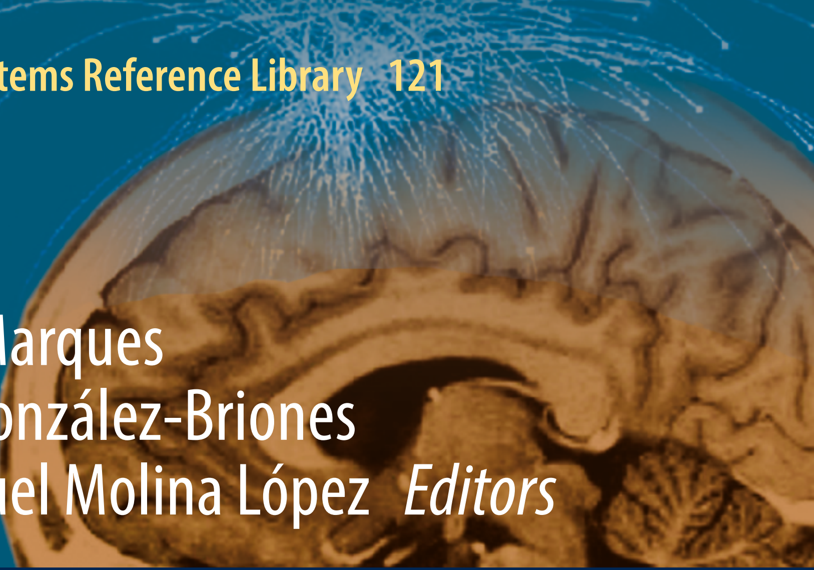
Intelligent Systems Reference Library 121



Gonçalo Marques
Alfonso González-Briones
José Manuel Molina López *Editors*

Machine Learning for Smart Environments/Cities

An IoT Approach

Springer

Intelligent Systems Reference Library

Volume 121

The aim of this series is to publish a Reference Library, including novel advances and developments in all aspects of Intelligent Systems in an easily accessible and well structured form. The series includes reference works, handbooks, compendia, textbooks, well-structured monographs, dictionaries, and encyclopedias. It contains well integrated knowledge and current information in the field of Intelligent Systems. The series covers the theory, applications, and design methods of Intelligent Systems. Virtually all disciplines such as engineering, computer science, avionics, business, e-commerce, environment, healthcare, physics and life science are included. The list of topics spans all the areas of modern intelligent systems such as: Ambient intelligence, Computational intelligence, Social intelligence, Computational neuroscience, Artificial life, Virtual society, Cognitive systems, DNA and immunity-based systems, e-Learning and teaching, Human-centred computing and Machine ethics, Intelligent control, Intelligent data analysis, Knowledge-based paradigms, Knowledge management, Intelligent agents, Intelligent decision making, Intelligent network security, Interactive entertainment, Learning paradigms, Recommender systems, Robotics and Mechatronics including human-machine teaming, Self-organizing and adaptive systems, Soft computing including Neural systems, Fuzzy systems, Evolutionary computing and the Fusion of these paradigms, Perception and Vision, Web intelligence and Multimedia.

Indexed by SCOPUS, DBLP, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at https://link.springer.com/bookseries/8578

Gonçalo Marques · Alfonso González-Briones ·
José Manuel Molina López
Editors

Machine Learning for Smart Environments/Cities

An IoT Approach

Springer

Editors
Gonçalo Marques
Polytechnic of Coimbra
ESTGOH, Rua General Santos Costa
Oliveira do Hospital, Portugal

José Manuel Molina López
Universidad Carlos III DE Madrid
Colmenarejo, Spain

Alfonso González-Briones
Grupo de Investigación BISITE
Departamento de Informática y Automática
Facultad de Ciencias
Instituto de Investigación Biomédica de Salamanca
Universidad de Salamanca
Salamanca, Spain

ISSN 1868-4394 ISSN 1868-4408 (electronic)
Intelligent Systems Reference Library
ISBN 978-3-030-97515-9 ISBN 978-3-030-97516-6 (eBook)
https://doi.org/10.1007/978-3-030-97516-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The Internet of Things (IoT) consists of numerous "things" that are connected and managed across the Internet. Machine Learning is a research field that focuses on the development of predictive computer-aided applications which are not explicitly programmed. IoT and Machine Learning are mutually contributing to the advancement of technologies involved in smart city and smart environment development. Furthermore, the use of such communication and computing technologies is enabled by miniaturized microcontrollers, sensors, and actuators, which have a lower cost and greater energy efficiency.

Thus, IoT and ML provide smart cities and smart environments with countless advantages by enabling interaction with the physical world through interoperable applications, supported by smart city and smart environment systems. These applications aim to improve the citizens' quality of life.

This book aims to introduce Machine Learning and its applications for smart environments/cities supported by IoT technologies. Its chapters have been written by researchers located in different countries across the globe. The book consists of 12 chapters. Chapter 1 presents an introduction to the book scope and main topics. The rest of the chapters are organized into two parts. On the one hand, Part I comprises Chaps. 2–6 and focuses on smart environments. On the other hand, Part II comprises Chaps. 7–12 and presents relevant studies on smart cities. These chapters have been contributed by several authors from across the globe, namely Australia, Brazil, Colombia, Germany, Ghana, Italy, Malaysia, Portugal, Spain, Turkey, the Netherlands, and the USA. This makes the content of the book geographically diverse, as its authors come from 12 different countries, spread over 6 different continents.

The advent of the IoT and Machine Learning in recent decades has brought about developments in smart sensing and actuating technologies, which have been adopted in so-called smart environments, such as smart homes, smart farms, and other smart city settings. In this regard, keeping track of applications that use IoT and Machine Learning for smart environments has become an important aspect of research. A lot of research effort has been put into reviewing aspects of smart environments/cities, such as technologies, architectures, and security. However, there is not enough research on approaches that would combine IoT and Machine Learning in smart environments/cities. Chapter 1 "**An Introduction and Systematic Review on Machine Learning for Smart Environments/Cities: An IoT Approach**" presents a systematic review of the combination of IoT and Machine Learning in smart environments/cities. Furthermore, recommendations for the implementation of IoT and Machine Learning in smart environments/cities are presented. It is expected that the recommendations may be used as a basis for the successful implementation of IoT and Machine Learning in smart environments/cities.

Chapter 2 "**Model-Based Digital Threads for Socio-Technical Systems**" defines the MBSE-based methodology for the creation and maintenance of the digital thread of physical systems and their digital twins. The authors illustrate this with a case study in which the proposed methodology has been applied to design a digital thread of a Traffic Monitoring System (TMS) for a smart city. The methodology uses the SysML (Systems Modeling Language), which is adequate for the specification of socio-technological (cyber-physical) systems, such as TMSs. This digital thread represents both the physical and virtual entities of the system, enabling the development of digital twins for simulating, testing, monitoring, and/or maintaining the system. SysML is currently being redesigned, and the new SysML v2 aims to offer precise and expressive language capabilities to improve support to system specification, analysis, design, verification, and validation. The chapter also discusses how SysML v2 is expected to facilitate the development of digital threads for socio-technological systems, such as TMSs.

Since the beginning, smart cities have been predicated on the concept of IoT; however, there is a gap between theory and practice, making effective implementation challenging. Smart cities are not just for urban dwellers; they are also for suburban dwellers who need effective methods and systems to foster a higher quality of life. The principal pollutants that are the subject of this study are industrial suburbs. Chapter 3 "**IoT Regulated Water Quality Prediction Through Machine Learning for Smart Environments**" presents research on data from recycled wastewater obtained from industrial use, which would otherwise be dumped directly into rivers. The authors used IoT sensors to gather data which has enabled the inspection of water and the maintenance of its quality. The data were used to measure a series of factors that indicate and influence water quality; they were utilized to calculate the water quality index. Finally, three Machine Learning algorithms were used to train and predict this quality index.

Chapter 4 "**The Power of Augmented Reality for Smart Environments: An Explorative Analysis of the Business Process Management**" contains an explorative analysis of the use of augmented reality for smart environments. It demonstrates that augmented reality improves business process management in numerous fields, such as military, medicine, architecture, automotive, and retail. Additionally, the study furnishes historical evidence for the evolution of augmented reality for smart environments, reviews the contributions of top research in this field, and evidences the effect of augmented reality on the business process and on the transformation of the business environment into a smart environment through IoT devices. This type of devices is capable of monitoring important variables in the users' environments, from a value co-creation perspective. Indeed, there is evidence of the use of augmented reality in Information Management, Planning and Control, Change Process, Knowledge Management, Performance Management, People, and Customer Management. Finally, the chapter reveals how this technology can support innovative practices in business processes. In the search for competitive advantage, firm managers could indeed exploit this research to explore the impact of smart components on the improvement of smart environments investigating new industries.

IoT technologies are an opportunity for humanity to provide a wide range of e-educational technologies that have the potential to change educational systems. For example, students can now access online labs and libraries on their smart devices, which increases productivity in learning activities. Thanks to the rapid growth of information and communication technologies, online learning is now a practical choice, provided there is Internet availability. By combining real and virtual aspects in the learning process, IoT will allow for the expansion of learning contexts. The adoption of IoT can provide a wide range of e-educational technologies that have the potential to cover the shortfalls of the current education system. Through IoT, students can interact with their colleagues and teachers, share ideas, and find answers to problems. Chapter 5 "**Internet of Things Applications for Smart Environments**" considers IoT in education, the benefits of IoT technology in e-learning, online self-study, smart collaboration, IoT and e-learning, smart homes, smart home service adoption, and critical factors for smart home service.

A smart environment should be self-aware, using IoT technologies applied to health, education, and justice. Following the COVID-19 pandemic, the gains from applying technologies in healthcare and medical decision-making environments became evident. Chapter 6 "**Exploring Interpretable Machine Learning Methods and Biomarkers to Classifying Occupational Stress of the Health Workers**" focuses on the monitoring of stress in healthcare professionals through wearable devices, using biomarkers and Machine Learning to develop models that can aid in decision making. Challenges related to Explainable Artificial Intelligence are also addressed, as well as to the definition of stress classification, enabling the identification of impact on health professionals. An intelligent system is proposed to recommend actions in response to the professionals' stress level, in a way that is explainable, transparent, and feasible. This is an outstanding solution that may be adopted by the managers of health centers. Challenges related to information security and to the privacy of the health professionals are also discussed.

Chapter 7, "**Smart Cities, The Internet of Things, and Corporate Social Responsibility**" explores a plethora of IoT studies to identify how such technologies improve operational efficiency and infrastructure service and create an ecosystem in which the economic, environmental, and societal challenges associated with increased urbanization and smart cities may be addressed. Their inherent risks, issues, and challenges are also explored. Building on CSR literature, the author argues for a re-orientation of the smart city design toward decisional and governance process(es), and a shift away from technocentric and top-down approaches. A call is made for increased collaboration between decision-makers, community, and citizens in IoT implementation. A top-down/bottom-up multi-staged collaborative approach is proposed for evolving Corporate Social Responsibility governance and engagement. It recognizes the importance of creating shared value, in the selection and deployment of IoT devices. Consequentially, addressing and resolving the challenges faced by communities and citizens in the adoption of IoT in smart cities.

Chapter 8 "**Intelligent Techniques for Optimization, Modelling and Control of Power Management Systems Efficiency**" analyzes the issue of climate change and smart grid, from the point of view of efficiency in power converters. Power distribution in the Spanish electric system has changed over the years and will further change with the introduction of the smart grid, opening up new possibilities for the distribution of energy. Power electronics aim to interconnect the different parts of the electric system and control the energy flow from point to point. In addition, the main objective of smart cities is efficiency and the optimal use of energy. Power converters are a crucial element, as they interconnect electricity generators with electricity consumers. The introduction of Machine Learning and AI in this field helps to optimize the switching converter, allowing for the reduction of power loss.

Chapter 9 "**Intelligent Simulation and Emulation Platform for Energy Management in Buildings and Microgrids**" presents a Multi-Agent-based Real-Time Infrastructure for Energy (MARTINE). MARTINE is a platform that enables the study of the physical components of buildings and microgrids, including emulation capabilities, multi-agent and real-time simulation, and intelligent decision support models and services, based on Machine Learning approaches. MARTINE enables the study and management of energy resources, considering both physical and intelligent virtual components. Hence, it provides a real platform for the continuous improvement of the synergies between IoT and Machine Learning solutions.

Smart cities collect data using IoT technology and use the information obtained from this data to manage resources and services efficiently. Thus, the living conditions of the people living in the cities are facilitated by the quality services offered to them. Chapter 10 "**Machine Learning Applications and Security Analysis in Smart Cities**" presents smart city IoT applications; especially, smart parks, smart buildings, smart homes, smart health, smart business, and smart environment applications are widely used. It is possible to benefit from Machine Learning methods depending on the data obtained while developing these applications. With these methods, routine operations, especially for the city administration, can be made more practical and rational. However, while everything is smart, environmental influences should be taken into account. Systems that can be implemented in smart cities with the smart environment should be completely sensitive to the environment. The protocols used for the communication of data of IoT applications developed in smart environments and cities must be secure in terms of security.

Chapter 11 "**Recent Developments of Deep Learning in Future Smart Cities: A Review**" reviews the use of Deep Learning techniques in Artificial Intelligence applications, oriented toward multiple smart city domains, including smart transportation, smart services, smart governance, environment, security, and public safety. The difficulties associated using Deep Learning on smart city data have also been addressed.

This chapter is concluded by describing a series of Deep Learning techniques that help to better understand the smart city concept and follow the current trends in smart cities.

Chapter 12 "**Smart and Sustainable Cities in Collaboration with *IoT*: The Singapore Success Case**" explores the collaboration of the IoT paradigm with the Sustainable and Smart Cities (SSC) concept and looks at their success cases. The need for an adequate flow of information is emphasized, so that the state of a particular urban area may be known in real time. Accordingly, the challenges in the interconnection of highly sensitive sensors, as well as the transfer of data, require a hybrid cloud architecture that would allow for the large-scale processing of daily citizen data and for the prediction of environmental factors. However, the conceptualization and creation of an SSC must be considered in technological, scientific, social, and state policies, aspects that translate into Governance, Mobility, Sustainability, Economic Development, Intellectual Capital, and Quality of Life. Moreover, adding to the technological utopia, the modern concept of economic and social development entails the creation of an SSC for the promotion of entrepreneurship, innovation, and social justice: a new dimension of urban resilience focused on a city caring society.

Recent advances and comprehensive reviews have been included, aiming to provide background for future research initiatives. It is hoped that the book will be of support in the research and development of future IoT architectures for smart environments/cities. Finally, we thank everyone involved in this project for their contribution and for giving us the opportunity to edit this book. Furthermore, we would like to thank all the professionals from Springer who have worked with us, in particular, Prof. Lakhmi C. Jain, for their help and support during the development of this book.

Oliveira do Hospital, Portugal — Gonçalo Marques
Salamanca, Spain — Alfonso González-Briones
Colmenarejo, Spain — José Manuel Molina López

About This Book

This book introduces Machine Learning and its applications in smart environments/cities. At this stage, a comprehensive understanding of smart environment/city applications is critical for supporting future research. This volume includes chapters written by researchers from different countries across the globe and identifies critical threads in research and also gaps that open up new and challenging lines of research for the future. Recent advances are discussed, and thorough reviews introduce readers to critical domains. The discussion on key research topics presented in this book will accelerate smart city and smart environment implementations based on IoT technologies. Consequently, this book will support future research activities aimed at developing future IoT architectures for smart environments/cities.

Contents

About the Editors

Gonçalo Marques holds a Ph.D. in Computer Science Engineering and is Member of the Portuguese Engineering Association (Ordem dos Engenheiros). He is currently working as Assistant Professor lecturing courses on programming, multimedia, and database systems. Furthermore, he worked as Software Engineer in the Innovation and Development Unit of Groupe PSA automotive industry from 2016 to 2017 and in the IBM group from 2018 to 2019. His current research interests include Internet of Things, Enhanced Living Environments, Machine Learning, e-health, telemedicine, medical and healthcare systems, indoor air quality monitoring and assessment, and wireless sensor networks. He has more than 80 publications in international journals and conferences, is a frequent reviewer of journals and international conferences, and is also involved in several edited books projects.

Alfonso González-Briones holds a Ph.D. in Computer Engineering from the University of Salamanca since 2018; his thesis obtained the second place in the 1st SENSORS+CIRTI Award for the best national thesis in Smart Cities (CAEPIA 2018). At the same university, he obtained his Bachelor of Technical Engineer in Computer Engineering (2012), Degree in Computer Engineering (2013), and Master in Intelligent Systems (2014). He was Project Manager of Industry 4.0 and IoT projects in the AIR Institute, Lecturer at the International University of La Rioja (UNIR), and also a "Juan De La Cierva" Postdoc at University Complutense of Madrid. Currently, he is Assistant

Professor at the University of Salamanca in the Department of Computer Science and Automatics. He has published more than 30 articles in journals, more than 60 articles in books and international congresses and has participated in 10 international research projects. He is also Member of the scientific committee of the *Advances in Distributed Computing and Artificial Intelligence Journal* (ADCAIJ) and *British Journal of Applied Science and Technology* (BJAST), and a reviewer of international journals (*Supercomputing Journal*, *Journal of King Saud University*, *Energies, Sensors, Electronics or Applied Sciences* among others). He has participated as Chair and Member of the technical committee of prestigious international congresses (AIPES, HAIS, FODERTICS, PAAMS, KDIR).

José Manuel Molina López is Full Professor at the Universidad Carlos III de Madrid. He joined the Computer Science Department of the Universidad Carlos III de Madrid in 1993. Currently, he coordinates the Applied Artificial Intelligence Group (GIAA). His current research focuses on the application of soft computing techniques (NN, Evolutionary Computation, Fuzzy Logic, and Multi-agent Systems) to radar data processing, air traffic management, e-commerce, and ambient intelligence. He has authored up to 100 journal papers and 200 conference papers. He received a degree in Telecommunications Engineering in 1993 and a Ph.D. degree in 1997 both from the Universidad Politécnica de Madrid.

Chapter 1
An Introduction and Systematic Review on Machine Learning for Smart Environments/Cities: An IoT Approach

José Joaquín Peralta Abadía and Kay Smarsly

Abstract Over the last centuries, human activities have had a significant impact on the environment, usually damaging and exploiting land, water bodies, and air all around us. With the advent of the Internet of Things (IoT) and machine learning (ML) in recent decades, developments in smart sensing and actuating technologies have been adopted for the environment. As such, interactions between humans and the environment have become more synergetic and efficient, creating so-called "smart environments". In recent years, smart environments, such as smart homes, smart farms and smart cities, have matured at an increasing rate. Therefore, keeping track of applications for smart environments has become an important aspect of research. Although several research efforts have targeted reviewing aspects of smart environments, such as technologies, architectures, and security, a gap is identified. Reviews focusing on approaches using a combination of IoT and ML in smart environments/cities are lacking. In this chapter, a systematic review of the combination of IoT and ML in smart environments is presented. Moreover, a summary of approaches to combine IoT and ML in smart environments is provided. The findings achieved in this chapter materialize into recommendations for the implementation of IoT and ML in smart environments. It is expected that the recommendations may be used as a basis for successful implementations of IoT and ML in smart environments.

Keywords Smart cities · Internet of Things (IoT) · Machine learning · IoT applications · Smart environments

1.1 Introduction

Human activities have had an impact on the environment over the last centuries, damaging and exploiting land, water bodies, and the air around us. Urban areas cover around 2% of the surface of the earth while accounting for approximately 75%

J. J. Peralta Abadía (✉) · K. Smarsly
Institute of Digital and Autonomous Construction, Hamburg University of Technology, Hamburg, Germany
e-mail: joaquin.peralta@tuhh.de

G. Marques et al. (eds.), *Machine Learning for Smart Environments/Cities*, Intelligent Systems Reference Library 121,
https://doi.org/10.1007/978-3-030-97516-6_1

of the world's natural resources [1]. However, recent developments in smart sensing and actuating technologies, such as Internet of Things (IoT) and machine learning (ML) algorithms, are key enablers of monitoring and control activities, sensing and acting upon the environment. Interactions between humans and the environment have become more synergetic and efficient, creating so-called "smart environments". Smart environments have the ability to obtain knowledge from the surroundings and to act and adapt according to the needs of the inhabitants, improving the experience of living beings in the smart environments [2]. As such, smart environments, such as smart homes, smart farms and smart cities, have been developing at increasing growth rates in recent years. Therefore, keeping track of applications for smart environments has become an important aspect of research.

Several surveys have covered the usage of IoT in smart environments, focusing on different aspects and fields. Security aspects of smart environments, such as intrusion detection systems, have been reviewed, providing insights into security vulnerabilities in IoT architectures for smart environments [3]. Information and communication technology (ICT) aspects of smart environments have also been thoroughly reviewed, covering networking technologies and standards [2], platforms and frameworks [4], communication technologies and architectures [5], and the integration of augmented reality with IoT [6]. Other reviews have focused on specific fields of smart environments, such as the adoption of smart devices in waste management in smart cities [7]. Nevertheless, ML algorithms, representing a driving technology of smart environments, have not been the focus of reviews in the field of IoT and smart environments. Thus, a review focusing on approaches that combine IoT and ML, representing innovative cornerstones of smart environments, is lacking. In this chapter, research related to the combination of IoT and ML in smart environments/cities is reviewed. Approaches used to combine both technologies are summarized and compared, revealing insights into aspects of the implementation and efficiency.

The remainder of the chapter is structured as follows. Section 1.2 presents concepts related to smart environments and machine learning, as a basis for the review process. Section 1.3 summarizes and compares the results of the review of approaches that combine IoT and ML in smart environments. Section 1.4 discusses the results of the review and the implications on future research. Finally, Sect. 1.5 provides conclusions and potential future work obtained from the results of this chapter.

1.2 Smart Environments, Internet of Things, and Machine Learning Concepts

This section introduces the basic concepts necessary to perform the systematic review of the combination of IoT and ML in smart environments. First, smart environment concepts are elaborated. Thereupon, concepts of IoT and ML are presented.

1.2.1 *Smart Environments*

Smart environments are the link between computers and everyday settings and tasks, powered by recent advances in ML, IoT, and pervasive computing [8]. The four main features of smart environments are (i) remote control of devices, (ii) device communication, (iii) information acquisition and dissemination from intelligent sensor networks, and (iv) enhanced services provided by intelligent devices. Remote control of devices is the most basic feature of smart environments, freeing humans from the necessity of physically interacting with devices. Subsequently, device communication allows machine-to-machine (M2M) communication, using standardized protocols and retrieval of information from Internet sources, thus building informed models of smart environments. Following, information acquisition and dissemination from intelligent sensor networks provide the ability to perform automated adjustments in smart environments, based on sensor readings and device communication. Finally, enhanced services provided by intelligent devices offer advanced capabilities, such as washing machines equipped with smart sensors that determine appropriate washing cycle times.

Smart environments can be subdivided according to the application domains. The main application domains include smart cities, smart homes, smart buildings, smart health, smart grids, smart transportation, and smart industries (also referred to as "smart factories") [2]. The application domains are characterized by varying characteristics, such as personal or business use and single-user or multi-user oriented [5]. Therefore, the aforementioned application domains will be taken during the review process to categorize studies involving the combination of IoT and ML in smart environments.

1.2.2 *Internet of Things*

The IoT is a paradigm in which IoT devices, denoted also as "things", are interconnected in a worldwide network. As such, IoT is powered by the use of heterogeneous IoT devices, enabling the design of applications that involve virtually both humans and IoT devices [9]. A scalable, layered IoT architecture is necessary to connect the physical and the digital world, as IoT applications may involve a large number of heterogeneous IoT devices [10]. Figure 1.1 shows an example of an IoT architecture, based on the architecture proposed by Guth et al. [11], with an additional vertical layer of security, contemplating the recent needs in data privacy and security. The application layer serves the visualization and analysis of data as well as the control of the IoT devices of the sensing layer. The IoT integration middleware layer comprises data storage, business logic, and the definition of networking and communication protocols, as well additional services, such as alerts and machine learning algorithms. The sensing layer includes IoT edge gateways, IoT devices, and sensors and actuators, which interact with the real world.

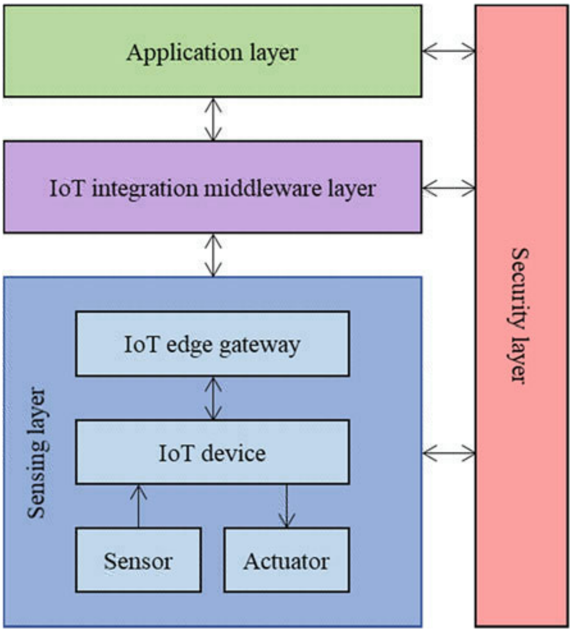


Fig. 1.1 IoT architecture with an additional security layer

1.2.3 Machine Learning

Machine learning is a subfield of computer sciences and a branch of artificial intelligence. Essentially, ML studies computer algorithms that learn and improve by using data, gaining experience, and generalizing knowledge [12]. ML algorithms observe data from datasets, build ML models based on the data and use the ML models to formulate hypotheses aiming to solve problems and to automate tasks [13]. Furthermore, ML algorithms, according to [13], may be categorized into three categories: (i) supervised learning, (ii) unsupervised learning, and (iii) reinforcement learning.

In *supervised learning* algorithms, datasets used to train ML models require a target feature, which can be labels in classification problems or numerical values in regression problems. In *unsupervised learning* algorithms, datasets used to train ML models lack a target feature and, as such, the ML model attempts to identify patterns and relationship between elements of the datasets (also referred to as datapoints). In *reinforcement learning* algorithms, outputs of the ML model are evaluated according to scoring functions and datasets are continuously updated with new datapoints, thus improving the knowledge of the algorithm through an iterative process. Finally, some ML algorithms combine supervised and unsupervised learning algorithms, called *semi-supervised learning* algorithms, and are capable of improving the performance of ML models in cases where only few datapoints are labeled. Consequently, in semi-supervised learning, only a small number of datapoints are labeled and a large number of datapoints are unlabeled.

Having presented the basic concepts of smart environments, IoT and ML, the following section presents the results of the review of the combination of IoT and ML in smart environments.

1.3 Results

This section reviews the combination of IoT and ML in smart environments. First, the methodology pursued during the review process is presented. Second, the quantity of studies reviewed herein, and the mean citation count are summarized for each smart environment domain. Finally, the review of the combination of IoT and ML in smart environments is detailed by domain, summarizing, and comparing the ML algorithms used in the studies, the IoT layer in which the ML algorithms are implemented, and the problem solved by the ML algorithms.

The review methodology follows three steps, (i) data collection, (ii) data organization, and (iii) data analysis. The data collection step involves searching for journal papers indexed in the Web of Science Core Collection as well as journal papers and conference papers indexed in the Scopus database. In an initial search, 1020 indexed studies, published between 2016 and 2021 and involving IoT, ML and smart environments, have been found, using the search string ((“smart environment*” OR “smart cit*” OR “smart home*” OR “smart grid*” OR “smart building*” OR “smart transportation” OR “smart health” OR “smart industr*” OR “smart factor*”) AND “machine learning” AND “IoT”). Next, the 100 most representative papers have been chosen, based on citation count. However, 29 papers lacked description of the ML algorithms used in the studies or presented the terms in the abstract but did not dwell in the topic in the full text. Consequently, the 29 studies have been omitted, entailing 71 studies to review. Then, in the data organization step, the IoT and ML characteristics described in the studies have been tabulated, according to the concepts presented in Sect. 1.2. Finally, an analysis of the organized data has been carried out, as presented in the remainder of this section.

Figure 1.2 presents the predominant terms in the abstracts of the studies in the form of a word cloud, where common multi-word terms in ICT have been joined as a single word. The primary predominant terms are “Internet of Things”, “data”, and “smart”, which highly relate to the topic being reviewed. Terms related to the field of machine learning follow in predominance, where “deep learning”, “machine learning”, “detection” and “model” have secondary predominance.

The papers reviewed in this study are grouped by domain, and the mean citation count is presented in Table 1.1, ordered by study count. It may be observed that smart cities and smart homes are the domains targeted most in smart environments, each with 18 studies, and the smart industries domain, with 2 studies, is the least targeted domain. Regarding the mean citation count, the most cited studies, with a mean of more than 40 citations, belong to the smart cities and smart grids domains, corroborating the interest of research and industry in both domains. Furthermore, smart transportation and smart homes follow, having a mean of more than 30 citations. The

Fig. 1.2 Predominant terms in the abstracts of the papers reviewed in this study

Table 1.1 Quantity of studies and mean citation count for each smart environment domain

Domain	Study count	Mean citations
Smart cities	18	44.6
Smart homes	18	31.6
Smart buildings	10	17.5
Smart health	9	26.1
Smart grids	7	40.4
Smart transportation	7	35.9
Smart industries	2	19.0

least cited studies belong to the smart buildings and smart industries domains. Having presented a general summary of the studies by domain, the following subsections present, by domain, a quantitative summary and comparison of the ML algorithms used in the studies, the IoT layer in which the ML algorithms are implemented, and the problem solved by the ML algorithms.

1.3.1 Smart Cities

Smart cities represent the broadest smart environment domain, encompassing the other smart environment domains. Therefore, studies involving the combination of IoT and ML in smart environments – with a broad covering scope, such as network intrusion detection and environmental monitoring – will mainly fall in this domain. Table 1.2 presents the studies targeting the smart cities domain grouped by machine learning type, data processing task, and algorithm. It may be noticed that supervised classification tasks have been preferred, where neural network algorithms and support vector machine algorithms are the most used ML algorithms, as the majority of

Table 1.2 Studies targeting the smart cities domain are grouped by machine learning type, data processing task, and algorithm

Learning type	Data processing task	Algorithm	Studies
Supervised	Classification	Neural networks	[14–19]
		Support vector machines	[14, 16, 20–25]
		Random forest	[20, 26]
		Decision tree	[16, 20]
	Regression	Neural networks	[27]
		Random forest	[27]
		Decision tree	[27, 28]
		Gradient boosting	[27]
		Logistic regression	[29]
		ARIMA	[30]
		Support vector regression	[31]
		Linear regression	[31]
Unsupervised	Clustering	DBSCAN	[23]
		k-means	[29]
	Feature selection	PCA	[17, 28]
		Auto-encoder	[16]

tasks involve labeling data, such as intrusion detection, object tracking, and user authentication.

The IoT layer in which the ML algorithms have been implemented in the studies targeting smart cities is summarized in Fig. 1.3.4, 6 It may be found that seven studies have implemented ML algorithms in the middleware layer. Furthermore, other seven studies have implemented ML algorithms in the sensing layer, where a preference for implementing ML algorithms in the IoT sensing edge is apparent. Regarding the problems to be solved by the ML algorithms, Fig. 1.4 reveals that the majority

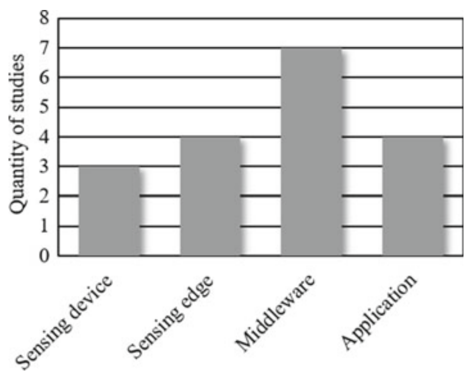


Fig. 1.3 IoT layer in which the ML algorithms are implemented