

**AN INTEGRATED RESEARCH FRAMEWORK OF SMART
CITY USING BIG DATA ANALYTICS: A STUDY.**

PULKIT



**BHARTI SCHOOL OF TELECOMMUNICATION
TECHNOLOGY & MANAGEMENT
INDIAN INSTITUTE OF TECHNOLOGY DELHI**

JANUARY 2024

**AN INTEGRATED RESEARCH FRAMEWORK OF SMART
CITY USING BIG DATA ANALYTICS: A STUDY.**

by

PULKIT

Bharti School of Telecommunication Technology & Management

Submitted

In fulfilment of the requirements of the degree of

Doctor of Philosophy

to the



Bharti School of Telecommunication Technology & Management

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JANUARY 2024

CERTIFICATE

This is to certify that the thesis entitled “An Integrated research framework of smart city using big data analytics: A study, the research being submitted by Pulkit to the Indian Institute of Technology Delhi for the award of Doctor of Philosophy (Ph.D.) degree is an authentic document of original research carried out by him. He put his research efforts under our supervision and has satisfied the necessity for the submission of the thesis, which has reached the excellence for Ph.D. degree of the institute. The proposed result of this thesis has not been submitted, in part or full, to any other university and institute for the award of any degree or diploma.

Prof. P. Vigneswara Ilavarasan

Department of Management Studies

Indian Institute of Technology Delhi

New Delhi, India

ACKNOWLEDGEMENT

I am grateful to my Ph.D. supervisor Prof. P. Vigneswara Ilavarasan, for his guidance, support, and encouragement. Without his active guidance, the thesis could not have been in its present form. He always become the great source of inspiration during this Ph.D. journey. He devoted his valuable time and motivated me whenever I was disheartened.

I express my gratitude to the Student Research Committee (SRC) comprising Prof. Arpan Kumar Kar (Chairman) Department of Management Studies, Prof. B. Lall (Internal Expert) Bharti School of Telecommunication Technology and Management, and Prof. Surya Prakash Singh (External Expert) Department of Management Studies, for giving their full support and valuable comments, and suggestions.

I express my gratitude to the faculty, staff and research scholars of the institute for contributing to completing this research work.

I would also like to thank the experts who supported the active participation of the questionnaire designed for this research work.

I especially thank my parents, my wife, and my little Master Maanvik for their stand, and love all along my Ph.D. progress. I am also obliged to all the encouragers for their support in completing this research work.

Pulkit

ABSTRACT

The increasing attention of various countries all over the world to adopt the smart city to maintain sustainability and resilience becomes the cause of interest for active participation in this domain by researchers. This research work aims to cover the literature on smart cities, analyze the available literature, and present the model in this domain by exploring various opportunities. The research explores the various extent to design the technological conceptual framework, which further leads to identifying and prioritizing the factors to develop and restructure a smart city. A qualitative analysis was accomplished in this research to explore the complexities of the successful execution of the process for making decision about smart city. Finally, the data-based machine learning framework for traffic management in smart cities was designed in this research work.

This research work performs the standardized literature to explore the technological requirement for execution of smart cities. It identifies the current trends, research sources, contributing authors, and countries using research investigation approaches like bibliometric analysis, network, and content analysis. The bibliometrics investigation demonstrates that the research publication trends have continuously shoot up over the last couple of years. The conceptual framework of a smart city based has been proposed using this literature. The priority given to the technical conviction in this digital world because the technology changes at a very fast rate and becomes the challenge to understand and execution. The advanced technologies figured out from the conceptual framework led to identifying the technical variables to administer the smart city design.

The study identifies and prioritizes the variables for smart city implementation. The Multi-Criteria Decision Making (MCDM) techniques were applied to prioritize the variables. In total, thirty-five factors were associated with publications, and based on prioritization process, waste management evaluated on the top position among all the variables of a smart city, the comprehensive list of factors identified for intelligent infrastructure. After focusing on aspects of urbanization, then issues faced by the Indian government to implement projects for smart cities, in most scenarios, government initiated the projects for urban development but could not complete them within the stipulated period.

This research further identifies the barriers that affect smart city implementation and the possible action required to implement the smart city infrastructure in India. The literature review was performed to collect information on smart cities, their influence on society, and the reasons why this mission of urbanization is delayed in completion in various cities of India. The Situation Actor Process – Learning Action Performance (SAP-LAP) analysis is applied to understand and manage this complexity systematically. The encouragement of the smart city design in India got delayed due to insufficient funds, lack of technology, poor coordination between local, state, and central government, and untrained employees of authorities. The research identifies the comprehensive list of components that influence the pace of smart city development to get the best possible solution to deal with this complex situation. The smart city implementation faced multiple challenges, but these can be minimized by using proper strategy and planning to use available resources and technology.

The final aspiration of this research is to draft a machine learning (ML) framework for traffic management in a smart city. The Long Short-Term Memory (LSTM) is the ML

algorithm used to forecast traffic patterns more accurately. The research covers the various parameters to implement smart city models using big data and other advanced technologies.

सार

विभिन्न देशों द्वारा स्मार्ट शहर संस्कृति को अपनाने के लिए स्थिरता और लचीलापन बनाए रखने के लिए शोधकर्ताओं द्वारा इस डोमेन में सक्रिय भागीदारी के लिए रुचि का कारण बन जाता है। इस शोध कार्य का उद्देश्य स्मार्ट शहर के मौजूदा साहित्य को कवर करना है, उपलब्ध साहित्य के आधार पर कुछ विश्लेषण करना और विभिन्न अवसरों की खोज करके इस डोमेन पर मॉडल प्रस्तुत करना है। यह शोध तकनीकी अवधारणात्मक ढांचे को डिजाइन करने के लिए विभिन्न साहित्य की खोज करता है, जो आगे शहरी क्षेत्रों के डिजाइन और पुनर्गठन के लिए कारकों की पहचान और प्राथमिकता देता है। क्वालिटेटिव विश्लेषण भी इस शोध में स्मार्ट शहर के सफल कार्यान्वयन के लिए जटिलता का प्रबंधन करने के लिए किया गया था, और अंत में इस शोध कार्य में प्रस्तुत पूर्वानुमान मॉडलिंग का उपयोग करके स्मार्ट शहर पर डेटा आधारित निर्णय समर्थन।

यह शोध कार्य स्मार्ट शहरों के तकनीकी पहलुओं पर व्यवस्थित साहित्य का प्रदर्शन करता है और वर्तमान रुझानों की पहचान करता है, अनुसंधान के सूत्रों, लेखकों को योगदान देता है और द्विपक्षीय विश्लेषण, नेटवर्क और सामग्री विश्लेषण जैसी विश्लेषण तकनीकों का उपयोग करके देशों को योगदान देता है। बॉयोमीट्रिक अनुसंधान प्रकाशन के रुझान को दर्शाता है जो पिछले कुछ वर्षों में लगातार बढ़ता है। उद्योग 4.0 पर आधारित स्मार्ट शहर की अवधारणात्मक रूपरेखा ने इस साहित्य का उपयोग करके प्रस्तावित किया है। यह शोध स्मार्ट शहर के सफल निष्पादन के लिए तकनीकी अवधारणाओं पर विचार करता है। उन्नत प्रौद्योगिकियों की अवधारणात्मक ढांचे से बाहर आंकड़ा स्मार्ट सिटी मॉडल को लागू करने के लिए तकनीकी चर की पहचान की ओर जाता है।

अध्ययन स्मार्ट सिटी कार्यान्वयन के लिए चर की पहचान और प्राथमिकता देता है। बुद्धिमान बुनियादी ढांचे के लिए आवश्यक कारकों की पहचान में साहित्य समर्थन। बहु मानदंड निर्णय लेने, तकनीकों को चर को प्राथमिकता देने के लिए लागू किया गया। कुल मिलाकर, साहित्य से पहचाने गए पैंतीस कारक और प्राथमिकता अपशिष्ट प्रबंधन स्मार्ट शहर के सभी चरों के बीच सबसे अधिक स्थान पर है। यह शोध स्मार्ट इन्फ्रास्ट्रक्चर के लिए कारकों की व्यापक सूची की पहचान करता है। एक बार जब कारकों की पहचान हुई तो यह शोध कार्य भारत में स्मार्ट शहर की विभिन्न परियोजनाओं को लागू करने के लिए जटिलता से संबंधित है।

यह शोध आगे उन बाधाओं को पहचानता है जो स्मार्ट सिटी कार्यान्वयन को प्रभावित करते हैं और भारत में स्मार्ट सिटी इन्फ्रास्ट्रक्चर को लागू करने के लिए संभावित कार्रवाई की आवश्यकता होती है। साहित्य समीक्षा स्मार्ट शहरों पर जानकारी एकत्र करने, समाज पर उनके प्रभाव और कारण क्यों शहरीकरण का यह मिशन भारत के विभिन्न शहरों में पूरा होने में देरी हो जाती है। सैप-ळप विश्लेषण ने इस जटिलता को एक व्यवस्थित ढांचे में समझने और प्रबंधित करने के लिए लागू किया। भारत में स्मार्ट सिटी मिशन का विकास अपर्याप्त धन, प्रौद्योगिकी की कमी, स्थानीय, राज्य और केंद्रीय सरकार के बीच खराब समन्वय और अधिकारियों के कर्मचारियों को प्रशिक्षित करने के कारण देरी हो गई। अनुसंधान उन कारकों की व्यापक सूची की पहचान करता है जो स्मार्ट सिटी डेवलपमेंट की गति को प्रभावित करते हैं और इस जटिल स्थिति से निपटने के लिए सर्वोत्तम संभव समाधान भी प्रदान करते हैं।

इस शोध कार्य का अंतिम उद्देश्य स्मार्ट शहर में यातायात प्रबंधन के डेटा निर्णय समर्थन प्रणाली को डिजाइन करना है। स्मार्ट शहरों में यातायात अवसंरचना और वायु प्रदूषण प्रबंधन के निर्णय लेने के लिए मशीन लर्निंग एल्गोरिदम पर आधारित पूर्वानुमान विश्लेषण मॉडल। लॉन्ग शॉर्ट टर्म मेमोरी मशीन लर्निंग

एल्गोरिथ्म है जिसका उपयोग ट्रैफिक पैटर्न के पूर्वानुमान के लिए किया जाता है। अनुसंधान में बड़े डेटा और अन्य उन्नत तकनीकों का उपयोग करके स्मार्ट सिटी मॉडल को लागू करने के लिए विभिन्न पैरामीटर शामिल हैं। स्मार्ट सिटी कार्यान्वयन ने कई चुनौतियों का सामना किया लेकिन इन्हें उचित रणनीति और उपलब्ध संसाधनों और प्रौद्योगिकी का उपयोग करने की योजना का उपयोग करके कम से कम किया जा सकता है।

Table of Contents

CERTIFICATE	ii
ACKNOWLEDGEMENT	iii
ABSTRACT	iv
संक्षेप	vii
LIST OF FIGURES	xv
LIST OF TABLES	xvi
LIST OF ABBREVIATION	xvii
CHAPTER 1 – INTRODUCTION	1
1.1 Introduction	1
1.2 Background	4
1.3 Need and scope of smart city	6
1.4 Basic concepts of big data and predictive modelling	8
1.4.1 Basics of big data	8
1.4.2 Predictive modelling	13
1.5 Smart city working process	15
1.6 Research motivation	17
1.7 Outline of the thesis	18
1.8 Proposed research methodology	24
1.9 Organization of thesis	26
1.10 Chapter Summary	27
CHAPTER 2 – LITERATURE REVIEW	28
2.1 Introduction	28
2.2 Developing smart cities: An integrated framework	29

2.3 State-of-the-art	32
2.4 Smart, sustainable cities of the futures: An extensive interdisciplinary literature review	33
2.4.1 Search strategy and scholarly sources	34
2.4.2 Inclusion and exclusion criteria	34
2.4.3 Specific purpose of literature	35
2.5 A brief literature review on a technological conceptual framework of smart city	36
2.5.1 Big data	36
2.5.2 Citizen-centered approach	36
2.5.3 Edge computing	37
2.5.4 IoT in smart cities	37
2.5.5 Compact and green cities	38
2.5.6 Future of waste management in smart cities	38
2.5.7 Industry 4.0	39
2.5.8 Industry 1.0 to Industry 4.0	40
2.5.9 Industry 4.0 and smart cities	42
2.6 Factors of smart cities	42
2.6.1 Waste management in smart cities	42
2.6.2 Smart infrastructure in smart cities	43
2.6.3 Information role in smart cities	43
2.6.4 Policies for urban development	43
2.6.5 Quality of life in smart city	44
2.7 Literature cover the complexity of smart city implementation	44

2.8 Literature for the decision support system of smart cities	47
2.8.1 Big data collection in smart cities	48
2.8.2 Big data support to intelligent cities	49
2.8.3 Big data modifies the living standard in smart city	49
2.8.4 Traffic identification in an intelligent city	50
2.8.5 Traffic congestion issues in smart city	51
2.8.6 Impact of traffic on the environment	52
2.8.7 The role of planning in smart city development	52
2.8.8 Decision making approach in smart cities	53
2.8.9 ML based decision-making approach to smart city	53
2.8.10 Classification of data based on machine learning method	54
2.8.11 Sustainability improvement using big data	54
2.9 Research Gaps	55
2.10 Research Questions	56
2.11 Research Objectives	56
CHAPTER 3 CONTENT ANALYSIS OF LITERATURE ON BIG DATA IN SMART CITIES	58
3.1 Introduction	58
3.2 Background and Problem statement	58
3.3 Research methodology	62
3.3.1 Bibliometric analysis	64
3.4 Network analysis	68
3.4.1 Big data technology	70
3.4.2 Smart city	70

3.4.3 Data mining	70
3.4.4 Edge computing	71
3.5 Content analysis	71
3.6 Observation	74
3.6.1 Conceptual framework	78
3.6.2 Reliability	82
3.6.3 Discussion	85
3.7 Chapter summary	88
CHAPTER 4 A METHODOLOGICAL APPROACH TO IDENTIFY AND PRIORITIZE THE FACTORS TO IMPLEMENT SMART CITY	91
4.1 Introduction	91
4.2 Research background and Problem statement	92
4.3 Factors of smart city	94
4.4 Proposed methodology	100
4.5 Proposed framework and analysis	105
4.5.1 DEMATEL analysis	106
4.6 Finding and discussion	113
4.7 Chapter summary	118
CHAPTER 5 – SAP-LAP AND ISM BASED ANALYSIS TO MANAGE THE COMPLEXITY OF SMART CITY DEVELOPMENT	120
5.1 Introduction	120
5.2 Background and Problem statement	120
5.3 Proposed methodology	123

5.4 Proposed framework	129
5.4.1 SAP-LAP analysis	129
5.4.2 Modeling using ISM	142
5.5 Finding and discussion	149
5.6 Chapter summary	150
CHAPTER 6 – THE MACHINE LEARNING FRAMEWORK FOR TRAFFIC MANAGEMENT	152
6.1 Introduction	152
6.2 Research background and Problem statement	152
6.3 Proposed methodology	155
6.3.1 Classification of algorithm	155
6.3.2 LSTM model	156
6.3.3 Accuracy and validation	159
6.4 Analysis of data	161
6.5 Machine learning framework	164
6.6 Finding and discussion	167
6.7 Chapter summary	170
CHAPTER 7 - SUMMARY OF RESEARCH FINDINGAND DISCUSSION	172
7.1 Introduction	172
7.2 Summary of research findings	172
7.3 Significant contribution of this research for academics and practising world	174
7.4 Outcome of this research	176
7.5 Limitations of the research	177

7.6 Future scope of research	178
7.7 Chapter summary	178
References	180
Publications from Thesis	223

LIST OF FIGURES

Figure 3. 1: Flow chart for the literature search and analysis	64
Figure 3. 2: Publication Trends	65
Figure 3. 3: Network of clusters	69
Figure 3. 4: Network of cluster of research papers	72
Figure 3. 5: Conceptual framework of smart city based on technical consideration	79
Figure 4. 1: DEMATEL cause-effect diagram visually categories factors of the smart city	110
Figure 4. 2: Enlarge portion of figure 4.1	111
Figure 4. 3: Proposed research flow framework	113
Figure 5. 1: SAP-LAP framework	127
Figure 5. 2: ISM model for smart city	146
Figure 5. 3: MIC-MAC analysis for smart city	147
Figure 6. 1: LSTM structure	157
Figure 6. 2: The process of data management in the smart city model	160
Figure 6. 3: Forecast of traffic based on smart on LSTM	163
Figure 6. 4: Machine learning framework for traffic management of smart city	166
Figure 6. 5: Transportation cost with supply chain	169

LIST OF TABLES

Table 1. 1: Research objective and methodology mapping	26
Table 2. 1: Literature search table	48
Table 3.1: Search table	63
Table 3. 2: Top authors	66
Table 3. 3: Top journals	67
Table 3. 4: Top contributing countries	68
Table 3. 5: Six separate technical clusters of smart city	73
Table 3. 6: Top publications	74
Table 3. 7: Reliability statistics	84
Table 3. 8: Variable statistics	84
Table 4. 1: Factors of smart city	94
Table 4. 2: Normalization matrix	107
Table 4. 3: Total relation result of DEMATEL	108
Table 4. 4: Influence is given and receive by each factor	111
Table 5. 1: Binary substitution	129
Table 5. 2: List of factors derived from SAP-LAP analysis	141
Table 5. 3: SSIM matrix	143
Table 5. 4: Reliability matrix	144
Table 5. 5: Partition levels	145
Table 6. 1: Accuracy of various models	162
Table 6. 2: Accuracy result	164

LIST OF ABBREVIATION

AI: Artificial intelligence

BC: Blockchain

CC: Cloud computing

DEMATEL: Decision making trial and evaluation laboratory

EC: Edge computing

ICT: Information and communications technology

IoE: Internet of energy

IoS: Internet of services

IoT: Internet of things

ISM: Interpretive structural modeling

IT: Information technology

kNN: k nearest neighbor

LSTM: Long short-term memory

MAE: Mean absolute error

MAPE: Mean absolute percentage error

MCDM: Multi criteria decision making

ML: Machine learning

MoHUA: Ministry of housing and urban affairs

QoL: Quality of life

RF: Random Forest

RMSE: Root mean squared error

SAP-LAP: Situation, Actor, Process – Learning, Action, and Performance

SMELTS: Social, management, economics, legal, technology, and sustainability

SPV: Special purpose vehicle

SSIM: Structural self-interaction matrix