weighting or customizing certain parameters" in response to changes, with models of normality being "updated in an adaptive way as people's needs or health change".^{3;29} The ability to reprioritize features or redefine "normal" on-the-fly is especially crucial in longitudinal deployments: an individual's activity levels might change slowly over months or years, and an adaptive AAL anomaly solution will accommodate this evolution rather than raising false alarms. Ultimately, the consensus in recent literature is that adaptive, learning systems outperform static ones in Ambient Assisted Living contexts, by maintaining alignment between the model and the user's current state.^{2;3;7;24;29}

1.4.4 Research Done in Test Environments

1.4.4.1 Ecological Validity in Test Environments

A significant limitation in much of the existing AAL research is the use of laboratory test apartments or simulated home environments that lack ecological validity. In these artificial setups, participants are often asked to perform scripted daily tasks in a highly instrumented "smart home" laboratory.³⁰ While such controlled experiments are useful for initial evaluations, they do not accurately reflect authentic daily behaviours in real homes.³⁰ There are several reasons why test apartments fail to reflect genuine daily behaviour. Typically, they involve a small number of participants (sometimes young graduate students or healthy adults) who follow a predefined script of activities for a short duration.^{1-3;30} This setting cannot replicate the long-term habits, multitasking, and interruptions characteristic of normal living.^{1-3;30} Environmental factors are also simplified: the test home is usually tidy, with ideal sensor placements and no pets, visitors, or other occupants to introduce noise.^{1-3;30} In contrast, actual homes can be unpredictable, appliances usage may not be consistent, and daily routines can change due to mood or health. Moreover, test

environments typically lack the variability inherent in real-world settings, such as diverse home layouts, clutter, lighting conditions, and interruptions caused by visitors. These factors, while seemingly mundane, play a crucial role in determining how AAL systems interact with their environment and their users. For instance, motion sensors may behave differently in homes with irregular layouts or in spaces with obstructed views. Without real-world data, AAL technologies risk being ill-equipped to handle these challenges. As a result, algorithms fine-tuned to clean laboratory data can misidentify activities when faced with the richer, more chaotic sensor data from real residences.^{1-3;30} Indeed, an analysis by Cicirelli et al. (2021) found that many published activity recognition approaches were validated only on “limited contexts, such as in labs or selected use cases,” and often on different public datasets, making it hard to judge how they would perform in a true home environment.^{1-3;30} The variability of experimental setups means reported accuracy results may not carry over to new settings, underscoring an external validity problem.^{1-3;23;30} Perhaps most compelling are findings from studies that moved out of the lab and into real living spaces.^{1-3;23;30} Aloulou et al. (2013) describe a 14-month trial deploying an AAL system in a nursing home, an effort that revealed many issues not discovered during prior laboratory testing.²³ During the real deployment, unforeseen problems emerged.²³ For example, some sensors failed because the pipeline and system was not robust enough to collect data without consistent researcher supervision and fixing any minor errors.²³ In addition, the logic used to detect certain activities did not account for some of the variance the residents performed those activities.²³ This gap arises precisely because the controlled test apartments could not simulate the full complexity of residents’ day-to-day life. Such evidence highlights that evaluations confined to artificial settings can give a false sense of confidence. For AAL research to truly progress, field studies in actual homes, or highly faithful simulations thereof are needed to capture authentic activity patterns and challenges. In summary, the prevailing reliance on test apartments and scripted scenarios is a major limitation: it fails to ensure that AAL anomaly-based systems will work robustly in unconstrained, real-world environments where daily activity patterns are far less predictable.^{1-3;23;30}

1.4.4.2 Behavioural Distortions in Artificial Settings

Studies conducted in test apartments fail to fully replicate the naturalistic behaviour of that individual and environmental conditions encountered in real homes. When research participants know they are being observed, as is invariably the case in instrumented smart home labs, their behaviour changes in response.³¹ This phenomenon, known as the Hawthorne effect and is well documented in social and health research.³¹ Simply put, individuals tend

to alter or regulate their activities because they are aware of being monitored.³¹ An individual might stick closely to the instructed tasks, move more cautiously, or avoid certain spontaneous actions, thereby deviating from how they would normally conduct their routine at home.^{3;31} Such reactive behaviour introduces bias: the collected activity data no longer represents the person’s true habits, but rather a performance shaped by observation.^{1;3;31} This can mislead researchers or algorithms into modelling a “cleaner” or more structured version of daily life than actually occurs. Unfamiliarity with the environment can further distort behaviour.^{1;3;31} Participants in a test apartment do not have the same comfort and habit patterns as they do in their own dwellings.^{1;3;31} For example, an older adult in an unfamiliar kitchen might not know where utensils are stored or might hesitate more when cooking, whereas in their home they would perform the task on autopilot. They may also avoid using certain features of the space that they find confusing, whereas at home they have adapted their environment to their needs.^{1;3;31} Together, being in a new setting and under observation can lead to atypical sequences of actions, effectively, the person’s “daily activities” observed in the lab are not truly their daily activities as lived normally.^{1;3;31} Participants may also try to do everything the way they believe that the researchers expect them to do it, knowing they are part of a study, which reduces the natural variability and improvisation that characterize real-life behaviour.^{1;3;31} These behavioural distortions are not just hypothetical, they have been noted as a concern in the AAL community.^{1;3;31} For instance, a scoping review by Blackman et al. (2016) emphasizes human factors in AAL, implying that user experience and natural interaction with technology are critical.² In summary, artificial evaluation settings often lead to behaviour that is systematically different from genuine daily life; hence, making a real-world data collection of various activities very important to developing robust anomaly detection models.

1.4.4.3 Longitudinal Real-World Data and Limitations of Static Models

The concept that most AAL anomaly research is done in test environments point to another reason for its limitation, the inherent lack of data collection long-term. Test environment research forces research to be at-most a few hours long as its not feasible to collect over the long term. One recent work warns that models built on such data may fail to generalize to real homes and real behaviours by observing that solutions developed and tested only in lab settings “may not be suitable for real-world deployments” because they occur over the short-term.³² Extended observation periods allow systems to build a robust model of normal behaviour under realistic living conditions. Indeed, long-term monitoring is often necessary to discern gradual changes, one survey notes that “long periods of observation are required when changes in habits require detection,” and results can only be evaluated

after sufficient data has been gathered to establish reliable normal-behaviour models.³ Studies using real-world datasets have also revealed issues that static models overlook, such as sensor noise, irregular but benign behaviour variations, and personalization needs.^{3;32} The consensus is that longitudinal real-life data is indispensable for AAL: it not only improves the accuracy of activity recognition and anomaly detection, but also highlights the limitations of static snapshots.^{3;32} These findings imply that incorporating long-duration monitoring and continuously updating models will yield more robust and clinically useful health monitoring solutions.

1.4.4.4 Financial and Practical Barriers

One of the main reasons why in-home studies are under-represented in current research is that planning and collecting real-world ambient sensor data presents significant financial, practical, and ethical challenges. The financial and logistical demands associated with real-home sensor deployments are substantial, requiring extensive investment in hardware installation, continuous maintenance, and robust data management infrastructures capable of handling large-scale data collection over prolonged periods.^{2;3;33} However, with recent advancements in computing this barrier is continuing to decrease, which as a result allows for more research to be done relatively affordably in real-world environments.^{1;29}

By successfully overcoming these financial and ethical hurdles, research can unlock the transformative potential of AAL technologies, facilitating personalized, real-time interventions that dramatically improve quality of life and health outcomes in realistic settings.^{2;3}

1.4.5 Challenged with Interpretability with Current Anomaly Detection in AAL

Many Ambient Assisted Living systems rely on anomaly detection models (e.g. deep learning or other complex algorithms) that behave as “black boxes,” providing little insight into their reasoning. Researchers have noted that current AAL anomaly detection approaches often lack transparency and interpretability, which poses serious limitations for end-users and caregivers. For example, a recent survey of anomaly detection in smart environments observes that a lack of transparency/explainability in complex machine-learning models can hinder user acceptance – users are often reluctant to trust a system they “do not understand or cannot easily interpret”.^{34;35} Similarly, in the context of remote health monitoring for dementia patients, Bijlani et al. (2022) point out that “current methods for health care

anomaly detection” suffer from “low explainability to clinical monitoring teams and caregivers”.^{35;36} This opaqueness means when an alert is triggered in a smart home or AAL setting, participants and/or caregivers may have no clear explanation of why, they see that an anomaly was flagged but not which sensor, feature, or behaviour caused it. Trust in the system is eroded if users cannot understand the alerts. Jovanović et al. (2022) note that as AI models become more capable in AAL, “the absence of explanations behind automated behaviors raises uncertainty with users due to a lack of understanding of how specific decisions are. made”.³⁷ In other words, an unexplained anomaly alert can leave older adult participants and their caregivers unsure whether to believe or act on it. Moreover, the inability to interpret alerts makes it difficult to intervene effectively or provide corrective action. Yahaya et al. (2021) emphasize that most anomaly detectors in AAL “can identify the anomalies but not the sources of the anomalies.” As a result, the system cannot point to the precise behaviour or sensor reading that went wrong.^{35;38} This lack of insight impedes timely and appropriate intervention, a caregiver receiving a vague “anomaly detected” notification might not know whether it was triggered by a skipped medication, a fall, unusual sleep patterns, or a sensor error. In short, black-box alerts without context can lead to frustration, false alarms, and missed opportunities for preventative care. Because of these challenges, there is a strong consensus in recent literature that AAL anomaly detection models must become more interpretable and explainable. Across high-impact journals and conferences in health informatics and pervasive computing, researchers explicitly argue for “opening up” the black-box models to foster user trust and to enable meaningful responses to alerts.^{34;35;37} The literature is essentially calling for a shift from inscrutable anomaly detectors to “glass box” models whose alerts come with human-readable justifications.

By making anomalies transparent and contextualized, such systems support clinical decision-making and behaviour change interventions, the caregiver or clinician knows what to follow up on.^{34;35;37} In sum, recent high-impact works in AAL and health informatics strongly advocate for explainable anomaly detection, because it builds trust with users, enables caregivers to verify and understand alerts, and ultimately makes the technology more effective as a tool for health monitoring and intervention. The clear message is that next-generation AAL systems should prioritize “white-box” anomaly detection models.

1.5 Purpose of the Present Study

The preceding *Current Research Gaps* section highlighted five persistent limitations in AAL research: (i) dependence on supervised, heavily-labelled datasets; (ii) general, population-level anomaly models that ignore individual baselines; (iii) evaluations performed in short, arti-

ficial laboratory settings. (iv) interpretable anomaly models that allow for the participant to understand why something is an anomaly.^{2;3;24;29;35}

This study directly addresses each limitation by deploying an AAL system that does the following:

1. Reduce exhaustive labelling

- Provide validation to see if unsupervised learning and clustering can be used to identify some activities, *cooking, showering, sitting on the couch*, mainly using data from low-cost ZigBee/Z-Wave multi-modal sensors.¹

2. Personalize anomaly detection

- Train a one-participant baseline that learns what is “normal” *for that individual* and flags deviations such as prolonged room occupancy or atypical room-transition sequences, thereby acknowledging person-specific routines.²⁹

3. Maximize ecological validity

- Collect continuous, real-world data over ≥ 90 days in the participant’s own apartment, capturing natural variability, household clutter, visitors, and daily interruptions omitted from laboratory studies.^{30;32}

1.5.1 Rationale

Advancing smart-home systems for independent living and health monitoring requires overcoming limitations in both data collection and analysis. Therefore, this research employs a multi-modal sensor network implemented within an edge-based architecture. This design is crucial for capturing the necessary richness and granularity of real-world behavioural data over extended periods, maximizing ecological validity. This specific infrastructure provides the foundation for applying interpretable models that learn individual baselines and adapt to continuous data streams, thereby addressing key gaps identified in the AAL literature.

¹Sensors include motion, temperature, humidity, door-contact, and smart-outlet measurements.

1.5.2 Hypotheses

Hypothesis 1: The in-home sensor system will reliably collect multi-modal data with minimal disruptions, enabling robust and real-time data transfer to a centralized server.

Hypothesis 2: The collected data will support the development of identifying some activities, and developing unsupervised machine learning models capable of detecting behavioural anomalies.

1.5.3 Specific Objectives

Objective 1: Design and Implement the In-Home Sensor System

Develop a smart-home system for in-home data collection via a multi-sensor setup (e.g., motion, temperature, humidity, door contact, smart outlets) using a Raspberry Pi and storing raw data locally in InfluxDB and securely pushing encrypted de-identified data to a University of Waterloo server for analysis.

Objective 2: Analyze and Model In-Home Sensor Data

Utilize the collected data to perform exploratory and predictive analyzes, focusing on:

- Objective 2a: Identify key features for presence detection, room-level localization, and activity discovery (cooking, showering, couch-sitting).
- Objective 2b: Building *unsupervised* machine learning models for behavioural anomaly detection using some activities identified as well as other features.

Chapter 2

Methods

This section elaborates on the structured methodology employed to develop accurate presence recognition, activity recognition, anomaly detection and pattern recognition models using sensor data collected in a robust in-home study. The research integrates exploratory data analysis (EDA) with model development to detect room presence, recognize specific activities, and identify anomalies in an uncontrolled, real-world setting. This methodology ensures that insights gained from data patterns and feature extraction are translated into high-performing machine learning models capable of adaptive learning over time.

The study utilizes a quantitative, data-driven approach, progressing systematically from sensor system design and deployment to processing and visualizing the data to eventually unsupervised model development and validation. The methodology emphasizes the collection and analysis of multi-modal data from commercially available sensors over a three-month period. The results will help optimize sensor configurations and activity recognition techniques for practical deployment.

2.1 Study Design

2.1.1 Participant Selection

2.1.1.1 Inclusion Criteria:

- **Age Requirement:** Participants must preferably be aged 65 years or older to align with the study's focus on aging populations. If older adults fitting this criterion

are unavailable, participants aged 18–65 years will be recruited to evaluate activity recognition in younger demographics.

- **Mobility:** Participants must be able to perform their daily activities autonomously to ensure data reflects independent living conditions.
- **Living Situation:** Participants must live alone in a private residence to avoid data contamination from multiple occupants. The research team will work with participants to notify the research team to remove periods of data when the participant has visitors.
- **Consent and Comprehension:** Cognitive ability to provide informed consent and comprehend study procedures is required for ethical and practical reasons.
- **Language Proficiency:** Fluency in English is required to ensure participants fully understand the study’s procedures and their involvement.

2.1.1.2 Exclusion Criteria:

- **Mobility:** Candidates with significant physical impairments or disabilities that limit their ability to perform daily activities autonomously will be excluded.
- **Health Status:** Participants requiring regular intensive medical intervention or whose conditions severely impede daily functioning will be excluded.
- **Consent and Comprehension:** Individuals unable to provide informed consent or comprehend study requirements will not be included.
- **Language Proficiency:** Non-fluency in English will preclude participation to ensure effective communication and understanding of the study.

2.1.1.3 Justification for Exclusion:

The study focuses on autonomous activity recognition. Excluding individuals with significant mobility impairments, intensive medical needs, or cognitive impairments ensures data validity and aligns with the research objectives. Fluency in English is necessary to facilitate clear communication and understanding of the study’s procedures.

2.1.2 Participant Recruited

To conclude, unfortunately an older adult was not able to be recruited for this experiment. 1 participant was recruited for this study. The participant was a 25 year old working professional that met the rest of the inclusion criteria and exclusion criteria highlighted above.

2.1.3 Sensor Selection and Placement

The study will utilize commercially available ambient sensors to anomalies in behaviour . Each sensor type has been carefully chosen based on its ability to provide data relevant to specific activities and anomalies based both off those activities and just general behaviour.

- **Motion and Luminance Sensors (MLS):** Detect movement within specific rooms or zones and measure ambient light levels. These sensors are essential for tracking presence, transitions between rooms, and activities requiring specific lighting conditions.
- **Door/Cabinet/Drawer Open/Close Sensors (DO):** Monitor the open/close status of doors, cabinets, and drawers. These sensors can infer activities such as cooking (accessing cabinets) or cleaning (frequent drawer usage).
- **Electric Outlet Monitors (EOM):** Track power consumption of appliances such as microwaves, televisions, or hair dryers. This data is critical for identifying appliance use, providing insight into activities like watching TV, cooking, or grooming.
- **Environmental Sensors (ES):** Measure temperature and humidity levels. These sensors help identify activities like cooking (elevated temperature) or showering (high humidity spikes).
- **Vibration Sensors (VS):** Detect vibrations on surfaces or furniture. These sensors can identify activities such as sitting (vibrations on a chair) or operating household appliances like washing machines.

The sensor network in this study is carefully designed to capture diverse and context-rich inputs, including motion, environmental changes, cabinet open/close events, and electrical usage. These inputs are strategically chosen and placed based on domain knowledge to ensure the data is sufficiently informative for detecting presence meaningful patterns.

For instance, in a kitchen setting, motion detection (MLS), increases in temperature and humidity (ES), frequent cabinet or drawer openings (DO), and spikes in electrical usage from a stove (EOM) combined with the time stamp feature could correlate with cooking. While this setup reflects logical relationships, it avoids embedding explicit rules; instead, the unsupervised model will identify patterns and clusters autonomously. Below is a list of the various different types of sensors placed in each part of the apartment:

- **Living Room Sensors: MLS, VS, EOM** Purpose: Detect activities like sitting, watching TV, and presence in the living room. Motion and luminance inputs provide context for activity presence, while vibration sensors monitor furniture use, and EOM tracks appliance usage.
- **Kitchen Sensors: DO, EOM, MLS, ES** Purpose: Identify activities such as cooking, which appliances and drawers she uses and when for understanding routine, and presence in the kitchen. Door sensors monitor cabinet/drawer usage, EOM tracks appliance use, motion and luminance sensors detect presence and lighting conditions, and ES captures environmental changes indicative of cooking.
- **Bathroom Sensors: DO, MLS, ES, EOM** Purpose: Monitor activities like showering and presence in the bathroom. Door sensors track access, motion and luminance inputs detect presence and light changes, ES captures temperature and humidity variations during showering, and EOM identifies appliance use like hairdryers.
- **Bedroom Sensors: MLS, DO, ES, EOM** Purpose: Mainly to track presence in the bedroom and specific actions like when they access the dresser, outlets for understanding routine. Motion and luminance inputs monitor room activity and light changes, door sensors track closet access, ES records environmental changes, and EOM detects device usage.

This approach is designed to leverage inductive reasoning as a foundation, enabling the detection of broad patterns, behaviours, and anomalies.³⁹ The following examples illustrate the logic behind how sensor inputs can be used to infer activity which will then be confirmed by the participant, in the initial stage of the analysis.

- **Cooking:** A combination of motion in the kitchen (MLS), environmental changes (ES), cabinet and drawer access (DO), and increased electrical usage (EOM) may align with a pattern resembling cooking-related behaviours.

- **Showering:** Spikes in humidity and temperature (ES), along with motion detection in the bathroom (MLS), could suggest showering-like patterns.
- **Sitting on Couch:** A combination of the vibration sensor under the couch plus the motion sensors around the couch.
- **Watching TV:** The electrical output sensor of the TV plus a combination of motion and vibration sensors in the living room.

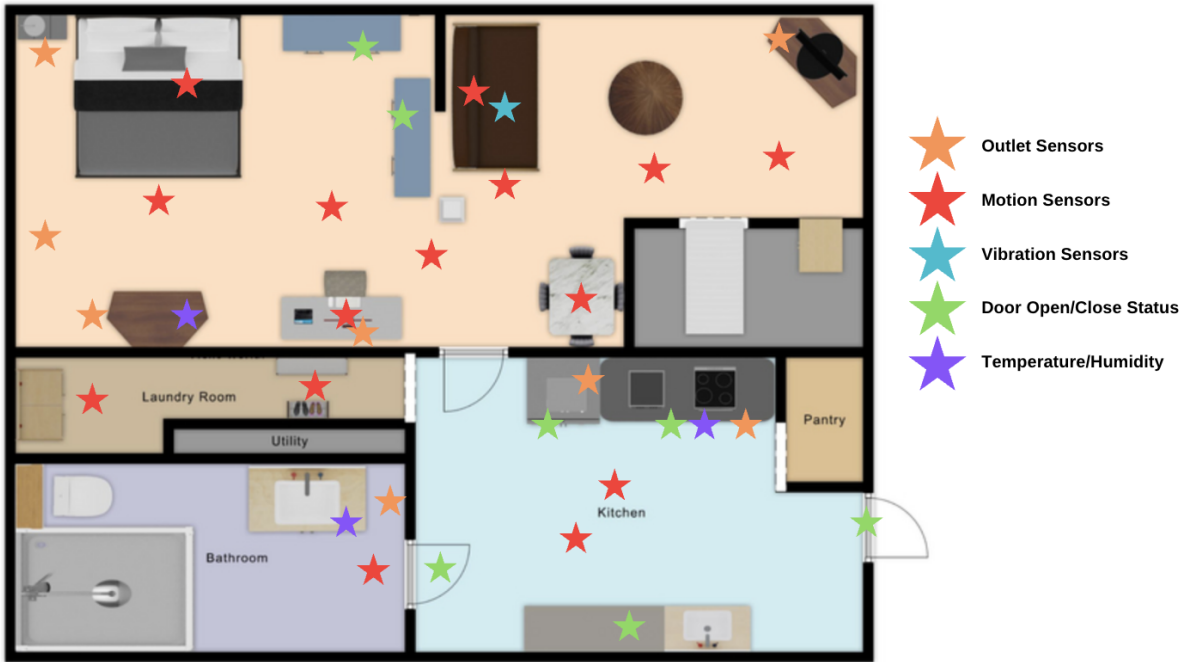


Figure 2.1: Sensor deployment layout in the smart home study environment. This schematic illustrates the spatial distribution and type of sensors installed within the mock apartment used for the study. Each icon represents a unique sensor modality: outlet sensors (red squares), motion sensors (red stars), vibration sensors (cyan circles), door open/close status sensors (green squares), and temperature/humidity sensors (blue circles). The floor plan serves as a reference for understanding spatial correlations between sensor placement and observed activities of daily living.

2.2 Data Collection Period

The sensor data is gathered continuously over a four-month period to ensure sufficient temporal coverage and variability in activities. Such a duration captures typical daily routines, occasional special events, and potential irregularities in participant behaviour.

2.3 System Configuration and Data Acquisition

2.3.1 Raspberry Pi Setup

A Raspberry Pi (Model 4B, 4GB RAM) was configured as the central hub for sensor acquisition and local data processing.⁴⁰ The Raspberry Pi OS was installed onto a microSD card, enabling built-in libraries necessary for automation and communication.⁴⁰ Secure Shell (SSH) access and a static IP address were configured to ensure reliable remote management and monitoring of the platform.⁴⁰ All system packages and firmware were updated to the latest stable releases to maintain security and software compatibility.⁴⁰

2.3.2 Domoticz Installation and Sensor Pairing

Domoticz was then deployed, an open-source home automation system, onto the Raspberry Pi to manage and unify data from multiple commercially available sensors.⁴¹ *Domoticz* was configured to automatically start on system boot, and its web-based dashboard was used for sensor pairing and ongoing data visualization.⁴¹

- **Z-Wave and Zigbee Integration:** USB dongles (Z-Wave and Zigbee) were attached to the Pi to manage wireless sensors for motion, door contacts, illuminance, temperature, humidity, etc.⁴²
- **Sensor Registration:** Each sensor was assigned a unique identifier (IDX), which *Domoticz* used to map sensor readings to semantic labels (e.g., “Kitchen motion,” “Bathroom humidity”).

This physical installation and device management approach follows established practice in smart-home and ambient-assisted living research, balancing local data processing with cost and privacy considerations.

2.3.3 Time-Series Data Pipeline

Following sensor registration in Domoticz, raw measurements (e.g., motion triggers, contact openings, humidity percentages) were continuously logged locally in *InfluxDB*, a time-series database optimized for high-frequency sensor data storage.⁴³ Data remains on the Raspberry Pi for privacy and local analytics. Periodic exports of the time-series data to CSV/Parquet formed the basis for subsequent machine learning pipelines.

2.3.4 Secure Upload to UWaterloo HealthCloud

Once all the data has been stored in InfluxDB a python script was constructed that automatically sends that data from the Raspberry Pi to a secure University of Waterloo server. Continuous pipeline monitoring ensures consistency and enables early detection of technical issues (e.g., sensor dropout). Additionally, anonymization protocols safeguard participant privacy and ensure ethical compliance.

2.4 Data Preprocessing and Cohesive Time-Series Construction

2.4.1 File Organization and Chronological Merging

Sensor readings exported from InfluxDB were split across multiple CSV files, each corresponding to a specific day or range of days. To consolidate these files into a single continuous time series, I began with a clearly ordered list of filenames (for instance, `12.02.2024.csv`, `12.03.2024.csv`). This ordering enforced a chronological sequence, ensuring that no events were processed out of temporal order.

After ignoring the first three header lines as they did not include data that was relevant, each CSV file was read into a `pandas` `DataFrame`. Only columns identified in a predefined *measurements* list, so sensor names relevant to door contacts, motion triggers, temperature, humidity, and power consumption were preserved. This filtering reduced extraneous variables and kept the dataset focused on the signals most pertinent to occupant activities.

Following parsing, all timestamps (originally in UTC) were converted into local time. By handling date ranges carefully, each CSV could capture data for either a single day or multiple days, which were then pivoted into a day-level wide format, rather than retaining the original long format.

To ensure a uniform temporal axis, the data was re-indexed so that all data is at a one-second frequency and carried forward the last known sensor reading whenever new data were absent. Ultimately, each daily subset was appended into a global DataFrame, making sure that the final row of one day would seamlessly link with the first row of the next. The resulting merged dataset was stored as a **Parquet** file, offering efficient on-disk access for large-scale processing.

2.4.2 Second-Level Time Index and Consistent Filling

By enforcing a one-second grid across all days, the pre-processing code achieved both temporal consistency and robust handling of missing data. In particular, door-contact sensors only transmit a reading when their status actually changes (e.g., a door-contact sensor toggling from closed to open), leaving idle periods with no new messages. So in this study, the last known value is filled until the sensor sends another a different valued input. The forward-fill step ensures that each sensor’s most recent known value is preserved until a subsequent update arrives, thereby preventing **NaN** gaps and capturing the correct status for each second. This carefully aligned time-indexed data frame forms the basis for all further analysis, such as occupant presence detection or activity classification.

2.5 Feature Engineering: Sensor Groupings and Semantic Labels

2.5.1 Door Grouping and Derived Variables

Because a single home may contain numerous door-contact sensors with overlapping or redundant coverage, the logic consolidated them into higher-level indicators. For example, all sensors labelled **DOOR-5-(contact)** and **DOOR-6-(contact)** (each tracking different sections of the fridge) were aggregated into a single variable **fridge_open**, clarifying whether the fridge was open at any time. Analogously, multiple kitchen cabinet sensors were combined under **kitchen_cabinets_open**, and bedroom dresser sensors were united in **bedroom_dresser_open**.

2.5.2 Column Renaming

To further enhance interpretability, each sensor’s name was mapped to a more descriptive label reflecting both its location and purpose. For instance, a raw name like `DOOR-8-(contact)` became *Main Door_Open*, signposting that this sensor corresponded to the home’s main door. Similarly, `OUT-5-Ampere:-power-consumption` was translated into *Kitchen: Microwave - OUT-5-Ampere:-power-consumption*, clarifying that it measured microwave electricity usage in the kitchen. A custom rename dictionary (`appliance_map`) handled these mappings programmatically, applying them after the door-grouping logic so that newly created columns, such as `kitchen_cabinets_open`, remained easy to identify. As a result, the final DataFrame offered human-readable column names that quickly directed attention to the associated room and device type, a crucial advantage for subsequent occupant-behaviour inference.

2.6 Occupant Presence, Multi-Resident Intervals, and Room-Level Feature Annotation

After pivoting and renaming the sensor dataset into `combined_df`, the following pipeline applies occupant-presence logic, locates the occupant in various rooms, flags multi-resident intervals, and annotates each interval with humidity and illuminance metrics.

2.6.1 Sorting by Time and Computing `row_dt_s`

I begin by sorting `combined_df` chronologically via the time column and resetting its index. For each row, I compute `row_dt_s`, the gap (in seconds) from the previous row, so that large discontinuities (≥ 2 s) can trigger logic resets (e.g., occupant location unknown). This step is mainly in preparation calculate accurate door cycles as will be shown in the next section.

2.6.2 Main-Door Cycles and Occupant Presence

Since the apartment has only one entry, occupant presence is derived from $0 \rightarrow 1 \rightarrow 0$ transitions of the “*Main Door_Open*” sensor. Whenever the logic detects the door changing from closed to open in the data ($0 \rightarrow 1$) and back to closed ($1 \rightarrow 0$), it records a “door

cycle” (start *vs.* end). A separate table, `cycles_df`, holds these intervals. To see if the occupant remains inside after a door cycle ends, the code automatically check for any interior motion sensor firing 0.4 minutes (24s) after the door closes. This buffer accounts for the fact that the motion sensors used do not change back to 0 after triggered for 20 seconds; hence, this buffer ensures short delays do not cause occupant presence in the apartment to be assumed true. If internal motion is found within this defined window (starting after the buffer and extending to the next door opening), occupant presence for that entire period (from the current door cycle’s start until the next one’s start) is set to 1; otherwise, it’s set to 0. To note, each and every major period of time that the occupant was away and in the apartment was confirmed by the participant’s logs for the first month to ensure reliability.

2.6.3 Incorporating Occupant Presence

Once the code compiles a list of door cycles and determines presence for each one, a helper function updates every row in the data frame with the final `occupant_presence` column equal to 1 or 0, filling the sensor DataFrame thoroughly. At this point, `combined_df` explicitly marks any second in which the occupant is thought to be in the apartment.

2.6.4 Generation of Binary ON/OFF Features from Illuminance and Outlet-Current Data

2.6.4.1 Illuminance Data

For each sensor channel, binary ON/OFF features were generated through a data driven clustering approach. Illuminance channels were first analyzed to classify ambient light into distinct ”ON” and ”OFF” states. To ensure computational efficiency, each illuminance sensor series was randomly sub-sampled to 70,000 observations. Preliminary experiments indicated that this sample size provided robust and stable clustering results, as demonstrated by silhouette scores consistently exceeding 0.70; further increasing the sample size yielded negligible improvements in clustering quality (silhouette increases below 0.02) while incurring quadratic growth in computational load.^{44;45}

Two-cluster partitioning was then conducted on these sub-samples using the *k*-means clustering algorithm with Euclidean distance, Lloyd optimization, and a fixed random state to ensure reproducibility. Clustering was executed in parallel across four CPU cores. The resulting clusters represented the two expected brightness modes—ambient lights ”OFF”

and "ON", as validated by silhouette scores ranging from 0.864 to 0.994, confirming strong cluster separation.⁴⁵ The cluster centroids (low and high illuminance values) were then used to define a midpoint threshold, classifying observations above this midpoint as "ON" and those below as "OFF". These binary illuminance states were integrated into the main dataset for subsequent analyses.

2.6.4.2 Outlet Data

Outlet-current data from smart plugs underwent a similar, but adapted, methodology to derive appliance usage states. Because amperage data exhibited greater variability and dynamic range than illuminance, a larger random sub-sample of 400,000 observations per channel was used. Clustering trials confirmed that this size also consistently produced silhouette scores above 0.70, again demonstrating minimal additional benefit from larger samples relative to computational cost.^{44;45} Before clustering, each outlet-current sample was standardized using z -score normalization (mean zero, standard deviation one) to eliminate bias from differences in magnitude across appliances.

Two-cluster k -means clustering was applied to the standardized current data, again using Euclidean distance, Lloyd optimization, and a fixed random seed for reproducibility. Parallel processing across four cores expedited computational efficiency. Similar to the illuminance channels, cluster quality was assessed using the silhouette coefficient, which revealed high-quality clustering with scores from 0.9349 to 1.000 across all appliance channels.⁴⁵ The resulting unscaled cluster centroids defined clear ON/OFF amperage levels. The midpoint of these centroids established a binary threshold, classifying observations above the midpoint as appliance "ON" states and those at or below as "OFF" states.

To reduce false positives from transient power spikes or measurement noise, an additional five-second run-length smoothing filter was applied: only appliance "ON" labels persisting for at least five consecutive seconds were retained, and shorter intervals were reverted to "OFF."

2.6.5 Occupant Presence and Room-Level Location Inference Using Multi-Sensor Fusion

Following the generation of binary ON/OFF features for ambient illuminance and appliance usage, occupant presence and precise room-level location inference were derived using sensor fusion logic.⁴⁶ This logic integrated multiple sensor types, including motion, outlet

activation, luminance transitions, door or cabinet openings, and specific furniture vibration events to accurately determine occupant location within the home.⁴⁶

Firstly, couch occupancy detection was performed to infer seated presence in the living room explicitly. Couch occupancy cycles were identified based on concurrent occupant presence and couch vibration sensor activations. A cycle was initiated when both occupant presence and couch vibration states were simultaneously positive (occupant present and couch vibration detected). To prevent false positives due to transient vibrations or sensor noise, the cycle was only considered valid if it persisted for at least 30 seconds. A couch occupancy cycle ended either when the occupant presence sensor indicated absence or when couch vibrations ceased in combination with external motion detected elsewhere in the apartment. External motion was defined as activity captured by motion sensors outside the living room. The start and end timestamps of these couch occupancy cycles were recorded, and periods within these intervals were labelled as "sitting_on_couch." This logic was corroborated and confirmed to all be correct by the participant confirming all couch sitting events that were far exceeded sitting for over 1 hour or longer over the data collection period.

Subsequently, the enhanced occupant-presence inference algorithm combined information from multiple sensors according to a prioritized set of logical rules to determine the occupant's current location at every timestamp:

- **Rule 1 – Explicit absence detection**

If the general `occupant_presence` sensor reported the occupant as not home (0), the location was set to *away*. No further rules were evaluated until presence (1) was re-established.

- **Rule 2 – Couch-occupancy priority**

When presence was confirmed, detected couch-sitting (via vibration cycles) overrode all other data: the occupant's location was immediately set to the *living room* because direct physical contact yields the highest certainty.

- **Rule 3 – New motion activation**

If couch occupancy was not detected, any sensor transition from inactive to active ($0 \rightarrow 1$) in a motion sensor defined the occupant's current room. Motion events were treated as real-time indicators of movement.

- **Rule 4 – Recent motion persistence**

In the absence of new motion, the algorithm retained the last room in which motion

had been sensed within the previous 30 min (1 800 s), assuming the occupant was unlikely to relocate without triggering motion again.

- **Rule 5 – Other sensor activations after motion lapse**

When motion data were stale for ≥ 30 min, the first new activation of any *outlet-usage*, *luminance*, or *door/open* sensor (inactive $\rightarrow$ active) reassigned location to that room, reflecting renewed activity despite prolonged motion silence.

If none of the above rules provided new location information, the algorithm defaulted to maintaining the occupant’s previously inferred location, ensuring continuity and avoiding unnecessary fluctuations in occupant location states. To avoid logical contradictions, an additional safeguard was implemented: if occupant presence was indicated but the current inferred location remained as "away," the location was corrected to "unknown" to reflect uncertainty rather than an explicit absence.

Each timestamp’s final inferred occupant location was recorded in the variable "occupant_location_enhanced," and boolean indicators were generated for ease of analysis. This comprehensive and prioritized logic provided robust, accurate occupant presence and room-level occupancy inference, essential for subsequent behavioural modelling, anomaly detection, and personalized AAL system interventions.

The resulting enhanced occupant location data, alongside previously generated binary ON/OFF variables, was finalized, sorted chronologically, and saved into a consolidated data file for subsequent analysis.

2.6.6 Data Cleaning, Multi-Presence Filtering, and Occupancy Session Identification

To handle scenarios where guests are present, the participant was asked to report approximate intervals of multi-resident presence (e.g., from 7:00 p.m. to 9:00 p.m.), without necessarily exact door event times. The logic then initially sets `multiple_presence=1` for all rows in those approximate intervals. Then, in order to refine them, the code looks for door cycles that occur fully within (or overlap) each approximate multi-resident block. The logic then "snaps" the multi-resident block’s start/end to match the earliest opening and the latest closing of the door found in that block, plus a brief 24s buffer. As mentioned, the 24 second buffer is to account for the motion sensors not being of 1 second frequency, so there is a delay in changing statuses by approximately 20 seconds as measured in the lab. If door events occur fully within that multi-resident time, the code extends or shortens the

block to match them. This final pass ensures `multiple_presence` is consistent with the occupant’s real comings and goings, further confirmed by participant feedback. Then rows flagged as having multiple people, `multiple_presence=1`, are removed from the dataset, to isolate single-occupant data. In the next dataframe, occupant presence sessions are identified by `presence_id`, incremented each time occupant_presence transitions from 0 to 1 or whenever a time gap (≥ 2 s) occurs, which is always when there are multiple people in the residence. I also track `occupant_time_in_interval` to measure how long a particular single-occupant session lasts in seconds.

This approach for tracking when the participant has other people in their apartment spares the participant from recalling precise times, which can be quite cumbersome especially when dealing with a dataset at 1 second frequency, meaning they have to be accurate at the seconds level. This methodology also reduces any errors due to errors in participant logging and exploits the fact that door cycles are highly logical events. To ensure accuracy the findings are then corroborated by the participant’s confirmation. In addition, since only one door exists in the apartment, occupant entries/exits are unambiguous, but this method could extend to multiple entry points if necessary. After refining the data, I subsequently exclude the rows that fall within the range of where the person had other people over and re-index sequentially, resulting in a new, clean and consistent dataset that purely contains single-occupant data.

Throughout this code, there are periodic checks to ensure no empty columns or rows remain. The code also produces periodic occupant presence summaries, listing start/end times of each occupant_presence block and verifying durations with the participant to ensure correctness. This offers a final consistency check before advanced modelling. Because row removal can introduce new time gaps, I recalculate `row_dt_s` to maintain internal consistency. This ensures that subsequent location or presence logic will correctly detect abrupt jumps because only the rows that follow rows that were cut because that time was removed have a `row_dt_s` greater than 1. I use this logic to reset any duration metrics at this row. So if, for example, the participant was in the kitchen before the time jump and was in the kitchen after the time jump it doesn’t calculate the duration as if they were in the kitchen the whole time including the time jump.

Two critical occupancy-related variables were subsequently created:

- ***presence_id***: A numeric identifier incremented with each distinct occupancy session, uniquely labelling uninterrupted intervals of continuous occupant presence. A new session began either upon transitions from absence to presence or when gaps between data points exceeded two seconds, indicative of potential brief monitoring interruptions.

- ***occupant_time_in_interval***: Within each identified occupancy session, this variable cumulatively recorded the occupant’s duration present, incrementing with each successive timestamp to accurately reflect continuous presence.

2.6.7 Unsupervised Bathroom Modelling

Intervals representing episodes of occupant presence within the bathroom were first extracted through the aggregation of contiguous temporal windows where binary presence indicator (`in_bathroom`) was 1, indicating the participant was in the bathroom. This aggregation was performed by iterating through each observation sequentially, identifying start and end timestamps where the binary indicator transitioned from inactive (0) to active (1) and vice versa. Each contiguous interval was thus defined by a pair of timestamps, (`interval_start`, `interval_end`), along with corresponding indices within the original data frame for detailed subsetting.

The intervals representing bathroom occupancy were systematically defined as:

- **`interval_start`**: the timestamp when occupancy transitioned from absence to presence.
- **`interval_end`**: the timestamp when occupancy transitioned from presence back to absence.

From each qualifying interval, domain-specific features were engineered to characterize showering events comprehensively:

2.6.7.1 Humidity-Based Features

Humidity Difference (`bathroom_hum_diff_end`): Calculated as the difference between the humidity at the end of the interval and the initial humidity at the interval onset. This captures net humidity changes potentially indicative of moisture accumulation due to showering.

Peak Humidity Difference (`bathroom_peak_diff_end`): Determined by identifying the maximum humidity recorded within the interval and subtracting the initial humidity. This feature quantifies the highest observed humidity rise, capturing significant transient moisture events.

Maximum Humidity Slope (bathroom_max_slope_end): Derived by calculating the maximal rate of humidity increase per unit time within each interval, determined through sequential humidity observations. This metric captures the rate at which humidity accumulates, distinguishing rapid moisture producing activities such as showering from other slower changing humidity sources.

2.6.7.2 Environmental Average Features

Within each identified bathroom occupancy interval, additional environmental parameters were averaged over the duration of presence:

Average Temperature (temperature_mean): Computed by averaging temperature sensor readings (Bathroom: VIN-3-Temperature), providing a baseline thermal environment assessment during occupancy.

Average Humidity (humidity_mean): Derived by averaging all humidity observations (Bathroom: VIN-3-Humidity) during the interval, reflecting overall moisture levels.

Average Illuminance (illuminance_mean): Calculated from the mean of adjusted illuminance sensor data (Bathroom: VAL-16-ILLUMINANCE-adjusted), indicating general light conditions within the bathroom.

2.6.7.3 Duration and Temporal Features

Interval Duration (duration_sec): Measured directly from the interval start and end timestamps, indicating total occupancy time within the bathroom, crucial for distinguishing between brief and extended occupancy periods.

Hour-of-Day Encoding (hour_of_day, hour_sin, hour_cos): Extracted from the interval's end timestamp, this cyclic time representation was implemented using sine and cosine transformations, calculated as follows:

$$\text{hour\_sin} = \sin\left(2\pi\frac{\text{hour\_of\_day}}{24}\right), \quad \text{hour\_cos} = \cos\left(2\pi\frac{\text{hour\_of\_day}}{24}\right) \quad (2.1)$$

2.6.7.4 Unsupervised Shower Model

Showering detection was attempted using unsupervised modelling to see even in instances where labelling of showering events is unfeasible, showering can still confidently be detected. The K-Means algorithm was selected for this analysis.⁴⁷ The values for the overall

and peak humidity rise features were standardized to have zero mean and unit variance using standard Z-score scaling.⁴⁷ The K-Means clustering algorithm was then fitted to this standardized two-dimensional feature set.⁴⁷ Parameters for the algorithm included setting the number of clusters to two, using a fixed random seed for reproducible centroid initialization, and allowing for multiple initializations from which the best result based on inertia was chosen.⁴⁷ The output consisted of a cluster assignment label (0 or 1) for each input data interval. The internal validity of the resulting partition was assessed using the Silhouette score, which measures how similar an object is to its own cluster compared to other clusters. External validity was through comparing these two clusters to see if the model was able to distinguish between the correct showering and non-showering events as assessed by precision, recall, F-1 score, and reduced false positives and false negatives.

Predictions were also systematically analyzed to identify ambiguous classifications and potential labelling errors. Specific probability thresholds were defined: intervals with predicted shower probabilities above 0.70 were classified confidently as showers, while those below 0.30 were confidently non-showers. Intervals outside these confidence bounds were flagged for human expert review, so those above 0.20 and below 0.80. The reasoning to implement human expert review is particularly for capability to scale the model to more homes in the future and to reduce hallucination. It is based on the assumption that if the model gets it wrong they they are most likely unsure about it; hence, falling within the category that needs an individual to either confirm or deny the showering event. This ensure accurate data whenever the model is unsure, improving the model’s robustness.

To support targeted review, detailed SHAP-based explanations were provided for flagged cases, highlighting the top three influential features affecting the model’s predictive decision.⁴⁸ This is especially important because it directly addresses modelling with a priority for explainability especially when the model is not sure. The intervals the model is not sure about is sent to an output file. The idea is that these output file, in future work, allows for the participant to provide feedback by confirming or denying the showering event and as a result the model takes it in as a supervised label to continue to improve it.

2.6.7.5 Unsupervised Bathroom Behavioural Anomaly Model

The supervised random-forest classifier described in the preceding section was optimized to label individual bathroom intervals as *shower* or *non-shower*. The present section details the second-stage unsupervised model that searched for *behaviourally unusual* intervals within those confidently labelled classes. As a result of only considering those that are confidently labelled, led, $0.20 < P < 0.80$ were excluded from unsupervised training because mixed labels would confound the density-based detector. The resulting strata contained

3,864 confident showers and 11,217 confident non-showers. Within each stratum numeric features were median-imputed to replace the rare missing or infinite values ($< 0.1\%$ of entries) and then z-standardized using the stratum-specific mean and standard deviation. Medians were preferred because they are resistant to the very outliers the model seeks to identify.

Isolation Forest (IF) configuration A separate IF was fit to each stratum using the scikit-learn implementation.^{47;49} Key hyper-parameters were

- 100 trees (`n_estimators`)
- subsampling all observations (`max_samples = auto`)
- random feature subset of at most 75% of the available columns at each split (`max_features = 0.75`) to reduce tree correlation and improve generalization
- `contamination = 0.01`, reflecting the objective to detect the biggest anomalies to reduce any potential anomalies that were noise due to a potentially small longitudinal sample size.
- `random_state = 42` to guarantee reproducibility

The IF was trained on the full historical record for each stratum, thereby learning the participant-specific distribution of “normal” shower and non-shower behaviour without relying on external labels.

Model Explanation Logic The following model explanation is consistent with all anomaly models constructed for all rooms including the bedroom, kitchen, living room, and bathroom. It will be mentioned in detail in this section and referred to in subsequent sections. To render the unsupervised flags interpretable to domain experts, a two-tier explanation pipeline was attached:

1. **Local feature attribution.** `TreeExplainer` from the SHAP library was applied to the trained IF to obtain per-feature contributions for every interval.⁴⁸ The three features with the largest absolute SHAP values were extracted; their signed values indicate whether each feature pushes the sample toward or away from normality.⁴⁸

2. **Dynamic z -score context.** For each of the top three features, the raw value of the anomaly was compared with all confident intervals that began within ± 1 h of the anomaly (minimum of five comparison intervals). A z -score was calculated from the local mean and standard deviation. Values with $|z| > 1.5$ were annotated as “high” or “low” to place the anomaly in its circadian context.

The resulting explanation string therefore links the internal mechanics of the IF (SHAP) to an intuitive statistical framing, thereby facilitating clinician assessment of whether the event is medically relevant (for example an abnormally long but low-humidity “shower” suggesting prolonged toilet use rather than hygiene).⁴⁸ For example: *duration = 3120 s* (*SHAP +0.63, high, $z = +3.1$*); *bathroom_hum_diff = 1.2 %* (*SHAP -0.28, low, $z = -1.9$*); *hour_sin = 0.97* (*SHAP +0.18, high, $z = +2.0$*).

2.6.8 Unsupervised Living Room Modelling

2.6.8.1 Extraction of Living Room Activity Features

The next methodological step involved extracting detailed features representing living room occupancy and activity intervals based on ambient sensor data. The goal was to structure continuous intervals of living room occupancy into meaningful blocks and compute associated activity metrics to characterize typical living room usage patterns.

Data Preparation and Interval Identification The initial dataset underwent pre-processing to establish accurate intervals of living room occupancy. Data entries were first chronologically sorted and standardized to ensure temporal continuity. Occupancy was identified using an occupancy indicator specifically derived from sensor data signaling presence within the living room.

Consecutive data points indicating living room occupancy were grouped into distinct intervals using changes in occupancy status. Intervals were identified by marking transitions in occupancy status, enabling discrete block delineation. To address potential data resolution issues, the duration of each occupancy interval was computed by summing the temporal differences between consecutive timestamps within each block. Special attention was given to intervals with zero calculated duration, wherein the median temporal difference across the dataset was used as an imputed duration value to maintain consistency and accuracy.

Feature Engineering To enrich the interval data, several features were computed:

- **Temporal Features:** Cyclical representations of the interval start times were computed using sine and cosine transformations of the start hour, effectively capturing the periodic nature of daily activities.
- **Couch Occupancy Metrics:** The duration of couch occupancy was calculated by integrating sensor-derived indications of sitting activity. Specifically, timestamps labelled with couch occupancy were cumulatively summed within each living room occupancy interval, quantifying the total duration spent on the couch.

These engineered features provided nuanced insights into living room usage, differentiating between general occupancy and specific seated activities, thereby enhancing the descriptive capability of the dataset.

The final structured output dataset encompassed clearly defined living room occupancy intervals, their associated durations, cyclical temporal indicators, and specific couch occupancy metrics. Resulting features were rounded to appropriate precision levels, ensuring consistency and readability. Validation steps included checks for correct data types and completeness, ensuring reliability of subsequent analyses.

2.6.8.2 Unsupervised Living-Room Behavioural Anomaly Model

Every living-room visit identified by the location segmentation pipeline was summarized into four core attributes: total visit duration (s), couch-contact duration (s), and a two-dimensional cyclical encoding of clock time derived from the visit start hour. Where present, auxiliary sensor variables (for example mean ambient light, door-open counts or audio energy) were appended, but the core model relied on the four invariants above so that the approach generalized across apartments with heterogeneous sensor inventories.

Living-room behaviour is markedly different when the participant is seated on the couch compared with transient walk-throughs. To reduce within-class variance the data set was split on a binary flag, of whether the individual was sitting on the couch or not. Subsequent modelling was performed independently in the two strata.

Unsupervised detection of unexpected living room behaviours was performed with the same two-step IF pipeline that was described in the bathroom section above, so please refer to that section for specific details on the exact model and how explainability was derived from the model results. The model (i) accounts for the substantial behavioural differences

between visits that consist of substantial sitting on the couch and not by having two separate models for each condition and (ii) provides transparent, feature-level explanations.

This subgrouping, similar to non-showering and showering and cooking and not cooking, reduces within-class variance and yields more homogeneous distributions for the unsupervised models of the living room. The two independent IF models, one per subgroup, were then fit to the full historical dataset, exactly as they were for the kitchen and bathroom. To confirm, exactly like the bathroom and kitchen model, all numeric features were median imputed, standardized and trained on all historical visits.

2.6.9 Unsupervised Bedroom Modelling

2.6.9.1 Extraction of Bedroom Activity Features

Data Preparation and Interval Identification Raw sensor records were ordered by local timestamp and converted to a uniform `datetime` format. Bedroom presence was indicated by the binary flag `in_bedroom_enhanced`. Consecutive rows with this flag equal to 1 were aggregated into occupancy blocks by a single-pass state machine: the block start was recorded at the first row whose flag switched from 0 to 1, and the block end at the first subsequent row whose flag returned from 1 to 0. Blocks that extended to the final row of the file were closed retrospectively.

A pre-computed column `time_diff_seconds` (the inter-row time delta) was summed within each block to obtain visit duration. Zero-length blocks—produced when two consecutive timestamps were identical—were assigned the median positive inter-row delta to preserve temporal consistency.

Feature Engineering For every validated interval the following attributes were computed:

- **Duration** of the visit (seconds).
- **Light-usage indicators:**
 1. a binary flag `lights_on` that equals 1 if any ceiling or bedside fixture was ever switched on during the interval;
 2. the cumulative `light_on_duration_seconds`, obtained by multiplying the unified “any-light-on” series by `time_diff_seconds`.

- **Dresser activity:** the count of distinct dresser-drawer openings, detected as rising edges in `bedroom_dresser_open`.
- **Circadian context:** start-hour sine and cosine (`hour_sin`, `hour_cos`) together with the raw hour value (`hour_of_day`) for later dynamic baseline comparisons.

All numeric fields were rounded to three decimal places for readability and exported both as Parquet and CSV to support downstream modelling.

2.6.9.2 Bedroom Behavioural Anomaly Model

Feature Matrix The unsupervised detector used six invariant features:

- `duration_sec`
- `lights_on`
- `light_on_duration_seconds`
- `dresser_opens_count`
- `hour_sin`
- `hour_cos`

These measures together describe time-in-room, light-use intensity, interaction with furniture and circadian timing—factors. The complete dataset was then z-standardized with `StandardScaler`.⁴⁷

Due to a lack of specific activity detection within the bedroom as was the case in all three aforementioned rooms, like cooking in the kitchen, all bedroom data was processed through a single-step IF pipeline with the same properties as was described in the bathroom section above (please refer to that section for specific details). To confirm, exactly like aforementioned room models, all numeric features were median-imputed, standardized and trained on all historical visits.

2.6.10 Minimally Supervised Showering Model

The reason why showering was also explored via a minimally supervised model is because unlike the other activities like cooking and sitting on the couch as well as presence in each of the rooms as well as presence in the apartment, the unsupervised showering model approach may not generalize well to other homes. This is because humidity may fluctuate to a higher degree depending on the size of the bathroom. As a result, this section is about exploring whether a few weeks of data may be enough to detect showering for months on end and whether it is possible to identify whenever the model is unsure and prompt the user to confirm. This is to basically say that even if a completely a unsupervised showering model is not possible to scale to other homes, then perhaps a few weeks of supervised labelling instead of the alternative of just no showering activity at all or labelling at all times. This still addresses the overall objective of reducing exhaustive labelling by exploring how many weeks of labelling are at minimum required to achieve great showering modelling results.

2.6.10.1 Innovative Approach for Supervised Showering Labels

Participants were instructed to maintain weekly logs documenting the approximate timing of their shower events, as shown in appendix A. These logs included the specific days of shower occurrences, approximate time frames (e.g., "late evening", 10-11 pm), and estimated duration (e.g., "approximately 10 minutes"). These entries served to validate the unsupervised models and labels for the supervised model also tried.

For each participant-reported approximate shower event, I scrutinized corresponding sensor data segments, systematically identifying periods characterized by distinct humidity spikes, abrupt temperature elevations, and filtering any bathroom occupancy that is safely below the reported duration of the showering event on that day and in that time period reported. These precisely engineered features enabled the objective determination of exact event timings, directly corroborating the participants' sometimes approximate logs (i.e. "Showered around 5-7pm on Monday"). Participants were subsequently presented with these precise sensor-derived showering intervals to confirm or refine exact event timings. This iterative validation not only confirmed sensor accuracy but also served as a critical cross-verification of self-reported data, effectively mitigating potential participant-driven biases and errors.

Finally, the intervals were annotated based on whether showering occurred within the interval or not through a binary feature called, the showering labels (**showering**).

2.6.10.2 Supervised Shower Model

Two supervised classification algorithms were compared: Random Forest (RF) and Logistic Regression (LR).^{47;50;51} Both models utilized balanced class weighting to address class imbalance explicitly. The RF classifier was instantiated with 100 trees, parallelized execution, and reproducibility enforced by a fixed random state.⁴⁷ LR employed a similar strategy with extended iterations (1000 iterations) to guarantee convergence.⁵¹

Model validation employed a stratified 5-fold cross-validation approach, preserving original class proportions in each fold, and ensuring generalized performance estimation. The dataset was partitioned into various stratified training sets from 40% training split all that way 5% training split with the rest being test splits, using temporal split.⁴⁷ Temporal allows for the model to learn from the data in chronological order. So for example if 20% training in temporal split may mean the first month of data is used to train the dataset and the test will be on the next few months of data.⁴⁷

Chapter 3

Results

3.1 Dataset

The following is a summary of some of the things detected completely unsupervised in this study.

Table 3.1: Summary of detected activity instances between 2 December 2024 and 20 March 2025

Activity	Event type	Count	Percent
Bathroom stays	All stays	707	—
	Showers	58	8.0%
Kitchen events	All events	3002	—
	Cooking events	157	5.2%
Bedroom stays	All stays	3221	—
Living-room stays	All stays	2632	—
	Couch-sitting	727	27.6%
	TV sessions	9	0.3%

To note, this was over a data collection period of about 3 and a half months (December 2nd - March 20th); however, as described in the methods section some days and most weekends were removed as the participant had people over. About 6 showering events were incorrectly classified as non-showers using the unsupervised model. All sitting on the couch

and cooking events were loosely validated by having them confirm the anomalies derived from couch and cooking activities as well having them confirm the visualizations produced for what normal cooking and couch activity looks like. Couch and cooking they did not require strict validation as they were more so reliant on sensors that directly measured these activities through creative, non-traditional usage of commercially available sensors. Specifically the vibration sensor used to detect sitting on the couch directly detects couch activity by automatically activating whenever they were on the couch. In addition, cooking was mainly defined as when key appliances were on which were specifically detected using outlet sensors which is a direct measurement, and leaves little room for error.

3.2 Unsupervised Appliance On/Off Detection

Table 3.2 below specifically describes how successfully the K-means clustering modelling was able distinguish when the appliances connected to a specific outlet was on or off. This is clearly demonstrated by the very high silhouette coefficients. The "Center Low" columns demonstrate the average value of the low cluster, which shows when outlet is off and the reciprocal for "Center High".

The plausibility of the derived appliance states was systematically validated against periods of known apartment vacancy for five high-usage appliances (microwave, stove, rice cooker, oven, television). This is because the participant indicated that they never turn these appliances on or keep them on when away from the apartment. The outlet-current measurements during these unoccupied periods were compared to see whether the k-means clustering model would be able to automatically detect on or off without knowing any labels that match up. The results are shown in the table below in the "Agreement" column, exceeding 97% across all evaluated appliances. This methodology shows the strong validity of the k-means clustering methodology to identify on/off for outlets without any labels at all.

Table 3.2: K-means clustering of outlet-level current draw

Appliance / Outlet	Sil.	Center Low (A)	Center High (A)	Agreement (%)
Bathroom: OUT-14	0.9991	0.0127	6.7924	N/A
Bedroom: Lamp (OUT-1)	1.0000	0.0000	0.0200	N/A
Bedroom: Laptop Charger (OUT-4)	0.9195	0.0050	0.2741	N/A

Table 3.2 (continued)

Appliance / Outlet	Sil.	Center Low (A)	Center High (A)	Agreement (%)
Bedroom: Monitor (OUT-3)	0.9993	0.0000	0.1412	N/A
Bedroom: OUT-10	0.9977	0.0063	9.6811	N/A
Bedroom: OUT-11	0.9895	0.0007	9.7610	N/A
Bedroom: OUT-9	0.9349	0.0196	0.4881	N/A
Bedroom: Phone Charger (OUT-8)	0.9493	0.0044	0.2210	N/A
Kitchen: Electric Stove (OUT-6)	0.9971	0.0080	7.0631	99.87
Kitchen: Microwave (OUT-5)	0.9998	0.0103	11.4649	99.76
Kitchen: Oven (OUT-7)	0.9998	0.0009	13.0924	99.68
Kitchen: Rice Cooker (OUT-13)	0.9871	0.0269	2.6090	100.00
Living Room: TV (OUT-2)	0.9868	0.0173	0.6882	97.24

Two-cluster ($k = 2$) solutions were fitted to a random 400 000-sample subset of one-second amperage readings for each outlet. Silhouette coefficients above 0.7 confirm well-separated OFF/ON clusters, and the large gap between *Center Low* and *Center High* supports deriving binary ON/OFF indicators using a midpoint threshold. Where available, the *Agreement* column reports the percentage concordance between the K-means classification and known times; entries without such known times are marked “N/A.”

3.3 Machine Learning for Showering Detection

3.3.1 Minimally Supervised Model for Shower Detection

This section explores a minimally supervised approach for detecting showering events. Identifying showers with ambient sensors, such as humidity sensors, can be challenging because humidity patterns vary greatly between homes, influenced by bathroom size, ventilation, and usage habits. This variability brings up important questions about how effectively a detection model can generalize across different environments.

To address this potential challenge, the study evaluates a supervised learning method to see how it compares with more and more labels. The primary goal is to test whether minimal initial labelling can sustain accurate shower detection over an extended period. Specifically, the evaluation considers two practical questions:

Can a model trained with a small amount of initial labelled data (for example, just a few weeks) reliably detect showers over several months?

Is the model capable of indicating uncertainty, clearly identifying when it has low confidence in its predictions? This feature would help determine when additional verification or labelling might be necessary.

Assessing this minimally supervised strategy aligns with the study’s broader objective of reducing the need for ongoing manual labelling. This supervised approach serves as a practical alternative, particularly useful in scenarios where completely unsupervised models might face challenges in generalizing effectively across diverse home environments.

3.3.1.1 Performance of Models at Various Labelling Levels

To assess the impact of how a few weeks to about 6 weeks worth of labelling by the participant may influence the model’s results, two models were evaluated across different temporal train/test split ratios. The dataset was partitioned chronologically, allocating the earliest 5% to 40% of observations for training and the remaining contiguous block for testing. The date ranges for the 5% and 40% are indicated so viewers can get an idea of how many days were labelled and the impact of the amount of days labelled on the model performance. Performance metrics achieved on the respective test sets for each configuration are detailed in Table 3.3. Both models demonstrated high performance across the tested splits, with peak F1-scores for the ‘Shower’ class observed when using 30% of the data for training.

Table 3.3: Model performance on a temporal test set across six training–test splits.

Split configuration	Random Forest (RF)						Logistic Regression (LR)					
	Acc.	AUC	P_{sh}	$F1_{sh}$	FN	FP	Acc.	AUC	P_{sh}	$F1_{sh}$	FN	FP
Days of Labelled Data Used												
25 (40%)	0.998	1.000	0.973	0.986	0.000	1	1	0.999	0.947	0.973	0.000	2
20 days (30%)	0.998	1.000	0.977	0.989	0.000	1	1	1.000	0.935	0.966	0.000	3
14 Days (20%)	0.995	1.000	0.942	0.970	0.000	3	1	1.000	0.875	0.933	0.000	7
7 Days (10%)	0.995	0.999	0.950	0.970	0.000	3	1	0.997	0.850	0.910	2.000	9
3 Days (5%)	0.996	0.999	0.980	0.970	2.000	1	1	0.993	0.620	0.750	2.000	35

P_{sh} is the precision for the Shower class. Recall is not shown because it was 1.000 for every split and model. FN = false negatives (missed showers over the 3 1/2 months of data collection); FP = false positives (non-showers classified as showers over the 3 1/2 months of data collection). The percentage indicated in the first column indicates the exact amount of bathroom instances that were used to classify the data alongside the number of days those bathroom instances fell under. Please keep in mind the days are approximate calculations as some days had more bathroom instances than other as well as some days, only partial data was available as participant had other people over at their house the other hours of the day.

The best performing model from the table above is illustrated in the table below. It is the Random Forest Model that uses about 14 days of data from Dec. 2 - Jan. 7. The reason why it is only 14 days even though the date range is indicates longer is because the participant was away from the home from Dec. 20

to Jan. 1 as well as weekends from Dec. 2 to Dec. 20. As a result just using about 14 days of data between that date range, the model was able to achieve exceptional results. More importantly to note; however, even using just 3 days gives pretty high results, specifically for the random forest model. In addition, any uncertainties, which all the false negative and false positive instances fall under, are automatically detected and can be set up in future studies to be sent to the participant to confirm. This is if future studies want to eliminate any possibility of false classifications to ensure the subsequent anomaly models are 100% accurate. The logic is that only a few of the models the models will be unsure about rather than all of them, so in turn it saves time from manually labelling all instances.

Five intervals ($< 1\%$ of the dataset) met this criterion of the model being unsure about, for the best performing model that used 20 days of labelled data. One of the 5 is the one that was actually incorrectly classified as a shower when the participant originally labelled that as a non-shower event.

Table 3.4: Best Performing Model

Class	Precision	Recall	F_1	Specificity	Accuracy	ROC-AUC	Support
Non-Shower (0)	1.000	0.9978	0.9989	0.9978			451
Shower (1)	0.9773	1.000	0.9885	1.000			43
Overall					0.9980	1.0000	494

Random Forest Model at the 30/70 train/test split results

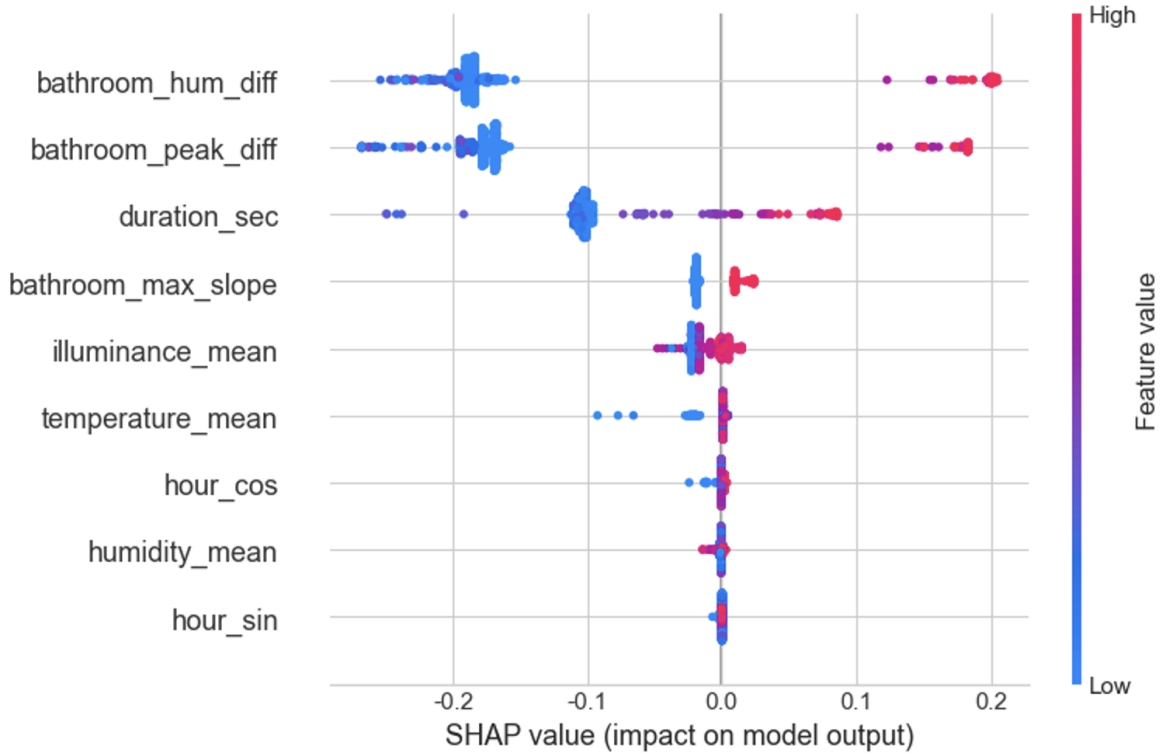
Table 3.5: Confusion matrix for Best Performing Supervised Machine Learning on the Test-Set

True label	Predicted label	
	Non-Shower	Shower
Non-Shower (0)	450	1
Shower (1)	0	43

3.3.1.2 Model Interpretability

Figure 3.1 presents a SHAP bee-swarm plot generated from the hold-out samples, confirming that rapid humidity rise (`bathroom_peak_diff` and `bathroom_hum_diff`) dominates the model's decision process, followed by event duration.

Figure 3.1: SHAP summary plot of the best performing shower-detection model.

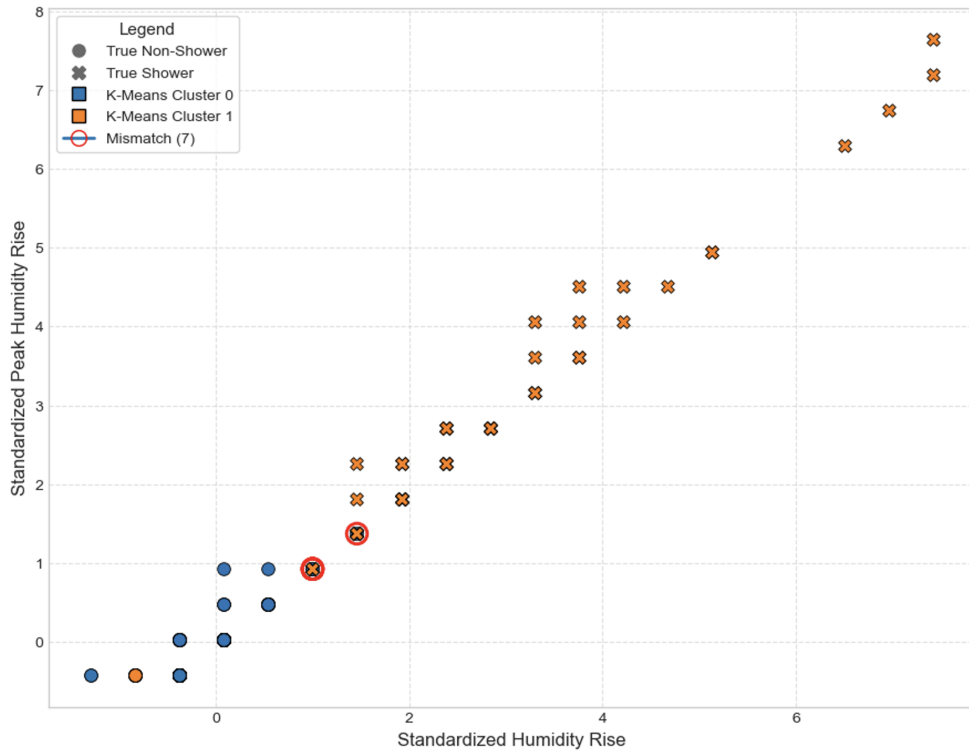


Each dot in this figure is a single bathroom interval; the x-position shows that feature’s contribution to the predicted probability of a shower (>0 shifts the decision toward “shower,” <0 toward “non-shower”). Colour encodes the raw feature value (blue = low, red = high). Large positive impacts from `bathroom_hum_diff`, `bathroom_peak_diff`, and `duration_sec` confirm that rapid and sustained humidity rises are the principal determinants of model output, while secondary features (e.g. `illuminance_mean`, circadian terms) play smaller, contextual roles.

3.3.2 Unsupervised Machine Learning Model for Shower Detection

This results section shows the best performing unsupervised machine learning model, a model that was not given any labelled data whatsoever. So this model was not given any labels at all on showering and non-showering events within the bathroom events across the 3 1/2 months of data collections. It purely takes all bathroom events without any labels and attempts to classify them into two clusters that represent showering and non-showering events.

Figure 3.2: k -means clustering of bathroom intervals ($k = 2$).



Points show the two standardized humidity-rise features used for clustering; colours denote the cluster assignments, while marker style (circle = true non-shower, cross = true shower) indicates the ground-truth label. Red-outlined markers highlight the seven intervals where the unsupervised cluster label disagrees with the ground truth. Please note that there are 7 mismatches but only 2 red circles are shown because they are all overlapping in the same region. (see Table 3.6). The silhouette score of 0.867 suggests a well-separated two-cluster solution.

Table 3.6: Confusion matrix comparing ground-truth labels with two-cluster k -means assignments.

True label	Predicted (Cluster label)	
	Non-Showering (Cluster 0)	Showering (Cluster 1)
True Non-Showering	642	1
True Showering	6	58

This clearly shows the 7 mismatches referenced in the figure above. 6 of them were incorrectly classified as non-showering events and 1 was incorrectly classified as showering. The number of mismatches for unsupervised is less than the model that uses labels for at least 3 days.

Table 3.7: Side-by-side comparison of the best-performing minimally labelled model and the no labelled model

Metric	20 days of labels (RF)	Unsupervised(k-means)
<i>Non-Shower (0)</i>		
Precision	1.0000	0.9907
Recall	0.9978	0.9984
F ₁ -score	0.9989	0.9946
Specificity	0.9978	0.9984
<i>Shower (1)</i>		
Precision	0.9773	0.9831
Recall	1.0000	0.9062
F ₁ -score	0.9885	0.9431
Specificity	1.0000	0.9984
Overall accuracy	0.9980	0.9901
ROC-AUC (overall)	1.0000	—
Support (total N)	494	707

This table demonstrates that the unsupervised model does not show a huge dip in model performance compared to the model that used 20 days of labels.

This section demonstrates that a completely unsupervised approach also performs really well in accurately detecting showering and non-showering events in the bathroom. Naturally, a completely unsupervised approach will result in more mismatches, 7, compared to even just 3 days of labels which resulted in only 3.

3.3.3 Comparative Analysis of Misclassifications and Achieving Full Accuracy

While both the minimally supervised and unsupervised approaches demonstrated strong performance in identifying shower events, a critical aspect for practical deployment is understanding and managing the inevitable misclassifications. This section directly compares the errors produced by the best-performing supervised model and the unsupervised model, highlighting a key shared capability: the identification of uncertain predictions to flag potential errors.

A crucial finding is that predictions can be flagged where the model exhibits low confidence or ambiguity to correctly identify mis-classifications. For the supervised RF model, a probability threshold was used to identify uncertainty. While k -means clustering does not inherently produce probabilities, metrics such as the distance of a point to its assigned cluster centroid versus the other centroid were effectively able to be served as an analogous measure of ambiguity. The misclassified points in Figure 3.2, particularly those near the decision boundary between the clusters, represent instances where such ambiguity is high.

Table 3.8: Comparison of Misclassifications and Uncertainty Flagging

Characteristic	Supervised (RF, 30% Train)	Unsupervised (k -means)
Dataset Scope	Test Set ($N = 494$)	Full Dataset ($N = 707$)
Misclassifications		
False Negatives (FN)	0	6
False Positives (FP)	1	1
Total Errors	1	7
Uncertainty Handling		
Mechanism	Prediction Probability	Cluster Assignment Ambiguity
Instances Flagged	5	7
Errors Captured	1 out of 1	7 out of 7

*Requires defining an ambiguity metric (e.g., distance-based) for k -means.

Note: Supervised results are on the held-out test set; Unsupervised on the full dataset.

Therefore, regardless of whether an initial supervised or unsupervised strategy is employed, leveraging model uncertainty offers a powerful mechanism to significantly reduce manual verification efforts. Instead of reviewing all detected bathroom events (over 700 in this study), attention can be focused solely on the

small subset of instances flagged as uncertain (e.g., the 5 flagged by the RF model). This targeted approach is key to achieving highly accurate activity logs efficiently, capturing most, if not all, misclassifications with minimal human intervention. Table 3.8 provides a direct side-by-side summary of these error characteristics and uncertainty handling aspects.

3.4 Anomaly Detection

This section tabulates the results of the various room-based unsupervised anomalies detected. The results only show the anomalies that were detected but fail to show all the various features that contributed to these anomalies. As such, certain features if there was a clear established pattern and no significant deviations from the normal routine, then that will not be reflected as one of the three most contributing features to that anomaly. Below of room-based anomaly tables, some figures were included to clearly illustrate the reasoning behind why some anomalies were detected. Not all anomalies from the table have clearly indicated graphs included as some of the anomalies are far more complex to illustrate via a 2 variable scatter plot. However, these relationships are discussed in detail in the discussion section.

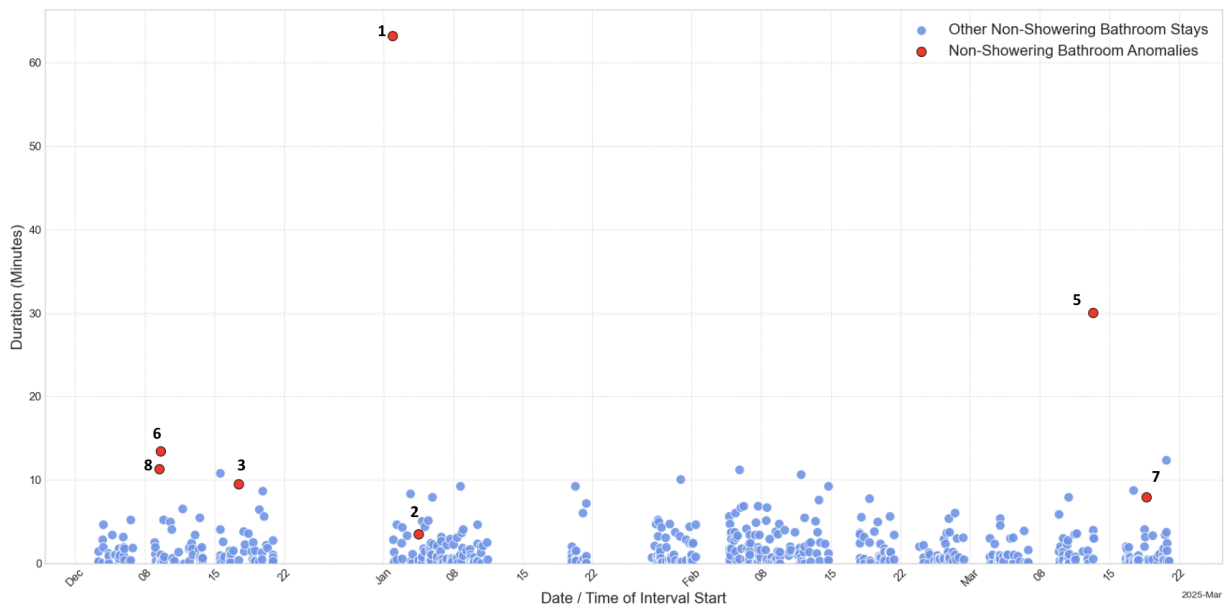
Table 3.9: Bathroom-Detected Anomalies with Participant Feedback

Interval	Anom. Score	Raw Anomaly Reason	Anomaly Explanation to Participant	Conf.	Imp.
1/1/2025 21:47–22:50 (1)	-0.0402	duration_sec = 3790 s (SHAP = -2.39 → Anom.; Z = 33.8); bathroom_hum.diff = -1.0 (SHAP = -1.03; Z = -1.9); illuminance_mean = 1.0 (SHAP = -0.81; Z = -2.2).	Bathroom activity lasted just over an hour without the humidity spike of a shower, making this very long non-shower stay clearly abnormal.	Yes	Yes
1/4/2025 12:14–12:18 (2)	-0.0335	bathroom_peak.diff = 3.0 (SHAP = -1.98; Z = 5.8); bathroom_hum.diff = 3.0 (SHAP = -1.14; Z = 5.2); bathroom_max_slope = 1.0 (SHAP = -0.71; Z = 1.8).	Humidity rose far higher than normal for a non-shower event at this time of day, so the system flagged the interval.	Yes	No
12/17/2024 09:30–09:40 (3)	-0.0235	duration_sec = 571 s (SHAP = -1.17; Z = 2.1); bathroom_peak.diff = 2.0 (SHAP = -1.08; Z = 2.7); bathroom_hum.diff = 2.0 (SHAP = -0.65; Z = 2.8).	A 9.5-minute stay plus a 2-unit humidity spike are both above typical non-shower levels for mid-morning.	Yes	Yes
1/21/2025 07:40–08:00 (4)	-0.0229	bathroom_hum.diff = 17.0 (SHAP = -1.05; Z = 2.5); bathroom_peak.diff = 18.0 (SHAP = -0.80; Z = 2.8); temperature_mean = 24.3 °C (SHAP = -0.28; Z = 2.0).	A very hot, humid shower: humidity rose 17–18 units and temperature hit 24.3 °C, both well above early-morning norms.	Yes	Yes
3/13/2025 08:16–08:46 (5)	-0.0168	duration_sec = 1802 s (SHAP = -2.06; Z = 15.3); bathroom_hum.diff = -1.0 (SHAP = -1.09; Z = -1.9); bathroom_max_slope = 1.0 (SHAP = -0.83; Z = 1.3).	A 30-minute non-shower stay, far longer than usual, with a slight humidity drop was flagged as abnormal.	Yes	Yes
12/9/2024 14:25–14:38 (6)	-0.0118	duration_sec = 809 s (SHAP = -1.60; Z = 6.2); bathroom_peak.diff = 2.0 (SHAP = -1.15; Z = 4.2); bathroom_hum.diff = 2.0 (SHAP = -0.75; Z = 5.3).	A 13.5-minute afternoon stay plus a 2-unit humidity rise are far above normal non-shower values.	Yes	Yes
3/18/2025 17:29–17:37 (7)	-0.0029	bathroom_peak.diff = 2.0 (SHAP = -1.13; Z = 3.5); duration_sec = 479 s (SHAP = -1.09; Z = 3.5); bathroom_hum.diff = 2.0 (SHAP = -0.65; Z = 3.3).	A 2-unit humidity spike and an 8-minute stay, both high for late afternoon, drove the anomaly.	Yes	No
12/9/2024 11:02–11:13 (8)	-0.0002	duration_sec = 682 s (SHAP = -1.40; Z = 5.2); bathroom_peak.diff = 2.0 (SHAP = -1.15; Z = 2.7); bathroom_max_slope = 1.0 (SHAP = -0.71; Z = 1.5).	An 11.4-minute late-morning stay and a 2-unit humidity spike are longer and more humid than typical.	Yes	No

The bathroom anomaly model singled out 8 anomalies, that are strong deviations from the participant’s baseline behaviour in the bathroom. Each explanation is framed against the relevant *3-hour time window*, and the participant confirmed and considered every interval important. The "Conf." and "Imp." represent whether the participant confirmed that that event actually happened or not and whether the participant considered those anomalies as important or not, respectively. The "Conf." columns serves as validation that the anomaly model makes sense to the participant.

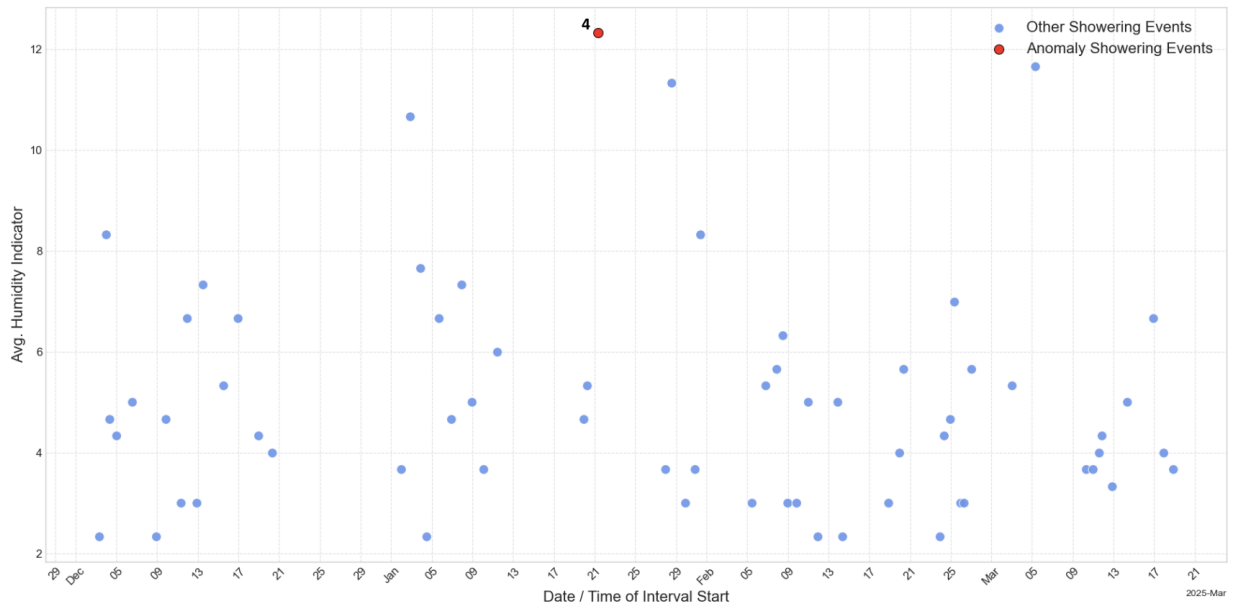
The two figures below illustrate how all anomalies detected in the bathroom differ from bathroom events that were not classified as anomalies. This is to illustrate some of the model’s reasoning behind why they were anomalies. Keep in mind that these models use over 4 features to identify anomalies, so although some are easily explained through the scatter plots below, some anomalies are anomalies because of the relationship of multiple features like humidity, temperature, duration at the same time. So they can not simply be illustrated via scatterplots. To best understand the exact reasoning, please refer to the table above.

Figure 3.3: A Comparison of the Duration of Non-Shower Bathroom Stays



Scatter-plot of every non-shower bathroom stay (blue) versus the seven detected non-shower anomalies (red, numbered 1–8). This plot clearly demonstrates how some of the non-showering anomalies were identified to be anomalies by the model. In this example, anomaly number 1 and 5 clearly had very high durations in the bathroom compared to normal non-showering events.

Figure 3.4: A Comparison of the Average Humidity Indicator of All Showering Events



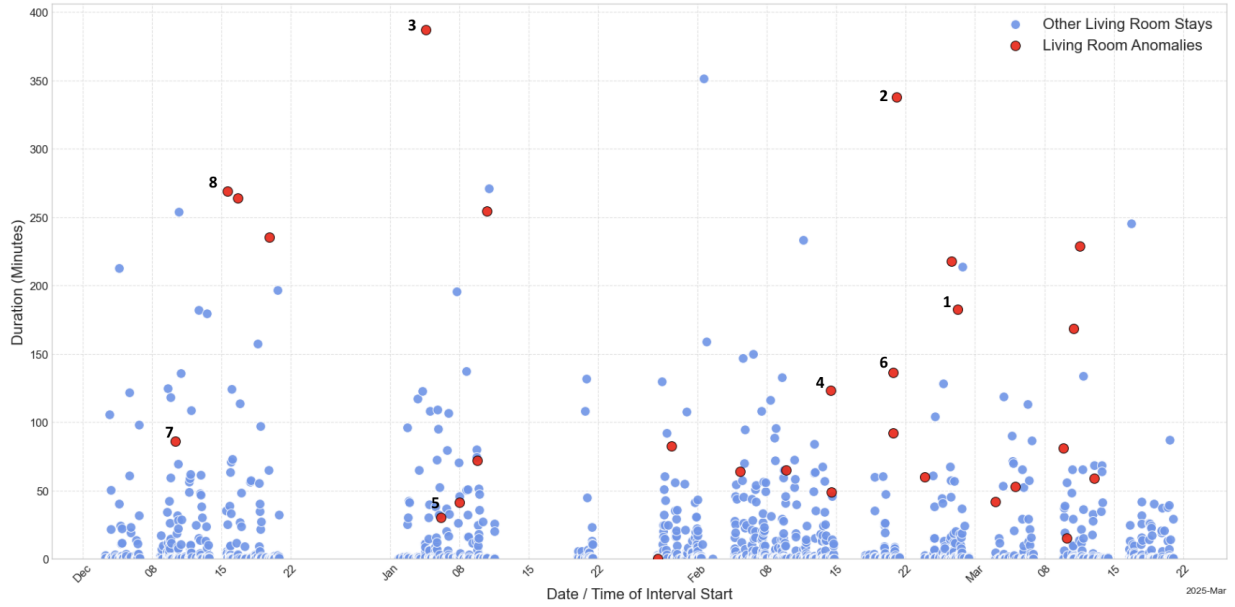
Scatter-plot of every showering event (blue) versus the fourth detected bathroom anomaly (red point labelled “4”). This model clearly demonstrates how the bathroom anomaly model determined anomaly number 4 to be an anomaly. This is because anomaly number 4 was very high in terms of humidity compared to all other showering events.

Table 3.10: Living Room-Detected Anomalies with Participant Feedback

Interval	Anom. Score	Raw Anomaly Reason	Anomaly Explanation to Participant	Conf.	Imp.
27/02/2025 06:33–09:35 (1)	-0.1176	duration_seconds = 10 942 s (SHAP = -4.82; Z = 102.3); hour_sin = 1.00 (SHAP = -1.69; Z = 3.3); hour_cos = 0.00 (SHAP = -0.60; Z = 2.4).	Stayed in the living room for roughly 3 hours straight during the early, morning window, far longer than is typical for that time of day.	Yes	Yes
21/02/2025 02:14–07:51 (2)	-0.0935	duration_seconds = 20 249 s (SHAP = -3.20; Z = 11.2); couch_duration_seconds = 20 250 s (SHAP = -3.00; Z = 13.4); hour_sin = 0.50 (SHAP = -1.10; Z = 0.9).	A continuous 5 h 37 m stay on the couch through the pre-dawn window is rare; hence, the model interpreted as anomaly.	Yes	Yes
04/01/2025 14:12–20:39 (3)	-0.0696	couch_duration_seconds = 23 185 s (SHAP = -3.31; Z = 5.8); duration_seconds = 23 204 s (SHAP = -3.25; Z = 5.8); hour_cos = -0.87 (SHAP = -0.41; Z = -0.3).	Spending over six hours on the couch from mid-afternoon into the late-evening window exceeds usual daytime lounging pattern.	Yes	Yes
14/02/2025 10:20–12:23 (4)	-0.0664	duration_seconds = 7 390 s (SHAP = -5.21; Z = 17.6); hour_sin = 0.50 (SHAP = -0.66; Z = -0.2); hour_cos = -0.87 (SHAP = -0.21; Z = -0.5).	About 2 hours of continuous living-room use during the late-morning window (10:00–13:00) is well above normal short visits around that time.	Yes	Yes
06/01/2025 03:15–03:45 (5)	-0.0519	duration_seconds = 1 815 s (SHAP = -3.83; Z = 79.1); hour_sin = 0.71 (SHAP = -1.37; Z = 0.5); hour_cos = 0.71 (SHAP = -0.63; Z = -0.3).	A 30-minute visit to the living room in the middle of the night (03:00–06:00 window) is highly atypical for you and therefore flagged.	Yes	Yes
20/02/2025 18:33–20:50 (6)	-0.0463	duration_seconds = 8 189 s (SHAP = -4.94; Z = 21.3); hour_sin = -1.00 (SHAP = -0.41; Z = -1.9); hour_cos = -0.00 (SHAP = -0.31; Z = 0.0).	Spending about 2 h 15 m in the early-evening window (18:00–21:00) is much longer than your usual evening routine; hence, an anomaly.	Yes	Yes
10/12/2024 09:31–10:56 (7)	-0.0459	duration_seconds = 5 156 s (SHAP = -5.29; Z = 8.7); hour_sin = 0.71 (SHAP = -0.46; Z = -0.1); hour_cos = -0.71 (SHAP = +0.10; Z = -0.3).	A 1 h 25 m stay in the mid-morning window (09:00–12:00) is much longer than your usual quick morning visits.	Yes	Yes
15/12/2024 14:54–19:23 (8)	-0.0360	couch_duration_seconds = 16 141 s (SHAP = -3.08; Z = 3.4); duration_seconds = 16 142 s (SHAP = -2.84; Z = 3.4); hour_cos = -0.87 (SHAP = -0.40; Z = -0.3).	Roughly 4 h 30 m on the couch across the mid-afternoon to early-evening window (15:00–18:00) is far longer than normal for that timeframe.	Yes	Yes

The living-room activity model flagged 28 anomalous intervals; this table highlights eight of the strongest anomalies. For each interval, the top three SHAP contributors, a plain-language explanation framed against the *3-hour time window* around that period, and the participant’s confirmation and perceived importance.

Figure 3.5: A Comparison of the Duration of all Living Room Stays



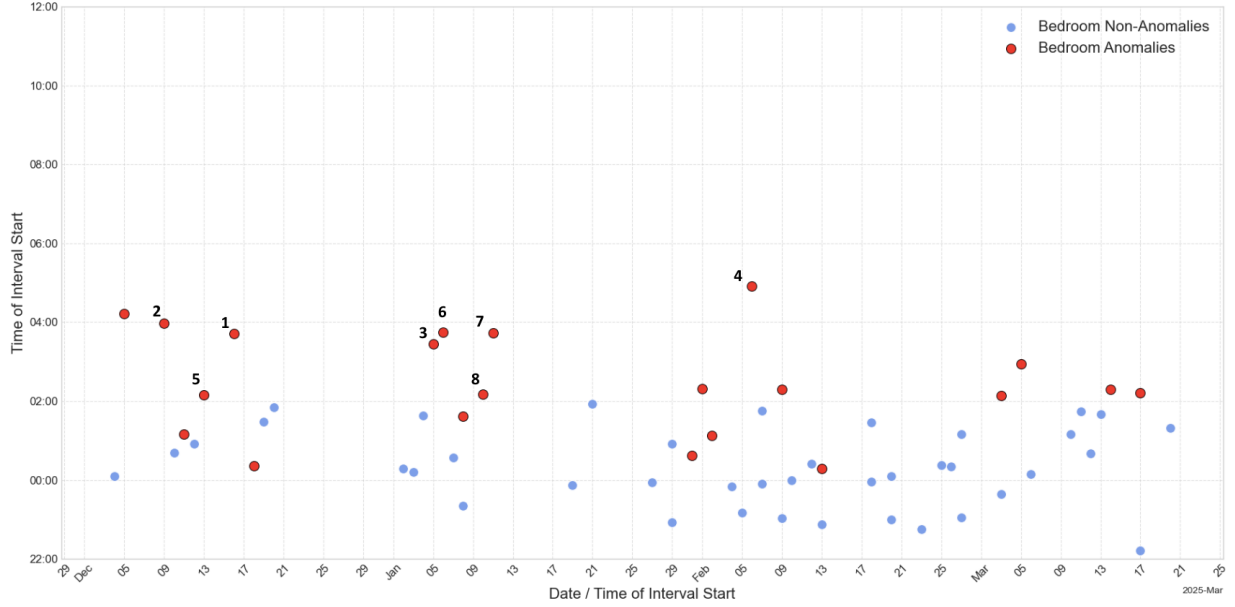
Scatter-plot of all living-room stays (blue) and the eight highlighted intervals (red points numbered 1–8; see Table 3.10). The y-axis gives stay duration in minutes, the x-axis the start date/time. Some points are clearly obvious why they are anomalies simply through indicating a high importance factor for duration in contributing to the status of the interval being an anomaly. Others however highlight the nuanced capabilities of the model to capture duration with consideration to the time of duration as well as couch duration in contributing to an anomaly. The models capability to capture these nuances are shown via the "Raw Anomaly Reason" column and the subsequent explanation in the next in 3.10.

Table 3.11: Bedroom-Detected Anomalies with Participant Feedback

Interval	Anom. Score	Raw Anomaly Reason	Anomaly Explanation to Participant	Conf.	Imp.
16/12/2024 03:43–12:23 (1)	-0.0678	duration_seconds = 31 226 s (SHAP = -5.15; Z = 2.5); hour_sin = 0.71 (SHAP = -1.15; Z = 0.6); hour_cos = 0.71 (SHAP = -0.57; Z = -0.3).	Stayed in the bedroom for nearly 8 h, extending well into the late-morning window (09:00–12:00) without leaving, far longer than your usual overnight rest.	Yes	Yes
09/12/2024 03:58–11:01 (2)	-0.0546	duration_seconds = 25 402 s (SHAP = -4.87; Z = 1.8); hour_sin = 0.71 (SHAP = -1.19; Z = 0.6); hour_cos = 0.71 (SHAP = -0.55; Z = -0.3).	Roughly 7h in the bedroom, ending after 11:00, is much later than your typical wake-up time, so the model marked it as unusual.	Yes	Yes
11/01/2025 03:44–09:21 (3)	-0.0507	duration_seconds = 20 209 s (SHAP = -4.78; Z = 1.3); hour_sin = 0.71 (SHAP = -1.19; Z = 0.6); hour_cos = 0.71 (SHAP = -0.55; Z = -0.3).	A 5 h 37 m stretch of uninterrupted bedroom time, with no presence recorded earlier in the night, deviates from the usual overnight pattern.	Yes	Yes
05/01/2025 03:27–09:38 (4)	-0.0502	duration_seconds = 22 315 s (SHAP = -4.78; Z = 1.5); hour_sin = 0.71 (SHAP = -1.19; Z = 0.6); hour_cos = 0.71 (SHAP = -0.55; Z = -0.3).	Entered the bedroom unusually late (after 03:00) and remained for over 6 h, well beyond your normal overnight routine.	Yes	Yes
06/02/2025 04:55–10:28 (5)	-0.0489	duration_seconds = 20 005 s (SHAP = -4.85; Z = 1.3); hour_sin = 0.87 (SHAP = -1.08; Z = 0.7); hour_cos = 0.50 (SHAP = -0.57; Z = -0.5).	Late-night entry near 05:00 followed by 5 h 33 m of bedroom stay is highly atypical for your sleep schedule.	Yes	Yes
06/01/2025 03:45–08:58 (6)	-0.0477	duration_seconds = 18 809 s (SHAP = -4.73; Z = 1.1); hour_sin = 0.71 (SHAP = -1.19; Z = 0.6); hour_cos = 0.71 (SHAP = -0.53; Z = -0.3).	Presence starts around 03:45 with no earlier bedroom activity and lasts until almost 09:00, longer than your usual overnight period.	Yes	Yes
09/02/2025 06:34–09:58 (7)	-0.0418	duration_seconds = 12 296 s (SHAP = -4.22; Z = 9.0); hour_sin = 1.00 (SHAP = -1.51; Z = 3.2); hour_cos = 0.00 (SHAP = -0.61; Z = 2.4).	A late-morning block of 3 h 25 m bedroom time after 06:30—uncommon compared with your normal wake-up pattern—was flagged as anomalous.	Yes	Yes
03/03/2025 06:33–09:32 (8)	-0.0408	duration_seconds = 10 737 s (SHAP = -4.20; Z = 7.4); hour_sin = 1.00 (SHAP = -1.51; Z = 3.2); hour_cos = 0.00 (SHAP = -0.62; Z = 2.4).	Similar late-morning bedroom stay of just under 3 h mirrors few previous nights and differs markedly from typical morning behaviour.	Yes	Yes

The bedroom model detected 32 anomalous intervals; this table lists the eight strongest. For each interval, the top three SHAP contributors, a plain-language explanation framed against the *3-hour time window* around that period, and the participant’s confirmation and perceived importance are shown.

Figure 3.6: A Comparison of When Individual Entered the Bedroom of All Night-Time Bedroom Stays



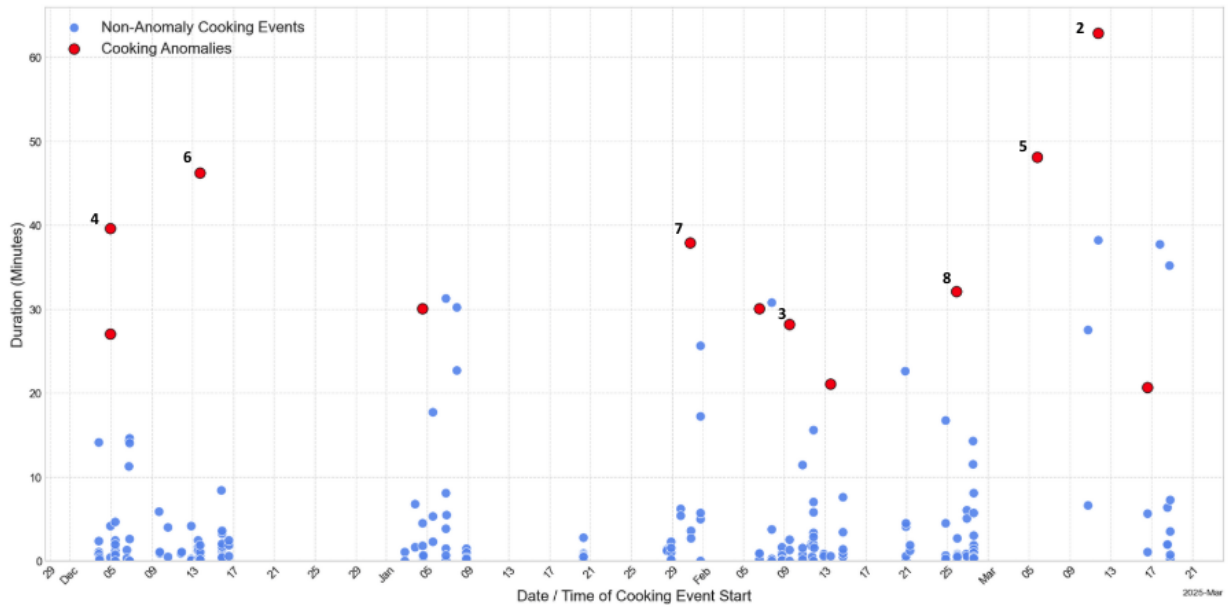
Scatter-plot of the final bedroom entry before each night’s continuous stay until the morning. Each point marks the clock-time at which the participant last entered the bedroom for the night/very early morning (y-axis) and the calendar date of that entry (abscissa). Blue points are routine nights; red points are all the anomalies detected with the top 8, labelled 1–8, to match the intervals listed in Table 3.11. This figure is shown to illustrate why the model marked some of the top 8 models as anomalies to illustrate a bit more as to its reasoning. The anomalies stand out because the participant either entered the bedroom well after the usual bedtime window (typically 22:00–02:00) or left and re-entered the bedroom during the night, both behaviours diverge from the participant’s baseline pattern and drove the strong anomaly scores.

Table 3.12: Kitchen-Detected Anomalies with Participant Feedback

Interval	Anom. Score	Raw Anomaly Reason	Anomaly Explanation to Participant	Conf.	Imp.
04/12/2024 21:36–21:50 (1)	-0.1281	fridge_opens = 11 (SHAP = -3.48; Z = 7.5); cabinet_opens = 3 (SHAP = -2.90; Z = 6.6); duration_seconds = 859 s (SHAP = -1.36; Z = 2.3).	Late-evening burst of fridge (11) and cabinet (3) access within 14 min is far above normal, suggesting unusually intense food prep at this hour.	Yes	Yes
11/03/2025 18:04–19:06 (2)	-0.1271	fridge_opens = 26 (SHAP = -3.09; Z = 11.7); Oven_on_duration_seconds = 1818 s (SHAP = -2.06; Z = 3.5); Microwave_on_duration_seconds = 60 s (SHAP = -1.31; Z = 1.5).	Early-evening cooking session: 26 fridge openings plus a 30 min oven run and microwave use signal an unusually busy dinner prep.	Yes	Yes
09/02/2025 11:37–12:05 (3)	-0.1027	Microwave_on_duration_seconds = 120 s (SHAP = -2.17; Z = 2.4); fridge_opens = 9 (SHAP = -1.52; Z = 6.5); Stove_on_duration_seconds = 680 s (SHAP = -1.06; Z = 2.2).	Just before noon, simultaneous microwave use, 11 min on the stove, and 9 fridge opens mark a heavier-than-usual lunch preparation.	Yes	Yes
04/12/2024 22:23–23:03 (4)	-0.1017	cabinet_opens = 12 (SHAP = -3.56; Z = 4.9); duration_seconds = 2375 s (SHAP = -0.99; Z = 3.8); fridge_opens = 5 (SHAP = -0.78; Z = 1.9).	A 40 min late-night stay with 12 cabinet and 5 fridge opens indicates a sizeable food-related task uncommon for this hour.	Yes	Yes
05/03/2025 19:21–20:09 (5)	-0.0956	Stove_on_duration_seconds = 2405 s (SHAP = -2.62; Z = 6.5); Oven_on_duration_seconds = 1 943 s (SHAP = -2.09; Z = 4.5); fridge_opens = 8 (SHAP = -1.26; Z = 2.1).	Lengthy concurrent stove (40 min) and oven (32 min) usage with multiple fridge opens signals an extended dinner prep far beyond typical.	Yes	Yes
13/12/2024 17:00–17:46 (6)	-0.0526	Oven_on_duration_seconds = 2020 s (SHAP = -2.50; Z = 5.1); Microwave_on_duration_seconds = 60 s (SHAP = -1.22; Z = 1.2); duration_seconds = 2772 s (SHAP = -1.21; Z = 3.5).	Late-afternoon oven run (34 min) with microwave use and a 46 min stay suggest heavier cooking than usual before dinner.	Yes	Yes
30/01/2025 19:07–19:45 (7)	-0.0526	Oven_on_duration_seconds = 2050 s (SHAP = -2.35; Z = 4.8); Stove_on_duration_seconds = 1 596 s (SHAP = -1.83; Z = 3.7); duration_seconds = 2 274 s (SHAP = -0.79; Z = 0.0).	Early-evening cooking: 34 min on the oven and 27 min on the stove during a 38 min kitchen stay marks a substantially heavier dinner prep than usual.	Yes	Yes
25/02/2025 21:34–22:06 (8)	-0.0273	Stove_on_duration_seconds = 1 686 s (SHAP = -2.15; Z = 8.4); Oven_on_duration_seconds = 1 063 s (SHAP = -1.02; Z = 3.1); fridge_opens = 6 (SHAP = -0.82; Z = 3.0).	Late-night double-appliance cooking—nearly 28 min on the stove plus 18 min of oven use and 6 fridge opens, a longer-than-normal kitchen session.	Yes	Yes

The kitchen model flagged 17 anomalous intervals; the eight strongest are listed here. For each interval, the top three SHAP contributors, a plain-language explanation framed against the *3-hour time window* around that period, and the participant’s confirmation and perceived importance are provided.

Figure 3.7: A Comparison of the Duration of All Cooking Events



Scatter-plot of every recorded cooking event between December 2024 and March 2025. Each point plots the calendar date/time that a cooking episode began (abscissa) against its duration in minutes (ordinate). Routine cooking episodes are shown in blue; red points mark episodes flagged as anomalous by the kitchen-stay model, with the eight highest-scoring anomalies labelled 1–8 to match the intervals listed in Table 3.12. The anomalies stand out because they either lasted far longer than typical meal-preparation windows (often exceeding 30–40 minutes) or occurred at times of day that diverge from the participant’s usual cooking pattern, both behaviours driving the elevated anomaly scores.

Chapter 4

Discussion

This study aimed to address persistent limitations in AAL research, specifically the dependence on supervised, heavily-labelled datasets (Gap i), the use of general, population-level anomaly models that ignore individual baselines (Gap ii), the need for models that can change over time rather than being static (Gap iii), and evaluations often performed in short, artificial laboratory settings (Gap iv). To overcome these, we pursued two primary objectives: (1) To design and implement a robust, multi-modal sensor system for long-term, real-world data collection in a participant’s home; and (2) To analyze this data using primarily unsupervised machine learning techniques to first identify key activities (cooking, showering, couch-sitting) and features, and subsequently build personalized, interpretable models for behavioural anomaly detection. This discussion evaluates the extent to which these objectives were met by directly tackling the identified gaps, and considers the broader implications of the findings.

4.1 Objective 1: System Implementation and Data Collection

The first specific objective of this study was to design and implement a robust smart-home system for ecologically valid, in-home data collection. This involved deploying a multi-sensor setup (including motion, temperature, humidity, door contact, and smart outlet sensors) managed by a Raspberry Pi, storing raw data locally in InfluxDB, and establishing a secure pipeline to automatically push encrypted, de-identified data to a University of Waterloo server for analysis.

This objective was successfully achieved. The core infrastructure demonstrated exceptional reliability and resilience throughout the entire 3.5-month data collection period (December 2, 2024 – March 20, 2025). The Raspberry Pi operated with virtually no interruptions or critical errors requiring manual intervention. Furthermore, the complete data pipeline, encompassing data acquisition from Zigbee/Z-Wave sensors via Domoticz, local storage in InfluxDB, and automated, secure transfer to the remote server, functioned seamlessly and autonomously as designed. The system also proved resilient to real-world disruptions; following intermittent power outages within the apartment, the Raspberry Pi and the associated data

collection processes automatically restarted and resumed operation without data loss or the need for manual resets. This level of stability aligns with findings from other research utilizing Raspberry Pi for AAL data collection, confirming its suitability for long-term, unattended deployments.^{40;41;52}

Sensor performance varied across types, highlighting practical aspects of real-world deployments. A subset of sensors exhibited outstanding reliability: the smart outlet sensors, door contact sensors, and the air quality sensor (measuring temperature and humidity) functioned flawlessly for the entire study duration. These sensors consistently transmitted data without any instances of failure, data gaps, or requiring battery replacements, providing a continuous and high-fidelity stream of data crucial for detecting activities like cooking, showering, and appliance usage discussed later.

Conversely, challenges were encountered with some of the motion sensors. Early in the deployment, approximately four motion sensors in the bedroom and one in the living room exhibited intermittent functionality or ceased transmitting data altogether. This was attributed partly to a lack of consistent, real-time system health checks that could have flagged these issues sooner. To maintain the integrity of the dataset and ensure the accuracy of downstream analyses, particularly room presence detection, which is sensitive to missing motion data, a strict protocol was adopted: any motion sensor showing significant periods of data loss was entirely excluded from the final analysis dataset.

Despite the removal of these specific motion sensors, the overall system successfully captured a rich, multi-modal dataset over an extended period in the participant’s natural living environment. The successful implementation of this robust data collection platform directly addressed Objective 1 and provided the high-quality, ecologically valid foundation necessary for the subsequent unsupervised activity recognition and anomaly detection analyses (Objective 2). The challenges encountered with specific motion sensors, while managed effectively in this study, underscore the importance of incorporating robust sensor health monitoring and potentially exploring alternative sensor hardware in future long-term AAL deployments.

4.2 Objective 2a: Unsupervised Activity and Feature Identification

Objective 2a focused on using the collected data to identify key features and discover specific activities (cooking, showering, couch-sitting) primarily through unsupervised methods, thereby reducing the reliance on exhaustive labelling. Although, some minimal labelling and an uncertainty flagging and correct by the participant approach were also discussed for showering detection in the results section. The primary reason, as referred to in the results section, these two approaches were discussed was because a completely unsupervised showering detection results in some false classifications, and these approaches further minimize these classification whilst reducing exhaustive labelling.

4.2.1 Outlet-Based Activity Recognition and Additional Features

Addressing Objective 2a, this study demonstrated the successful unsupervised identification of cooking and television usage by repurposing low-cost smart outlet sensors, thereby reducing the need for intrusive validation or burdensome ground-truth labelling (addressing Gap i).

4.2.1.1 Cooking Activity Recognition

Cooking was operationally defined based on the activation state (ON/OFF) of the stove and/or oven, detected via dedicated outlet sensors. Within this, a distinction was made between "active cooking" (cooking appliances on while the participant was present in the kitchen) and "passive cooking" (appliances on while the participant was elsewhere); however, subsequent anomaly detection models primarily focused on active cooking, as passive cooking occurred too infrequently in this participant's routine to establish meaningful patterns for analysis. The determination of the core appliance ON/OFF states was achieved through a purely data-driven, unsupervised approach using k-means clustering applied to the one-second amperage readings from each relevant outlet. This analysis yielded remarkably strong cluster separation between ON and OFF states for key kitchen appliances (stove, oven, microwave), confirmed by the exceptionally high silhouette scores presented in the Results (Table 3.2). This strong bimodal distribution affirmed that unsupervised clustering could reliably distinguish appliance states without labelled training data or predefined thresholds.

To validate these findings using a minimally intrusive method aligned with the unsupervised goal, periods of known home vacancy were leveraged (based on participant confirmation that these appliances were never left on while away). Comparing the cluster-derived states to these known OFF periods confirmed exceptionally high agreement rates, over 99.6% for the actively used stove, oven, and microwave, and 100% for the unused rice cooker. Minor discrepancies were attributed to inherent sensor noise in high-resolution amperage readings. Based on this validated unsupervised detection, 157 (5.2%) of the 3002 detected kitchen presence events were identified as cooking periods involving these appliances.

4.2.1.2 Television ON Detection

A similar unsupervised k-means clustering methodology was successfully applied to the outlet sensor dedicated to the television (OUT-2). This again resulted in strong cluster separation between ON and OFF states (Table 3.2). Validation using the home vacancy method yielded a high agreement rate of 97.24%, with the slightly lower rate compared to kitchen appliances likely due to the TV's lower power consumption making it more susceptible to sensor noise.

4.2.1.3 Contribution of this Research to Outlet Modelling

Beyond these specific activities, the unsupervised clustering performed robustly across other monitored outlets in the bedroom and bathroom (silhouette scores greater than 0.93), suggesting broader potential for inferring context or other activities. The primary contribution of this work lies in validating that unsupervised clustering of data from low-cost, commercially available smart outlets provides a viable, scalable, and privacy-respecting method for accurately determining appliance states and recognizing associated activities like cooking and TV viewing. This approach directly addresses the challenge of Gap (i) by eliminating the need for expensive, specialized hardware with built-in state detection or reliance on manual, time-consuming data labelling, thus offering significant advantages for practical, long-term AAL deployments.

While many sensors more expensive than the affordable ones used in this study can automatically determine the ON/OFF state of connected appliances, the unsupervised analysis presented here demonstrates

significant value.⁵³ It opens the door for any basic outlet sensor, regardless of built-in intelligence, to be effectively used for detecting appliance states and inferring activities pertinent to anomaly detection.^{53;54} Methodologies enabling low-cost sensors to detect ON/OFF states via electrical current analysis have been explored previously.^{53;54} This study builds upon and extends such work by not only classifying appliance states but explicitly linking these states to defined activities (e.g., cooking) relevant for behavioural monitoring, validating the approach with specific, commonly available sensors. Furthermore, the unique validation method employed, leveraging periods of known absence, adds to the robustness of these findings and offers a practical technique for future validation efforts in similar unsupervised contexts.^{53;54}

4.2.2 Couch Sitting Detection Findings

Further advancing Objective 2a, this study successfully identified periods of couch-sitting using an innovative approach leveraging vibration and motion sensors. Sustained vibration detected at the couch, combined with simultaneous motion sensor confirmation of living-room presence, proved a simple yet highly discriminative proxy for seated behaviour. A 30-second persistence criterion was implemented to filter out transient bumps, while a specific requirement that vibrations cease and motion be detected elsewhere before closing a sitting cycle, ensured graceful handling of posture shifts and room transitions. This dual-sensor logic echoes the trajectory-propagation principle observed in other work where temporally coherent binary events mark movement between zones.⁵⁵ Critically, by its design, this algorithm was entirely unsupervised, requiring no labelled training data, and inherently avoided privacy-invasive methods like vision or audio capture, representing a key advantage for acceptable real-world AAL deployments.

Regarding performance, across 2,632 detected living-room presence episodes, this unsupervised method classified 727 intervals (27.6%) as sitting on couch. The validity of these classifications was confirmed through participant spot-checks covering every detected sitting bout that lasted longer than one hour; these checks reported zero false detections for these extended periods. This accuracy is consistent with preliminary findings in the literature suggesting the promise of such methodologies.⁵⁶ Furthermore, this study contributes by being one of the first to specifically validate this approach for couch presence using a commercially widely available and affordable vibration sensor. It is important to note, however, that shorter sitting bouts (less than 10 min) were not individually validated in this study, leaving open the possibility of occasional missed micro-sits. Nonetheless, their potential omission or mis-classification is unlikely to have biased the subsequent anomaly detection models, as no anomalies related to unusually short couch durations were flagged by those models, indicating that these short events, whether perfectly captured or not, occurred with sufficient consistency to be considered normal within the participant’s routine.

4.2.2.1 Prior Work and Specific Implications

Contextualizing this within prior work, a few studies have indeed exploited structural vibrations to infer occupant state while sidestepping the privacy trade-offs associated with cameras or wearables.^{55;56} These studies demonstrated that sustained low-frequency oscillations, characteristic of static postures like sitting or standing, can be reliably detected provided that transient spikes are suppressed, a function precisely fulfilled by the 30-second persistence filter employed here. This study’s results extend these findings by

showing that even a single vibration sensor, when augmented with appropriate rule-based logic incorporating motion data, is adequate for effectively monitoring presence specifically on a couch.

The implications of reliably generating these couch-sitting labels are significant for health monitoring. Firstly, they enable the estimation of sedentary time with room-level resolution, an indicator increasingly linked to cardiometabolic risk, particularly in aging populations.⁵⁷ Secondly, coupling sofa occupancy data with other sensor streams, such as television power data derived from the smart outlets, allows for disentangling passive screen time from other forms of sitting, thereby enriching the interpretability of downstream anomaly scores. For instance, an unusually lengthy sitting bout without concurrent TV usage might flag potential lethargy or mood changes, providing more nuanced behavioural insights.⁵⁷ With minimal instrumentation and no privacy compromise, this vibration-based detection method successfully captured nearly one-third of the participant’s living-room presence as confirmed seated leisure. These granular, unsupervised labels feed directly into the personalized, interpretable anomaly detection pipelines, strengthening the case for vibration sensors as a cornerstone technology for unobtrusive home health monitoring.

Finally, certain limitations should be acknowledged with this specific activity that were not addressed in this study. The single-occupant nature of this study simplified the cycle termination logic, as any motion detected outside the couch area definitively implied the participant had left. Additionally, the 30-second threshold, while effective for filtering noise in this instance, might obscure clinically meaningful micro-breaks within longer sedentary periods. However, identifying such micro-events was not an objective of the current research, making this limitation acceptable within the study’s defined scope.

4.2.3 Showering Detection Findings

Detecting showering activity using ambient sensors like air quality monitors presents unique challenges due to variability in home environments (e.g., bathroom size, ventilation) and individual habits. Therefore, specifically for this activity and as part of achieving Objective 2a (activity discovery), this study explored and compared both unsupervised and minimally supervised learning approaches. This comparative analysis aimed to understand the performance trade-offs and directly address the significant limitation of reliance on heavily labelled datasets in AAL research (Gap i).

First, an unsupervised approach was evaluated using a two-cluster k-means model trained solely on unlabelled humidity features derived from bathroom presence intervals. This model achieved a high overall accuracy of 99% but exhibited limitations in recall, failing to correctly identify 9% (6 out of 64) of the true showers (Table 3.6). These missed events were typically shorter showering sessions clustering near the decision boundary where humidity changes were less pronounced. Nonetheless, the potential for applying the study’s interpretability methods to flag uncertain cluster assignments suggests that even this fully unsupervised approach could achieve high precision by prompting user confirmation for ambiguous cases, thus minimizing false positives.

To evaluate the impact and necessity of labelling, minimally supervised models (Random Forest and Logistic Regression) were explored. These classifiers were trained on progressively larger, yet still very small, initial portions of chronologically ordered, participant-labelled data (ranging from 5% to 40% of intervals, corresponding to approximately 3 to 25 days of labelled data). The Random Forest (RF) classifier, in particular, demonstrated near-ceiling performance, achieving essentially error-free shower recognition

after being trained on only the earliest 30% of the dataset (approx. 20 days), with negligible gains from incorporating more history (Table 3.3). Even with just 3 days of labelled data (the 5% split), the RF model performed remarkably well. This rapid performance plateau underscores the data efficiency of such models for this task; high accuracy, precision, recall, and F1-scores were achievable with just a few weeks (or even days) of initial labelling. SHAP analysis (Fig. 3.1) provided interpretability, confirming domain intuition: the supervised model relied almost exclusively on the magnitude and velocity of humidity rise, followed by event duration, to identify showers. These features capture the archetypal shower signature (rapid condensation, sustained moisture), explaining the model’s robustness and the complete absence of false negatives across all tested data splits, even short rinses produced humidity spikes distinct from non-shower baselines.

Critically, a key finding applicable to both the unsupervised and supervised approaches is the ability to identify uncertain predictions, forming the basis of a practical semi-supervised or active-review strategy. For the supervised RF model, predictions with low confidence (e.g., probability near 0.5) could be flagged. For the unsupervised k-means model, instances close to the cluster decision boundary represent similar ambiguity. As detailed in Table 3.8, strategically flagging these few uncertain instances (less than 1% of intervals for the best RF model, the 7 misclassified instances for k-means) successfully captured all misclassifications made by each respective model. This demonstrates a powerful, practical approach: deploy either an unsupervised or minimally supervised model, and use its inherent uncertainty metrics to prompt targeted user verification (“active review”) only for the small subset of ambiguous events. This strikes an effective balance between autonomy and accuracy, drastically reducing labelling effort while virtually eliminating classification errors and ensuring the robustness of subsequent models relying on these activity labels.

4.2.3.1 Prior Work and Specific Implications

Comparing this work with prior research, there has been limited exploration of using non-intrusive air quality sensors for robust shower detection, whether supervised or unsupervised.^{58;59} This study validates and bolsters recent suggestions that such sensors can indeed detect showering with high confidence, offering effective methodologies and strong validation for both labelled and unlabelled scenarios.^{58;59} A distinguishing feature of the presented models is their demonstrated ability to recognize prediction uncertainty and leverage it for a highly efficient active review process. One potential limitation, acknowledged here but not observed to cause issues in this specific single-occupant apartment, is the possibility that prolonged steam from cooking in open-plan living spaces could theoretically mimic a shower’s humidity signature, requiring further investigation in diverse settings.

In summary, the specific takeaway from this comparative exploration is that highly accurate shower detection is achievable using ambient humidity sensors with minimal or potentially zero ongoing annotation burden. While the fully unsupervised approach was promising, a minimally supervised model trained on just a few days of data yielded near-perfect results. More importantly, leveraging model uncertainty via a semi-supervised active-review strategy allows either approach to achieve state-of-the-art accuracy while requiring user input for only a tiny fraction of events. This balance of accuracy, interpretability, and scalability, demonstrated effectively even with the added challenge of the participant primarily taking cold showers (which produce less distinct humidity signals), directly addresses the need to reduce labelling dependence (Gap i) and positions these humidity-based shower detection models as robust and transferable components for future AAL systems.

4.3 Objective 2b: Personalized and Interpretable Unsupervised Anomaly Detection

Building upon the successfully identified features and activities under Objective 2a, the study directly addressed Objective 2b: developing and evaluating unsupervised machine learning models for personalized and interpretable behavioural anomaly detection using non-intrusive sensors. As highlighted previously, the ability to reliably detect specific activities, like cooking, showering, and couch-sitting, and actions, like door/cabinet usage via contact sensors, without supervision provided a rich, contextual foundation. This context-aware feature set, encompassing activity labels, temporal information (time, duration), sensor readings (temperature, humidity), and presence in each of the rooms, allowed the anomaly detection models to move beyond identifying simple statistical outliers. Instead, they could pinpoint deviations specifically related to the participant’s established routines and behaviours, such as activities occurring at unusual times or lasting for atypical durations at either the activity or simply the room-level. This approach inherently learns individual baselines directly from the participant’s own long-term data gathered within their actual home, ensuring ecological validity (Gap iv) and directly addressing the need for personalized models (Gap ii) rather than relying on generic, population-level profiles. Furthermore, the use of interpretable unsupervised models (Isolation Forest) coupled with explainability techniques (SHAP analysis and dynamic Z-scores, detailed in Tables 3.9-3.12) was crucial for tackling the “black box” problem, providing clear reasons for flagged anomalies and enhancing trustworthiness. While the models implemented in this phase were static, the long-term data foundation and potential for incorporating user feedback back into the models, as discussed later, lay the groundwork for future adaptive systems (Gap iii). It is worth noting that while luminance features were included, they did not emerge as significant contributors to any of the top anomalies for this particular participant, though their relevance may vary for others. Additionally, due to insufficient usage data, an anomaly model specifically for television viewing was not developed in this study.

4.3.1 Bathroom Anomaly Model Findings

The bathroom model very effectively identifies and expresses the reasoning behind the top anomalies. The sub-modelling approach which was used to separate and find anomalies within bathroom events where showering took place and events where showering did not take place, allowed for anomalies that were specific. One of the anomalies listed in the table was because it was specifically abnormal during showering events as referenced in Figure 3.4. Whereas the other 7 anomalies were abnormalities within the non-showering category. Although the 1 showering anomaly is clearly depicted on a scatter plot (3.4) as to why it was an anomaly, most anomalies chosen are not that simple. This is because the foundational model used in this study is an isolation forest, which allows for the thorough examination of how multiple independent variables, at one time, can be used to identify patterns. Most anomalies were not just anomalies because of one main factor, they were anomalies because of the combination of different factors including duration, the time of duration, humidity levels, and temperature. As a result these types of anomalies are not very easily visualized, and would require 3 or 4 dimensional plots to account for the different independent variables.

Another key aspect of these models is their explainability. Due to the variety of complex work done on explainability of the artificial intelligence model, the model is able to semantically indicate not only

the most important factors that contributed to the anomaly but also whether they were too high or too low and what is the degree of difference the current variable's values are compared to what was expected. This is through a combination of SHaPley analysis and dynamic Z scores attained using a 3 hour window. The interpretations stated in the "Raw Anomaly Reason" column in the table 3.9, are the result of this explainability work, which can then be used to inductively develop clear explanations, either automatically through heuristic logic using Python, Large Language Models or manually. These explanations can then be used to explain the anomalies to the participant. As noted in the Table, none of the explanations used anything other than what was explicitly derived from the "Raw Anomaly Reason" column and the time of the bathroom stay. Overall, all anomalies have been confirmed to have occurred by the participant; and, some anomalies have also been stated to not be as important. The anomalies not considered important were because they had to do with humidity levels during non-showering events that the participant was not able to recall what they did differently at that time for it to result in an anomaly. This speaks to future work being focused on how participant feedback can be used as labelled inputs which the model can use to learn anomalies that are important to the participant and present anomalies accordingly.

4.3.2 Living Room Anomaly Model Findings

The living room overall flagged 28 anomalies compared to the bathroom anomalies being only 8. This is one inherent flaw to the modelling architecture. For context, all anomaly models had the same threshold of 0.01, which meant that it identified the top 0.01% of outliers, as related to the various different features within each room, to be labeled as anomalies. This means that in this study, for rooms like the living room which inherently had more visits overall than the bathroom, 0.01% of more of those visits means means more anomalies. However, the main way this can be assessed is through participant feedback. As mentioned in the previous subsection, future work will be focused on how participant feedback, once they have been notified that an anomaly took place and why it has taken place, can be used to develop models that can learn to apply filters on the anomalies, to choose only the most important anomalies, rather than simply the top 0.01%. The reasoning behind the 0.01% as explained in detail in the Methods section was due to this study prioritizing the biggest anomalies detected using the features and validating those feature for proof of concept. However, the actual threshold should be more so based on participant feedback, which the models used in this study have the capability of taking on as inputs in future studies. In addition, there was one sensor placed underneath the dining table in the living room; however, that sensor stopped sending in data halfway through the data collection period. As a result this sensor was removed and not used to analyze patterns related sitting at the dining table.

Similar to the bathroom anomalies, living room anomalies also had two sub-models. One model was to identify anomalies related to the participant sitting on the couch and the other was when they were in the living room and not sitting on the couch. This once again enables for specific anomalies that may identify patterns between features that are normal in one sub-model but not in the other. Figure 3.10 illustrates the key point made about how the model recognizes relationships between multiple features to identify some of these anomalies. For example, one of the most important features according to SHaPley as seen in Table 3.10 is duration. However, Figure 3.10 illustrates how it is not obvious just looking at a scatter plot with duration as the independent variable why they were anomalies. This is expected and, in fact, the goal. Anomalies, for example, should not just focus on just focused on feature like duration, but focus on abnormalities that may have occurred because of prolonged duration in the living room *and* it was at an odd time compared to what is normal for the participant. This logic holds true for all the

anomalies produced in this study. This enables true anomaly detection that moves beyond just looking at one main factor, leading to nuanced findings that could potentially help enable users to learn from their unique patterns and abnormalities and course-correct.

4.3.3 Bedroom Anomaly Model Findings

The bedroom anomaly was quite possibly the most different compared to the other rooms because there was no sub-model for the bedroom. There was only overall model because there were no specific activities without the bedroom to differentiate. Although there was one action related to the bedroom dressers in the bedroom opening and closing, they did not happen frequently enough nor were they directly linked to any particular health factor justify a distinct sub-model just for stays in the bedroom that included a dresser being opened as opposed to not. In addition, around 4 of the motion sensors placed in the bedroom initially stopped working early on, due to a lack of consistent checks by myself to ensure the sensors were up and running. This prevented some more detailed anomalies related to different areas of the bedroom, including the desk area.

In general; however, there were 32 total bedroom anomalies, of which, similar to the living room, the 8 biggest anomalies were presented in the results section 3.11. Bedroom anomalies were interesting as they all mainly focused on late night bedroom stay interruptions due to visits to the washroom late night which was unexpected and not very common for the participant. This trend is reflected by Figure 3.6, which clearly shows that the top 8 biggest anomalies were because the participant entered the bedroom at odd times as they are usually expected to be sleeping at the time.

4.3.4 Kitchen Anomaly Model Findings

The kitchen has most variety in terms of different types of sensors. A few key sub-models were derived in this model including kitchen intervals where the participant did cooking, no cooking but participant used the microwave, and no cooking or microwave use but the cabinets were accessed. This allowed for an overall kitchen model that detected anomalies related to things like abnormal cooking duration and time of cooking, abnormal kitchen patterns when using the microwave specifically as well as any abnormal behaviour related cabinet usage either by volume or time of day.

One key advantage of using sub-models is that it allows for the top 0.01% of anomalies within each sub-model's conditions rather than just overall. Overall models tend to drown out key event-based anomalies, simply because they don't occur as often.

Overall the participant found the kitchen intervals to be most interesting because they had such a variety of different types of conditions. One key learned aspect from the kitchen is that the usage of door contact sensors and outlets sensors enriched the ability to detect sub-models that would traditionally not be possible from simply motion sensors.

4.3.5 Comparison with Prior Work

Despite there being some research into unsupervised anomaly detection for AAL, there remain important gaps that the present anomaly model aims to fill. One clear gap is the granularity and completeness

of sensing in many prior deployments.⁷ Typically, as noted, smart home studies often have for example one motion sensor per room, no differentiation of multiple concurrent events.⁷ In contrast, this study employs a densely instrumented sensor suite including motion, luminance, vibration, door sensors, temperature/humidity, and electrical usage monitors in a single-occupant home. By processing and merging these multi-modal signals it allows for the determination of understanding which room the person is in at a high granularity and even detect activities, that would be invisible to sparser setups.^{25–28;60} This high-resolution behavioural model of routines allows this system to detect subtle anomalies, like an unusual pattern of brief room-to-room shuttling that might indicate restlessness, that would likely go unnoticed with fewer sensors or coarser data.^{25–28;60}

Another significant gap this research addresses is the lack of continuous validation and interpretability in prior unsupervised systems. This research builds on the insight that anomaly detection in healthcare contexts must be trustworthy and understandable. To that end, I employ interpretable unsupervised models that not only flag anomalies but also provide insight into which sensor readings or inferred activities contributed to the anomaly. Unlike black-box neural nets or just machine learning models as used in previous work, these models yield explanations such as “unusually long inactivity in the living room without the TV on and without sitting on the couch”. This interpretable approach directly addresses the black-box criticism in the literature and ensures that when this system raises an alert, it can be immediately contextualized by caregivers or by the occupant themselves.^{27;28} In practice, this means the participant provided feedback whenever the system thought something odd had happened, effectively serves as an oracle that labels events as true anomalies or false alarms on an ongoing basis.

The result is research that addresses the overall gap of research in unsupervised anomaly detection using non-wearable, non-camera based sensors and is interpretable. The research also results in a model that has the potential in future work to be tuned continuously by the end-user. Ultimately leading to high precision in distinguishing meaningful anomalies from benign routine variations.

4.4 Key Takeaways & Contributions

This study yielded several key takeaways and contributions that directly address persistent limitations in AAL research, particularly those outlined in the introduction: (i) the dependence on supervised, heavily-labelled datasets; (ii) the use of general, population-level anomaly models ignoring individual baselines; (iii) evaluations performed in short, artificial laboratory settings; and (iv) the associated lack of model interpretability and trustworthiness.

A significant methodological contribution lies in directly confronting the limitations associated with supervised methods.^{2;3;7;25;27} Crucially, this study successfully implemented and validated a fully unsupervised anomaly detection framework using non-intrusive sensors. This achievement demonstrates the feasibility of deriving meaningful behavioural insights without the practical and ethical burdens of continuous user labelling, thus directly addressing Gap (i) by offering a more scalable and participant-friendly approach for real-world, long-term deployments. Furthermore, the use of sub-modelling based on identified activities or contexts (e.g., showering vs. non-showering bathroom events, specific kitchen activities) represents a further methodological refinement within this unsupervised paradigm, allowing for more nuanced, context-aware anomaly detection that might be obscured in monolithic models.

Building on this, the research makes a theoretical contribution to behaviour modelling within ambient intelligent systems by championing personalized modelling, directly tackling Gap (ii), the reliance on generic, population-level models. Consistent with the literature’s emphasis on personalization, the unsupervised, personalized modelling approach inherently captured the unique routines and behavioural variability of the participant.^{3;29} Anomalies were identified as deviations from this individual’s established norm, advancing the understanding of how to effectively model idiosyncratic behaviour in real-world settings.

From a practical standpoint, the study provides important contributions by confronting the ecological validity limitations common in laboratory-based AAL research, thereby addressing Gap (iii). By conducting the research within a real-world residential setting over an extended period (greater than 90 days), it offers ecologically valid insights and demonstrates the practical feasibility, as well as the inherent challenges, like sensor reliability, of deploying such systems long-term in genuine living environments.^{2;3;23;30;31}

Furthermore, addressing growing concerns about transparency and ethical AI, this work makes significant ethical and interpretability contributions.^{34–37} The integration of explainable AI techniques, specifically SHAPley analysis combined with dynamic timed Z-scores, directly tackles the “black box” problem and addresses Gap (iv). It demonstrates how complex model outputs can be translated into understandable and actionable explanations, as seen in the “Raw Anomaly Reason” findings. This enables non-technical users or caregivers to comprehend detected anomalies, a critical factor for building user trust and ensuring ethical deployment, addressing the often-cited need for trustworthy AAL systems.

Finally, the findings highlight the necessity for adaptive systems, pointing towards solutions for the limitation of static models.^{2;3;7;24;29} Participant feedback regarding the importance of detected anomalies and the observed limitations of using a fixed anomaly threshold underscore that future methodological and practical advancements should focus on incorporating feedback loops. This points towards creating systems capable of learning user preferences and evolving over time, making the monitoring more relevant and effective.

4.5 Implications of the Research

The implications of this research are multifaceted, extending across technical, clinical, societal, and ethical domains. From a practical and technical perspective, the demonstrated unsupervised and explainable approach enables more scalable, real-time monitoring capabilities for smart homes, moving beyond the significant practical limitations of labour-intensive supervised methods.^{2;3;25} Societally and ethically, this work strongly supports the goal of aging-in-place and enhancing independence, autonomy, and quality of life, not only for older adults but potentially for a broader range of individuals needing support, such as those with chronic conditions or disabilities.^{1;3;16;17;19} Crucially, the system’s ability to detect subtle, personalized deviations from established routines facilitates preventive health strategies through the early recognition of potential health declines or emerging issues before they become acute.^{2;3;6;22} This capability, combined with the explainability of the detected anomalies, can better inform timely interventions and clinical decision-making.^{34;35;37} In practice, this could manifest through user-friendly interfaces, such as dashboards, presenting individuals or their designated caregivers with summaries of daily routines, activity logs, and clearly explained, flagged anomalies. Configurable alerts for significant deviations could notify family or care providers, prompting timely check-ins or facilitating proactive adjustments to care plans.

Clinicians, with patient consent, could leverage this longitudinal behavioural data to gain a richer understanding of a patient’s functional status and well-being between appointments, potentially aiding in the management of chronic conditions, tracking recovery trajectories, or supplementing cognitive assessments. Furthermore, by providing a reliable yet unobtrusive safety net, such technology has the potential to reduce caregiver burden, offering peace of mind through continuous, passive monitoring and delivering specific, actionable information when deviations occur, rather than requiring constant active supervision.^{1;15;17} The focus on generally acceptable non-intrusive sensors and the avoidance of complex labelling tasks may also promote health equity by offering a potentially more accessible monitoring solution, aligning with the broad global need for assistive technologies.¹⁸ Looking forward, the groundwork laid here for identifying personalized behavioural anomalies could also support public health applications. The heavy emphasis on unsupervised approaches allows for this type of system to have higher adoption due to less active participant participation, in turn allowing for improve community-level health monitoring.^{3;61} For example, aggregated and anonymized patterns might eventually contribute to community health analytics or early warning systems for population health trends.^{3;61} For instance, analyzing anonymized data across many households could reveal seasonal variations in activity levels or sleep patterns, identify widespread behavioural shifts potentially linked to environmental factors (like heatwaves) or community health events (like influenza season), and potentially highlight areas with higher needs for specific support services.^{3;61} Such insights, derived ethically and with robust privacy protections, could inform public health resource allocation, guide preventative campaigns, and provide evidence to support policies favouring the integration of effective AAL technologies into broader health and social care systems.^{3;61}

4.6 Limitations and Future Work

A primary limitation is that the study involved only a single participant. Although this allowed for in-depth, personalized anomaly detection and system creation that is able to be scaled, there needs to be validation that it can be scaled and applied to more individuals in the future. This is especially important to some of the features used effectively to detect activities like the shower need to be validated. Future work should prioritize replicating and extending this methodology with larger and more diverse participant cohorts to establish that this system can be deployed to develop the same individual-level models using the same logic developed in this study for many people.

On a technical level, the use of heterogeneous sensors introduces challenges. As noted, different sensors update their status at different rates (e.g., motion vs. door sensors). This temporal discrepancy can introduce noise or inaccuracies in algorithms relying on the precise timing and fusion of sensor events, particularly for complex logic like fine-grained presence detection within the apartment. Future iterations could explore more sophisticated sensor fusion techniques or state estimation algorithms robust to these timing variations.⁶² Furthermore, sensor reliability and data gaps emerged as practical limitations during this real-world deployment, with failures impacting the dining table and bedroom motion sensors, as discussed earlier. While addressed by removing affected sensors from analysis, future systems should incorporate robust strategies for detecting sensor failure, handling missing data streams gracefully, and potentially include automated system health checks.

Additionally, specific sensor choices impacted the granularity of insights. The lack of a vibration sensor on the bed, for instance, limited the ability to definitively distinguish between the participant simply being in the bedroom versus resting or sleeping in bed, similar to how couch presence was detected. Incorporating

such sensors in future deployments would enrich the activity recognition and anomaly detection capabilities within key locations like the bedroom.

Perhaps the most significant area for future work lies in closing the loop between anomaly detection and user value by incorporating participant feedback for model adaptation. As highlighted in the discussion, not all statistically detected anomalies were considered important by the participant, and the fixed anomaly threshold proved suboptimal. A crucial next step is the development of a user-facing application (front-end) that presents the detected anomalies and their explanations (generated via SHAP/Z-scores) to the participant or designated caregivers. This interface would allow users to confirm, reject, or rate the importance of flagged anomalies. This feedback can then be used to adapt the anomaly detection models dynamically, enabling the system to learn which types of deviations are subjectively meaningful to the individual. This could involve retraining models with labelled feedback, adjusting anomaly thresholds per user or context, or implementing filtering layers based on learned user preferences. This aligns directly with the need for adaptive AAL systems identified in the literature and moves towards a truly personalized and user-centered monitoring system.^{3;29}

Furthermore, the current anomaly detection focused primarily on deviations within specific room presence events or activities. Future research should explore the use of advanced machine learning models, particularly deep learning-based sequential models.²⁶ These models excel at capturing complex temporal dependencies within sequences of behaviour. Applying them could enable the detection of anomalies occurring over broader time scales, such as deviations in patterns across a 6-hour block or an entire day. This could reveal subtle shifts in circadian rhythms, changes in the sequential ordering of daily activities, or alterations in overall activity levels that might be missed by the current event-centric approach, potentially offering earlier indicators of health changes.²²

Chapter 5

Conclusion

AAL holds immense potential for supporting independent living and well-being, yet its widespread adoption has been hampered by significant challenges inherent in previous research paradigms. This thesis confronted some of these challenges head-on by designing an integrated AAL system designed explicitly to overcome these limitations through its core principles: edge-based processing, unsupervised learning, personalized modelling, built-in explainability, and validation via long-term deployment in a real-world home.

The successful execution of this study’s specific objectives confirmed the feasibility of this novel approach. The study demonstrated that a robust, privacy-preserving, edge-based sensor network can be effectively implemented for continuous, real-world data acquisition (Objective 1). Furthermore, the study showed that the rich data generated from this ecologically valid deployment can power unsupervised models capable of learning individual routines and detecting meaningful, personalized anomalies, coupled with interpretable explanations (Objective 2).

By integrating unsupervised methods, personalization, explainability, edge computing, and real-world validation, this research offers a tangible demonstration that effective unsupervised AAL systems that prioritize commercially affordable sensors can be developed and effectively learn to detect key activities and identify behavioural anomalies that may indicate key health indicators. It represents a concrete step towards realizing the potential of AAL technology to provide scalable, trustworthy, and genuinely supportive tools that enhance autonomy and facilitate proactive health awareness through non-intrusive technology. Specifically this research can contribute to future work where daily anomalies are presented in understandable ways via mobile applications or electronic screens around the house to help increase awareness of behavioural deviations. These anomalies can pave the way for the participants to confirm or deny each anomaly as accurate and important which the model then is able to take as feedback tailor the anomalies it presents. This thesis, therefore, not only addresses critical gaps in the literature but also contributes a validated blueprint for future AAL systems designed for real-world impact.

References

- [1] Guillermo Márquez and Carla Taramasco. Barriers and facilitators of ambient assisted living systems: A systematic literature review. *International Journal of Environmental Research and Public Health*, 20(6):5020, 2023.
- [2] Stephanie Blackman, Claudine Matlo, Charisse Bobrovitskiy, Ashley Waldoch, Mei Lan Fang, Piper Jackson, Alex Mihailidis, Louise Nygård, Arlene Astell, and Andrew Sixsmith. Ambient assisted living technologies for aging well: a scoping review. *Journal of Intelligent Systems*, 25(1):55–69, 2016.
- [3] Grazia Cicirelli, Roberto Marani, Antonio Petitti, Annalisa Milella, and Tiziana D’Orazio. Ambient assisted living: a review of technologies, methodologies and future perspectives for healthy aging of population. *Sensors*, 21(10):3549, 2021.
- [4] Gaya Bin Noon, Thokozani Hanjahanja-Phiri, Harishree Dave, Laura Fadrique, Jennifer Teague, Plinio P Morita, et al. Exploring the role of active assisted living in the continuum of care for older adults: thematic analysis. *JMIR aging*, 6(1):e40606, 2023.
- [5] Robert Stodczyk and Fischer-Hirchert Uhp. Ambient assisted living an overview of current applications, end-users and acceptance. *Biomedical Journal of Scientific & Technical Research*, 30(3):23374–23384, 2020.
- [6] Curtis S. Florence, Gwen Bergen, Adam Atherly, Elizabeth Burns, Judy Stevens, and Cynthia Drake. Medical costs of fatal and nonfatal falls in older adults. *Journal of the American Geriatrics Society*, 66(4):693–698, 2018.
- [7] Yan Wang, Xin Wang, Damla Arifoglu, Chenggang Lu, Abdelhamid Bouchachia, Yingrui Geng, and Ge Zheng. A survey on ambient sensor-based abnormal behaviour detection for elderly people in healthcare. *Electronics*, 12(7):1539, 2023.
- [8] Jiacheng Fang, Yuan Liu, Yuning An, and Kexin Zhou. The macroeconomic impact of demographic shifts: Aging populations and their socioeconomic consequences. *Law and Economy*, 2023.
- [9] World Health Organization. Ageing and health, 2024. Fact sheet, published 1 Oct 2024. Accessed from WHO Newsroom.

- [10] Hugo Saner, Narayan Schütz, Ashraf Botros, Prabitha Urwyler, Philipp Bulushek, Guillaume du Pasquier, and Tobias Nef. Potential of ambient sensor systems for early detection of health problems in older adults. *Frontiers in Cardiovascular Medicine*, 7:110, 2020.
- [11] Mir Nabila Ashraf, Natasha L. Gallant, and Cara Bradley. Digital Health Interventions in Older Adult Populations Living With Chronic Disease in High-Income Countries: Protocol for a Scoping Review. *JMIR Research Protocols*, 13(1):e49130, 2024.
- [12] Cother Hajat and Emma Stein. The global burden of multiple chronic conditions: a narrative review. *Preventive Medicine Reports*, 12:284–293, 2018.
- [13] Masud Chand and Rosalie L. Tung. The aging of the world’s population and its effects on global business. *Academy of Management Perspectives*, 28:409–429, 2014.
- [14] H. G. Kang, Diane Mahoney, H. Hoenig, V. Hirth, P. Bonato, I. Hajjar, and L. Lipsitz. In situ monitoring of health in older adults: Technologies and issues. *Journal of the American Geriatrics Society*, 58, 2010.
- [15] Bartosz Sawik, Sławomir Tobis, Ewa Baum, Aleksandra Suwalska, Sylwia Kropińska, Katarzyna Stachnik, Elena Pérez-Bernabeu, Marta Cildoz, Alba Agustin, and Katarzyna Wieczorowska-Tobis. Robots for elderly care: Review, multi-criteria optimization model and qualitative case study. *Health-care*, 11(9), 2023.
- [16] Andrea Monteriù, Alessandro Freddi, and Sauro Longhi. Special issue on advancements in ambient assisted living: integrating technology and human-centered design for enhancing user well-being and care. *Journal of Ambient Intelligence and Humanized Computing*, 15(5):2659–2660, 2024.
- [17] Marko Periša, Petra Teskera, Ivan Cvitić, and Ivan Grgurević. Empowering people with disabilities in smart homes using predictive informing. *Sensors*, 25(1):284, 2025.
- [18] World Health Organization and United Nations Children’s Fund. *Global report on assistive technology*. World Health Organization, 2022.
- [19] Inge Dhamanti, Ika M Nia, Krishnaraj Nagappan, and Balaji P Srikanth. Smart home healthcare for chronic disease management: A scoping review. *Digital Health*, 9:20552076231218144, 2023.
- [20] Taisa Kushner, Clara Mosquera-Lopez, Andrea Hildebrand, Michelle H Cameron, and Peter G Jacobs. Risky movement: Assessing fall risk in people with multiple sclerosis with wearable sensors and beacon-based smart-home monitoring. *Multiple sclerosis and related disorders*, 79:105019, 2023.
- [21] Mahsa Sheikh, Meha Qassem, and Panicos A Kyriacou. Wearable, environmental, and smartphone-based passive sensing for mental health monitoring. *Frontiers in digital health*, 3:662811, 2021.
- [22] Femke Koome, Clare Hocking, and Daniel Sutton. Why routines matter: The nature and meaning of family routines in the context of adolescent mental illness. *Journal of Occupational Science*, 19(4):312–325, 2012.

- [23] H. Aloulou, M. Mokhtari, T. Tiberghien, J. Biswas, C. Phua, J. Lin, and P. Yap. Deployment of assistive living technology in a nursing home environment: methods and lessons learned. *BMC Medical Informatics and Decision Making*, 13:42 – 42, 2013.
- [24] Bruna Maria Vittoria Guerra, Emanuele Torti, Elisa Marenzi, Micaela Schmid, Stefano Ramat, Francesco Leporati, and Giovanni Danese. Ambient assisted living for frail people through human activity recognition: state-of-the-art, challenges and future directions. *Frontiers in neuroscience*, 17:1256682, 2023.
- [25] Jens Lundström, Eric Järpe, and Antanas Verikas. Detecting and exploring deviating behaviour of smart home residents. *Expert Systems with Applications*, 55:429–440, 2016.
- [26] Matteo Ciprian, Matteo Gadaleta, and Michele Rossi. An unsupervised anomaly detection framework for smart assisted living via growing neural gas networks. *Journal of Ambient Intelligence and Smart Environments*, 16(3):365–387, 2024.
- [27] Oya Aran, Dairazalia Sanchez-Cortes, Minh-Tri Do, and Daniel Gatica-Perez. Anomaly detection in elderly daily behavior in ambient sensing environments. In *Human Behavior Understanding: 7th International Workshop, HBU 2016, Amsterdam, The Netherlands, October 16, 2016, Proceedings 7*, pages 51–67. Springer, 2016.
- [28] Fco Javier Ordóñez, Paula de Toledo, and Araceli Sanchis. Sensor-based bayesian detection of anomalous living patterns in a home setting. *Personal and Ubiquitous Computing*, 19:259–270, 2015.
- [29] Shruthi K Hiremath and Thomas Plötz. The lifespan of human activity recognition systems for smart homes. *Sensors*, 23(18):7729, 2023.
- [30] Jessamyn Dahmen, Diane J Cook, Xiaobo Wang, and Wang Honglei. Smart secure homes: a survey of smart home technologies that sense, assess, and respond to security threats. *Journal of reliable intelligent environments*, 3:83–98, 2017.
- [31] Luyi Sun, Bian Yang, Els Kindt, and Jingyi Chu. Privacy barriers in health monitoring: Scoping review. *JMIR nursing*, 7:e53592, 2024.
- [32] Zahraa Khais Shahid, Saguna Saguna, and Christer Åhlund. Detecting anomalies in daily activity routines of older persons in single resident smart homes: Proof-of-concept study. *JMIR aging*, 5(2):e28260, 2022.
- [33] Baraa Zieni, M. Ritchie, A. Mandalari, and F. Boem. An interdisciplinary overview on ambient assisted living systems for health monitoring at home: Trade-offs and challenges. *Sensors*, 2025.
- [34] Daniel Fährmann, Laura Martín, Luis Sánchez, and Naser Damer. Anomaly detection in smart environments: a comprehensive survey. *IEEE access*, 2024.
- [35] Merlin Kasirajan, MA Hannan Bin Azhar, and Scott Turner. Trustworthy insights: A novel multi-tier explainable framework for ambient assisted living. In *2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pages 2554–2561. IEEE, 2023.

- [36] Naina Bijlani, Ramin Nilforooshan, and Samaneh Kouchaki. An unsupervised data-driven anomaly detection approach for adverse health conditions in people living with dementia: Cohort study. *JMIR Aging*, 5(3):e38211, 2022.
- [37] Mladjan Jovanovic, Goran Mitrov, Eftim Zdravevski, Petre Lameski, Sara Colantonio, Martin Kampel, Hilda Tellioglu, and Francisco Florez-Revuelta. Ambient assisted living: scoping review of artificial intelligence models, domains, technology, and concerns. *Journal of Medical Internet Research*, 24(11):e36553, 2022.
- [38] Salisu Wada Yahaya, Ahmad Lotfi, and Mufti Mahmud. Detecting anomaly and its sources in activities of daily living. *SN Computer Science*, 2(1):14, 2021.
- [39] Geoff Dougherty. *Pattern recognition and classification: an introduction*. Springer Science & Business Media, 2012.
- [40] Gareth Halfacree. *The official Raspberry Pi Beginner’s Guide: How to use your new computer*. Raspberry Pi Press, 2023.
- [41] Domoticz. <https://www.domoticz.com/>, 2025. Open Source Home Automation.
- [42] Gollu Appala Naidu and Jayendra Kumar. Wireless protocols: Wi-fi son, bluetooth, zigbee, z-wave, and wi-fi. In *Innovations in Electronics and Communication Engineering: Proceedings of the 7th ICIECE 2018*, pages 229–239. Springer, 2019.
- [43] Influxdb. <https://www.influxdata.com/products/influxdb/>, 2025. Open Source Time Series Database.
- [44] Jiawei Han, Jian Pei, and Hanghang Tong. *Data mining: concepts and techniques*. Morgan kaufmann, 2022.
- [45] Ketan Rajshekhar Shahapure and Charles Nicholas. Cluster quality analysis using silhouette score. In *2020 IEEE 7th international conference on data science and advanced analytics (DSAA)*, pages 747–748. IEEE, 2020.
- [46] Ni-Bin Chang and Kaixu Bai. *Multisensor data fusion and machine learning for environmental remote sensing*. CRC Press, 2018.
- [47] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.
- [48] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.
- [49] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining*, pages 413–422. IEEE, 2008.
- [50] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

- [51] Peter McCullagh and John A. Nelder. *Generalized Linear Models*. Chapman and Hall, 2nd edition, 1989.
- [52] Marine Diniz, Vincent Zalc, and Dan Istrate. Cough sound recognition for covid-19 risk detection. In *JETSAN 2021-Colloque en Télésanté et dispositifs biomédicaux-8ème édition*, 2021.
- [53] Xuan Viet Cuong Nguyen, Hoai Son Nguyen, et al. Appliance classification on low-cost smart power outlets based on frequency feature of electric current. In *2021 26th IEEE Asia-Pacific Conference on Communications (APCC)*, pages 128–133. IEEE, 2021.
- [54] Mohammad Arif Ul Alam, Nirmalya Roy, Michelle Petruska, and Andrea Zemp. Smart-energy group anomaly based behavioral abnormality detection. In *2016 IEEE Wireless Health (WH)*, pages 1–8. IEEE, 2016.
- [55] Yue Zhang, Shijia Pan, Jonathon Fagert, Mostafa Mirshekari, Hae Young Noh, Pei Zhang, and Lin Zhang. Occupant activity level estimation using floor vibration. In *Proceedings of the 2018 ACM international joint conference and 2018 international symposium on pervasive and ubiquitous computing and wearable computers*, pages 1355–1363, 2018.
- [56] Fabio Salice, Marco De Vellis, Alessandro Gwyn Barbieri, Andrea Masciadri, and Sara Comai. Non-intrusive monitoring and detection of sitting and lying persons: A technological review. *IEEE Access*, 2024.
- [57] Grace H Coughlin, Maximilian T Antush, and Chantal A Vella. Associations of sedentary behavior and screen time with biomarkers of inflammation and insulin resistance. *Journal of Behavioral Medicine*, 47(5):828–838, 2024.
- [58] Lukas Sablica, Bettina Grün, Siamak Layeghy, Sara Dolnicar, and Marius Portmann. Ecoshower: Estimating shower duration using non-intrusive multi-modal sensor data via lstm and gated transformer models. *Expert Systems with Applications*, 277:127202, 2025.
- [59] David Marín-García, David Bienvenido-Huertas, Juan Moyano, Carlos Rubio-Bellido, and Carlos E Rodríguez-Jiménez. Detection of activities in bathrooms through deep learning and environmental data graphics images. *Heliyon*, 10(6), 2024.
- [60] Jérémie Saives, Clément Pianon, and Gregory Faraut. Activity discovery and detection of behavioral deviations of an inhabitant from binary sensors. *IEEE Transactions on Automation Science and Engineering*, 12(4):1211–1224, 2015.
- [61] Francisco Javier Parada Otte, Bruno Rosales Saurer, and Wilhelm Stork. Unsupervised learning in ambient assisted living for pattern and anomaly detection: a survey. In *Evolving Ambient Intelligence: AmI 2013 Workshops, Dublin, Ireland, December 3-5, 2013. Revised Selected Papers 4*, pages 44–53. Springer, 2013.
- [62] Yuxuan Shen, Zidong Wang, Hongli Dong, and Hongjian Liu. Multi-sensor multi-rate fusion estimation for networked systems: Advances and perspectives. *Information Fusion*, 82:19–27, 2022.

- [63] Amir Baniassadi, L. Lipsitz, Á. Pascual-Leone, B. Manor, and Wanting Yu. A platform to study the effects of home environment on health and wellbeing of older adults. *Innovation in Aging*, 5:965 – 965, 2021.
- [64] Julie Doyle, Andrea Kealy, John Loane, L. Walsh, B. O’Mullane, Carl Flynn, Andy MacFarlane, B. Bortz, R. Benjamin Knapp, and R. Bond. An integrated home-based self-management system to support the wellbeing of older adults. *J. Ambient Intell. Smart Environ.*, 6:359–383, 2014.
- [65] Julie Doyle, Niamh Caprani, Yusuke Komaba, and Akihiro Inomata. Healthcare professionals views on technology to support older adults transitioning from hospital to home. page 34, 2016.
- [66] Aftab Khan, Nils Y. Hammerla, Sebastian Mellor, and T. Plötz. Optimising sampling rates for accelerometer-based human activity recognition. *Pattern Recognit. Lett.*, 73:33–40, 2016.
- [67] Tatsuhito Hasegawa. Smartphone sensor-based human activity recognition robust to different sampling rates. *IEEE Sensors Journal*, 21:6930–6941, 2021.
- [68] Shaolin Weng, Luping Xiang, Weiwei Tang, Hui Yang, Lingxiang Zheng, Hai Lu, and Huiru Zheng. A low power and high accuracy mems sensor based activity recognition algorithm. *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 33–38, 2014.
- [69] Kenki Ueda, Hirohiko Suwa, Yutaka Arakawa, and Keiichi Yasumoto. Exploring accuracy-cost tradeoff in in-home living activity recognition based on power consumptions and user positions. In *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, pages 1130–1137, 2015.
- [70] World Health Organization. WHO Global Report on Falls Prevention in Older Age, 2008.
- [71] World Health Organization. Addressing the 18 million health worker shortfall – 35 concrete actions and 6 key messages, 2019. Departmental Update, 28 May 2019, World Health Organization.
- [72] Hannah Biermann, Julia Offermann-van Heek, Simon Himmel, and Martina Ziefle. Ambient assisted living as support for aging in place: Quantitative users’ acceptance study on ultrasonic whistles. *JMIR Aging*, 1(2):e11825, Dec 2018.
- [73] Mariagrazia Costanzo, Rossana Smeriglio, and Santo Di Nuovo. New technologies and assistive robotics for elderly: A review on psychological variables. *Archives of Gerontology and Geriatrics Plus*, page 100056, 2024.
- [74] Eva-Maria Schomakers and Martina Ziefle. Privacy perceptions in ambient assisted living. In *ICT4AWE*, pages 205–212, 2019.

APPENDICES

Appendix A: Approximate Shower Log

Table 1: Participant Logs of Recorded Showering Events

Date	Approximate Shower Time & Notes	Approx. Duration
2024-12-03	Morning around 7:40–7:50 AM. “Took a quick shower before work.”	~10 minutes
2024-12-03	Late night around 10:45–11:05 PM. “Showered before bed.”	~20 minutes
2024-12-04	Morning around 7:00–7:20 AM. “Quick shower to start the day.”	~15–20 minutes
2024-12-04–05	Late night around 11:50 PM–12:10 AM. “Rinsed off before sleeping.”	~20 minutes
2024-12-06	Early afternoon around 1:30–1:40 PM. “Showered after lunch.”	~10 minutes
2024-12-08	Night around 10:05–10:15 PM. “Quick rinse before bed.”	~10 minutes
2024-12-09	Night around 8:05–8:25 PM. “Showered after dinner.”	~20 minutes
2024-12-11	Morning around 8:10–8:25 AM. “Quick morning routine.”	~15 minutes
2024-12-11	Late night around 10:35–10:55 PM. “Felt sweaty, showered before sleep.”	~20 minutes
2024-12-12	Night around 9:50–10:05 PM. “Quick end-of-day shower.”	~15 minutes
2024-12-13	Late morning around 11:55 AM–12:20 PM. “Longer midday shower.”	~25 minutes
2024-12-15	Early afternoon around 12:50–1:15 PM. “Showered after brunch.”	~25 minutes

Continued on next page

Table 1 – *Continued from previous page*

Date	Approximate Shower Time & Notes	Approx. Duration
2024-12-16	Night around 9:30–9:50 PM. “Quick rinse-off at night.”	~15 minutes
2024-12-18	Late night around 10:20–10:50 PM. “Needed a long shower.”	~25–30 minutes
2024-12-20	Morning around 7:05–7:25 AM. “Showered before heading out.”	~20 minutes
2025-01-01–02	Late night around 11:45 PM–12:05 AM. “New Year’s shower!”	~20 minutes
2025-01-02	Night around 8:55–9:20 PM. “Shower before winding down.”	~25 minutes
2025-01-03	Night around 9:05–9:25 PM. “Rinsed off at night.”	~20 minutes
2025-01-04	Noon around 12:00–12:05 PM. “Very quick midday shower.”	~5–10 minutes
2025-01-05	Afternoon around 4:35–5:05 PM. “Showered after being out.”	~30 minutes
2025-01-06	Night around 9:35–10:00 PM. “Wanted a relaxing shower.”	~25 minutes
2025-01-07	Night around 9:50–10:20 PM. “Took my time in the shower.”	~25–30 minutes
2025-01-08	Night around 10:15–10:40 PM. “Longer night shower.”	~25 minutes
2025-01-10	Very early morning around 1:45–2:00 AM. “Late shower after a long day.”	~15 minutes
2025-01-11	Late morning around 10:40–11:00 AM. “Post-breakfast shower.”	~20 minutes
2025-01-19	Night around 10:10–10:30 PM. “Felt gross, needed a shower.”	~20 minutes
2025-01-20	Early morning around 6:25–6:45 AM. “Morning wake-up shower.”	~20 minutes
2025-01-21	Morning around 7:40–8:00 AM. “Got ready for the day.”	~20 minutes
2025-01-27	Late night around 11:20–11:50 PM. “Long shower before bed.”	~30 minutes
2025-01-28	Early afternoon around 12:45–12:55 PM. “Quick midday rinse.”	~10 minutes
2025-01-29	Night around 9:20–9:25 PM. “Super quick shower.”	~5–10 minutes
2025-01-30	Night around 9:00–9:20 PM. “Routine evening shower.”	~20 minutes
2025-01-31	Morning around 9:25–9:45 AM. “Post-workout shower.”	~20 minutes
2025-02-05	Late morning around 10:20–10:45 AM. “Refreshing shower.”	~25 minutes

Continued on next page

Table 1 – *Continued from previous page*

Date	Approximate Shower Time & Notes	Approx. Duration
2025-02-06	Evening around 6:50–7:10 PM. “Quick shower after work.”	~20 minutes
2025-02-07	Night around 8:55–9:15 PM. “Short shower to wind down.”	~20 minutes
2025-02-08	Late morning around 11:10–11:35 AM. “Took a bit longer.”	~25 minutes
2025-02-08–09	Late night around 11:40–11:50 PM. “Quick rinse.”	~10 minutes
2025-02-09	Night around 7:50–8:10 PM. “Needed a warm shower.”	~15–20 minutes
2025-02-10	Late night around 10:50–11:10 PM. “Showered before bedtime.”	~20 minutes
2025-02-11	Night around 10:15–10:35 PM. “Another night shower.”	~20 minutes
2025-02-13	Night around 9:55–10:10 PM. “Quick late evening shower.”	~15 minutes
2025-02-14	Morning around 9:00–9:10 AM. “Fast shower before leaving.”	~10–15 minutes
2025-02-18	Night around 9:10–9:15 PM. “Very quick rinse.”	~5–10 minutes
2025-02-19	Late night around 10:45–11:05 PM. “Wanted a warm shower.”	~15–20 minutes
2025-02-20	Morning around 8:00–8:15 AM. “Morning shower, short and sweet.”	~15–20 minutes
2025-02-23	Night around 9:40–9:45 PM. “Really fast rinse.”	~5 minutes
2025-02-24	Morning around 7:20–7:35 AM. “Woke up and showered.”	~15 minutes
2025-02-24	Night around 10:15–10:30 PM. “Needed a post-work shower.”	~20 minutes
2025-02-25	Morning around 8:20–8:30 AM. “Quick rinse before heading out.”	~10 minutes
2025-02-27	Late night around 12:10–12:20 AM. “Short late-night shower.”	~10 minutes
2025-03-02	Late night around 11:45–11:55 PM. “Quick rinse before sleep.”	~10 minutes
2025-03-05	Morning around 7:50–8:05 AM. “Took a decent morning shower.”	~15 minutes
2025-03-10	Morning around 7:30–7:40 AM. “Short morning shower.”	~10 minutes
2025-03-10	Night around 10:40–10:50 PM. “Tired, quick rinse.”	~10 minutes
2025-03-11	Early afternoon around 1:10–1:20 PM. “Lunch break shower.”	~10 minutes
2025-03-11	Night around 9:00–9:25 PM. “Wound down with a longer shower.”	~20–25 minutes

Continued on next page

Table 1 – *Continued from previous page*

Date	Approximate Shower Time & Notes	Approx. Duration
2025-03-14	Morning around 8:30–8:40 AM. “Quick shower to start the day.”	~10 minutes
2025-03-16	Night around 10:20–10:30 PM. “Rinsed off after a long day.”	~10–15 minutes
2025-03-18	Showered late night at 10–10:30 PM. “I approximated about 20 minutes.”	~20 minutes