Smart City Ranking Based on Big Data

A thesis submitted in fulfillment of the requirements for

the award of the degree of

Doctor of Philosophy

In

Information Technology

Submitted By

Sandeep Tayal

2K14/PHD/IT/01

Under the Guidance of

Prof. Kapil Sharma

Professor & Head of Department

Department of Information Technology

Delhi Technological University, Delhi - 110042, India

Delhi Technological University,

Shahbad Daulatpur, Main Bawana Road,

Delhi - 110042, India.

2020

DECLARATION

I, Sandeep Tayal, Ph.D. student (Roll No.2K14/PHD/IT/01), hereby declare that the thesis entitled **"Smart City Ranking Based on Big Data"** which is submitting for the award of the degree of Doctor of Philosophy in Information Technology, is a record of Bonafede research work carried out by me in the Department of Information Technology, Delhi Technological University. I further declare that this work is based on original research and has not submitted to any university or institution for any degree or diploma.

Date: _____

Place: New Delhi

Sandeep Tayal,
2K14/PHD/IT/01,
Department of Information Technology,
Delhi Technological University,
Delhi -110042.

CERTIFICATE

DELHI TECHNOLOGICAL UNIVERSITY

(Govt. of National Capital Territory of Delhi)

BAWANA ROAD, DELHI - 110042

Date: _____

This is to certify that the work embodied in the thesis titled "**Smart City Ranking Based on Big Data**" has been completed by Sandeep Tayal under my supervision and guidance towards fulfillment of the requirements for the degree of Doctor of Philosophy in Department of Information Technology of Delhi Technological University, Delhi. This work is based on original research and has not been submitted in full or in part for any other diploma or degree of any university.

Prof. Kapil Sharma
Professor & Head of Department,
Department of Information Technology,
Delhi Technological University,
Delhi -110042.

ACKNOWLEDGMENT

First and foremost, I would like to thanks Prof. Kapil Sharma, who has always extended unconditional support and guidance in carrying out my research work. He has always been very co-operative and understanding, and I am thankful for his constant support and encouragement throughout this work.

I would now wish to extend my deep sense of gratitude to Prof. Kapil Sharma, who has been instrumental in shaping my research and who has guided me from time to time by giving his valuable inputs and suggestions, which helped me in continuing with my work smoothly. Whenever I was in stress or found myself at sea, he encouraged and motivated me and brought the best out of me. Besides being my mentor, he has been a pillar of strength, and I am deeply indebted to him for guiding me with his extensive knowledge and insightful discussions.

I am deeply indebted to Prof. Yogesh Singh, Vice-Chancellor, Delhi Technological University, for providing necessary facilities.

I express my thanks to the office, technical, and library staff of the department for their help at various stages of my work. I would also like to thank all my teachers and friends for their direct and indirect help and support in this work.

Lastly, I would like to thank my family, who has been very loving and supportive throughout the research period.

Sandeep Tayal

Abstract

The smart urbanization is growing in international urban planning for the last decade. The smart city expresses to combine information and communication technology that defines smart living. A smart city has an infrastructure to provide life quality, a safe and clean environment to its citizens by smart technology. These days people want to live a smart life. So, the urban planner/ researcher is evolving “smart city models” based on the following six dimensions, “environment,” “economy,” “people,” “governance,” “mobility,” and “living.” This model was developed based on “American” and “European” cities.

“Make in India” Smart City is a scheme taken by the “Government of India” in 2014 to encourage the multi-national, among national companies, to help in making in Indian smart city. After the start of this program in India would emerge as the top focus in the world for foreign direct investment to make smart city projects. The major objective behind the start is to focus on development creation and smart skill enhancement in the sector of smart city development to living better. The start of Indian smart city also aims to achieve high-quality benchmarks combined with the lowest effect on the environment. The smart city hopes to attract technical as well as a capital investment in India.

Smart Cities India is going to grow into the most populated country. The growing population will be transferred to the top-level by cities of India. It will be challenging to cope up with the ever-increasing population in cities with the presently available urban area. So, the vision Prime Minister Narendra Modi is going towards developing more smart

cities, so the plan is to build 100 smart cities has started. The problem of ranking smart cities with the “Indian Smart City Model” is arising.

Big data is a solution to manage the massive volume of both structured and unstructured data. The processing of an extensive data set with verity and volume is not possible on traditional software of database and techniques. The generation of data is too large and frequently changes, so the database is becoming too heavy to analyze the information from traditional databases.

The objective starts with making a model for Indian smart city with the dimensions and indicators, according to India. This thesis proposed an India Smart City Model with the Eight Dimensions “Demographics Profile,” “Economic Profile,” “Infrastructure Profile,” “E-Governance & computerization,” “Finance,” “Environmental,” “Progress Track,” “Security,” and 80 indicators. These Dimensions and indicators based on the Indian smart city. The present smart city models are based on the European and American smart city indicators. The Indian smart city ranking is now calculating based on eight dimensions and eighty indicators.

The Ranking decisions are made on multiple criteria-based problems solved using the “distance-based algorithm (DBA)” to rank the Indian “Smart City.” The algorithm calculates the optimal value solution based on the distance to make an optimal ranking solution for the Smart Indian cities. The multiple indicators having different units and values, the DBA algorithm standardizes all indicators so that the optimal value of ranking can be provided.

The second algorithm, “Taxicab Distance-Based Approach (TDBA),” is proposed to rank the “Indian smart city.” The TDBA finds the optimal solution on the bases of “Indian smart city” indicators multiple values. The approach calculates the optimal distance solution using taxicab distance to find the best result.

The result is showing the ranking of Indian Smart City based on the proposed “Indian Smart City Model,” having eight dimensions and eighty indicators using the algorithms defined in DBA and TDBA. The individual ranking of each dimension for every city is done so that the growth of development can be visualized.

This information used for evaluation of city rank using the DBA and TDBA approaches. DBA and TDBA is a mathematical tool that has been used to aggregate and convert data into a standard form that is used to rank “Indian smart cities.” The result clarifies cities' vision and creates a blueprint of the cities that can be useful in the near future.

Keywords: Big Data, Distance-Based Approach, Ranking Model, Indian Smart City.

Table of Contents

List of Figures

List of Tables

Nomenclature

ICT	Information and Communication Technologies
SBM	Swachh Bharat Mission
AMRUT	Atal Mission for Rejuvenation and Urban Transformation
HRIDAY	National Heritage City Development and Augmentation Yojana
ICMI	IESE Cities in Motion Index
DBA	Distance-Based Algorithm
TDBA	Taxicab Distance-Based Approach
WOA-AMP	Whale Optimization Algorithm with Adaptive Multi-Population
RF-SVM	Random Forest with Support Vector Machine
ISO	International Organization for Standardization
TC	Technical Committees
DCPP	Data Computation and Processing Phase
FG-SSS	Focus Group on Smart Sustainable Cities
IUME	Integrated Urban Monitoring in Europe
ITU	International Telecommunication Union

Chapter One: Introduction

This chapter introduces the concept of Smart City and Big data. The motivation to make the "Indian Smart City Ranking Model" based on big data. The objectives of the research work define. Chapter-wise thesis reporting is brief at the end of the chapter.

The smart urbanization is getting popular in international urban planning for the last decade. The smart city word is to combine information and communication technology to define smart living. A smart city has an infrastructure to provide life quality, a safe and clean environment to its citizens by smart technology. Big data is a solution to manage the massive volume of structured data and unstructured data. The processing of an extensive data set with verity and volume is not possible on traditional software of database and techniques. The generation of data is too large and frequently changes, so the database is becoming too heavy to analyze the information from traditional databases. Big Data can smartly transform the entire process of a smart city.

1.1 Background

The smart city word is more popular from the last decade, and the urbanization is growing rapidly. The smart city word is to combine information and communication technology to define smart living. A smart city has an infrastructure to provide life quality, a safe and clean environment to its citizens by smart technology. These days people want to live a smart life. So, the urban planner/ researcher is evolving "smart city models" based on the following six dimensions, "environment," "economy," "people," "governance,"

“mobility,” and “living.” This model was developed based on “American” and “European” cities.

The urban area growth is increasing from the last decade, and people desire to stay in a smart city. In the previous 60 years, the metropolitan area expanded from 62 million to 377 million approximately in 1951-2011[1], the five-time extension of the urban area. India is rapidly increasing towards the urbanization, 30% of increment is shown from the year 2001-2011 from 1951 to 2011[2]. The Globally urban population is more than 50 % and generated nations having 80% of the urban population[3], this reality validates that India has more urbanization in the arriving decades. India becomes the market for the smart city project. It assumed that 600 million transaction economy would achieve up to 2030.[4] The smart city concept in India boosted by the prime minister of India, with the announcement to developed 100 smart cities in 2014. [5]. The Smart City[6] start the project for developing the smart city after the declaration of “Make in India” smart city. “Smart Cities Mission”, “Atal Mission for Rejuvenation and Urban Transformation (AMRUT)", [7]" this is start for the transformation of cities to become the smart city, the Swachh Bharat Mission (SBM)," [8] is initiative for environment to make clean and green smart city, Sardar Patel Urban Housing Plan (SPUHS) [9], is launched for the housing for all, to decrease the poverty area in the city. “National Heritage City Development and Augmentation Yojana (HRIDAY)”[7] is made for the fast development of the cities. The "Information and Communication innovation" (ICT)[10] is to developed technology about e-governance and smart services for smart cities. The "Information and Communication

innovation" (ICT)[10] to improve the gathering of information for the development of the city.

The “smart city model” was proposed by many researchers from the start of 2006. Some popular basic model is given by “Ruddolf Giffinger”[10]. This model is covering the indexing of mid-size European cities based on six basic dimensions of the smart city indicators taken according to the smart European cities.

“Boyd Cohen Dr. Cohen” [11] provided the model for “North America,” “Latin America,” and the “Asia Pacific” on the consideration of six basic dimensions of the smart city.

“IESE Cities in Motion Index(ICMI)”[12] is to measure the growth motion of the cities. This model gave the indexing of sites based on nine dimensions in 2014[13]. In recently 2018, update the tenth dimensions for the motion index.

There are more than fifty models, and a definition proposed by the researcher discussed in the literature review section. It shows the growth of smart city research among academic, industrialist, corporate, and researchers. The models discussed can not apply to Indian smart city as it is. So, the problem to develop Indian Smart City can be solved by proposing eight dimensions with eighty indicators “Indian Smart City Model” in this thesis.

1.2 Motivation

“Make in India” Smart City[14] is a scheme taken by the “Government of India” in 2014[15] to encourage the multi-national, among national companies, to help in making in Indian smart city. After the start of this program, India would emerge as the top focus

in the world for foreign direct investment to make smart city projects. [15]. The major objective behind the start is to focus on development and smart skill enhancement in the sector of smart city development for better living. The beginning of Indian smart city also aims to achieve high-quality benchmarks combined with the lowest effect on the environment. The smart city hopes to attract technical as well as a capital investment in India. [3]

Smart Cities India is going to grow into the most populated country in the world. The growing population will be shifting to the top tier cities of India. With this perspective, “Prime Minister Narendra Modi” vision “Digital India,” has set a motivated plan to construct 100 smart cities throughout the country[7]

Big data is a solution to manage the massive volume of both structured data and unstructured data. The processing of a large data set with verity and volume is not possible on traditional software of database and techniques. The generation of data is too large and fast changes, so the database becomes too heavy to analyze the information from traditional databases.[16]

India can smooth its operations, and the ability of fast processing of data makes smart decisions making with the help of Big Data analysis[17]. The data collected can be manipulated, stored, and analyzed to gain useful insight to grow the quality of life.[18]

The likelihood of smart city creativities, the improvement of enormous information, has had an indispensable impact on offering the potential for urban areas. For understanding the smart city, the choice and control administration is the most compelling

variable [19]. The term “huge information” is used to clarify an expansive number of both organized and unstructured data that it is perplexing to process utilization of customary database and software instruments since it is enormous and multifaceted [20]. Huge information investigation characterized as the way toward testing a huge informational index to uncover hidden designs, unknown relationships, and other key data, to settle on the choices that can be abused[21]. Prescient displaying, content investigation, machine getting the hang of, gauging, and factual examination are the propelled systems utilized by the huge information examination. Huge information examination will distinguish patterns, weak spots, or decide conditions though settling on better and speedier choices[22].

For the change and game plan of huge information applications for smart cities, different difficulties are encountered. The fundamental part is to maintain a strategic distance from or if nothing else restrain the difficulties influenced in smart sales aim and change for brilliant Smart City because the shrewd urban areas pondered unique and developing conditions [23][24]. The difficulties, for the most part, center on business and innovation by understanding the imperative smart condition includes that empowers cities to realize the idea, standards, also necessities of the smart cities’ entries [25]. One of the difficulties of constructing and managing the smart city is the comfort and size of such assets and their capacities. Administrative frameworks are further tests. The odds of progress extraordinarily influence by administrative frameworks[26]. Numerous applications are utilized to determine the difficulties, one of the fitting stages for exceedingly asset severe applications for dynamic collaboration between various requesting.

The utilization of enormous information learning on behalf of the city encourages viable storing information, also handling to create data that could upgrade diverse, savvy city offices. Huge information utilized to the leader's procedure for any expansion in smart city offices and highlights[27]. While enormous information frameworks use the information, the savvy city applications deliver a tremendous measure of information to give data, yet it will build up the shrewd urban area’s applications[28]. In the brilliant city, the enormous information applications can help numerous parts, for example, business, medicinal services, and transportation; better order in business, precautionary offices for social insurance, to advance courses and plan for transportation and increment conservational agreeability [29]. The significant objective of the smart city condition is meant to improve the nature of administrations open to the general public individuals by creating the use of resources productively and open administrations. The administrations involving smart home, transportation, wellbeing, etc., are improved by the self-representing information gathering. It is extremely fundamental to the procedure and manages a significant measure of the information offered by the administrations to the general public. Afterward, the tremendous change in information capacity, prior preparing, and systematic strategies are not skilled to suit the information handling prerequisites[30]. Big Data investigation techniques can productively recover data than prior information examination apparatuses. Thus, the coordination of Big Data investigation with the brilliant city is fundamental to manage the heterogeneous with exact information administration. Besides, there are various endeavors prepared through both the industry and the scholarly community to the grasp of conviction of the smart city. In any case, various mechanism

additionally displayed given the individual exertion, which covers smart city administration applications. [17]

Additionally, it must create practical autonomous execution, information calculation, and essential leadership. Be that as it may, there are various difficulties looked by Smart City and Big Data[31]. The online office-based assembling of heterogeneous information brings interoperability issues that are extremely trying for asset use. Furthermore, information handling and essential leadership ought to do in the remedy way. Anyway, cities are being concocted with online offices to acquire criticism information from city individuals. Performing the proficient Big Data investigation[32], it is fundamental to total information accurately also accomplish adequate pre-handling and distinguished novel highlights with their level of significance. In the same way, Big Data incorporating information in the dissimilar configurations essential to deal with the issue. Keeping in mind the end goal to address the previously revealed problems, various individual and joined methodologies are anticipated, yet at the same time missing non-specific and efficient engineering for the smart city model.[20]

The objective starts to make a model for Indian smart city with the dimensions and indicators, according to India. So, the thesis proposed an “India Smart City Model” with the Eight Dimensions “Demographics Profile,” “Economic Profile,” “Infrastructure Profile,” “E-Governance & computerization,” “Finance,” “Environmental,” “Progress Track,” “Security,” and 80 indicators.

1.3 Research Objectives

- It is proposed to develop a multidimensional Smart City model and its framework based on the big data paradigm that will include feedback of people from available facilities (e.g., online facilities).
- To identify and define a number of features with importance level and apply existing features selection techniques on newly proposed model.
- To develop an algorithm based on identified features for optimal utilization of resources.
- To perform the ranking of cities based on identified features and data collecting from different sources.

1.4 Contribution of Research Work

The major research contribution is summarized by:

- The "Indian Smart City Model" has developed on the based on the multiple dimensions with multiple indicators according to Indian information.
- "Indian Smart City Model" proposed with new dimensions, "Economic Profile," "Demographics Profile," e-Governance& Computerization, "Infrastructure Profile," "Environmental," "Finance," "Progress Track," and "Security" per sub-continent of India.
- The "Progress Track" dimension introduced to collect the feedback of the smart city progress work as the scope of continuous improvement and utilization of funds tracked.

- The Distance-Based Algorithm has defined to find the optimal identified features for optimal utilization of resources to ranking the smart cites.
- The Taxicab Distance-Based Approach is proposed to rank the “Indian smart city.” The TDBA finds the optimal solution on the bases of “Indian smart city” indicators multiple values. The approach calculates the optimal distance solution to find the best result in ranking the smart city.
- The Indian Smart cites data is collected and simulates the algorithms to find the rank of smart cities. The individual dimensions rank also finds to show the growth level of the individual cites in each dimension.
- The recommendation system model for a smart city to work as the feedback system for online information.
- An Adaptive Whale Optimization Algorithm Guided Smart City Big Data Feature Identification for Fair Resource Utilization. In this, the algorithm Whale Optimization Algorithm with Adaptive Multi-Population (WOA-AMP) system as an inquiry process in a wrapper display driven by the notable relapse demonstrates regression model Random Forest with Support Vector Machine (RF-SVM)

1.5 Thesis Outline

The thesis arranged in five-chapter, which defined as follow:

THE SECOND CHAPTER Literature Review in the field of smart city and the role of big data in smart cities with smart city models. It covered the importance of big data and

users of the information to developed smart cites. The existing smart city models explain in detail. The existing smart cities ranking model development. The methods used by the ranking models. The last is key findings from the literature review.

THE THIRD CHAPTER is about the Smart city Proposed Model. It defines all the dimensions of the proposed model with its indicators. The data set of Indian smart city defined in detail. The recommendation system for a smart city to work as the feedback system for online information. In this model, the movie database taken for the simulation. An Adaptive Whale Optimization Algorithm Guided Smart City Big Data Feature Identification for Fair Resource Utilization. In this, the algorithm Whale Optimization Algorithm with Adaptive Multi-Population (WOA-AMP) system as an inquiry process in a wrapper display driven by the notable relapse demonstrates regression model Random Forest with Support Vector Machine (RF-SVM). Our proposed calculation gives the exact method to choose the most agreeable feature blend, which prompts ideal asset usage.

THE FOURTH CHAPTER defines the "Smart City Ranking Model," using the DBA and TDBA. The chapter explained the distance-based algorithm(DBA) and the taxicab distance-based approach(TDBA) to rank the smart city using the indicators and dimensions of Indian Smart city Proposed Model. The simulation results explained with the result table and charts.

THE FIFTH CHAPTER is to conclude the research work with future research scope. The references have provided at the end of the thesis.

Chapter Two: Literature Review

This chapter presents an overview of recently published research on smart city, and smart city models, big data smart city ranking methods. It overviews the dimensions and indicators of the smart city.

2.1 Introduction

The smart city area is more famous for the researcher to form the last decade. More than 100 definitions have proposed by the researchers. The development of a smart city, the industrialist, corporate, and researcher had offered more than 50 models for smart cities. [33] The model with the year explained in the nest section. This area of growth has now come to India after the announcement of developing 100 smart cities in India.[34]

The smart city model proposed by many researchers from the start of 2006. Some popular basic model is given by "Ruddolf Giffinger"[10]. This model is covering the indexing of midsize European cities based on six basic dimensions of the smart city. The indicators are taking according to the smart European cities.

Boyd Cohen Dr. Cohen [11] provided the model for "North America," "Latin America," and the "Asia Pacific" on the consideration of six basic dimensions of the smart city.

"IESE Cities in Motion Index(ICMI)"[12] is to measure the growth motion of the cities. This model gave the indexing of sites based on nine dimensions in 2014[13]. In recently, 2018[35] update the tenth dimensions for the motion index. The IESE having the 10 dimensions and 49 indicators to compute the motion index of the cities in 2014 [36] in

2018[37] IESE used 9 dimensions and 49 indicators to compute the motion index of the cities The method for calculation used assigns a value of 0 too low with high to 100 for scaling the cites.

The 2THINKNOW Sites [38] developed a model on the base of 3 dimensions and 162 indicators for the city analysis. Cultural Assets, Human Infrastructure, Networked Markets the analysis is doing for commercial purposes. Innovation Cities Index [39] ranks the city to used three factors. Thirty-one segments, one hundred sixty-two indicators. This indexing is not for smart cities.

Ruddolf Giffinger defined the ranking of medium-sized European cities[40]. The model has six dimensions and 74 indicators, according to Europe[41]. The model has some update in indicator to 90 to rank the smart European cities. “Mapping Smart Cities in the European Union” [36]to perspective for the “European Innovation Partnership” on Smart Cities and Communities.

University of Vienna Ruddolf Giffinger[41] is calculating the rank for the smart city based on dimensions “smart mobility,” “smart economy,” “smart living,” “smart environment,” smart people,” and finally” smart governance.” The calculation of the rank is done by $z_i = (x_i - \overline{x}) / s$, average 0, standard deviation 1 of all dimensions.

Arcadis Sustainable Cities Index[42] model has come with the concept of sustainability of the city. The model calculates the rank based on information people, plants, and profits of the city. The raking method is used the averaging the value of indicators and give percentage score for all the dimensions from 0 to 100. City as a

knowledge tool model given by Edvinsson (2006)[43] having the six dimensions and 11 indicators to defining the cities. Siemens Green City Index[44] gives the city indexing based on eight subthemes and thirty indicators. The method used by the calculation is the average score equal to weight 0-10 per indicator and 0-100 overall. [45]

It is a "Hierarchic structure" [46] using hypothetical max/min values, and a "geometric mean" as the changed weight of groups 1-100. Networked Society City Index [47] defines the city based on two perspectives, six dimensions, fifteen variables, thirty-five indicators represented by 41 proxies. The model used a "Hierarchic structure"[48] by using a geometric mean take into account. In this model, the value is assigned from 1 to 100 to calculate the city. Networked Society City Index[44] developed a model with the two perspectives, six dimensions, fifteen variables, thirty-five indicators in model. [49] [50].

IUME Integrated Urban Monitoring in Europe[51] used four dimensions with fifty-six indicators in model. The headline indicators of average model standardization in the variety 0-3, $x<1$ under average, $x> 1$ above average.

"ISO 37120" [39] & "Global City Indicators Facility"[39] is having 100 indicators with the 17 themes to calculate the rank. It gives the individual score to each indicator with its themes to rank the city. "ISO 37120" & "Global City" Indicators Facility gives a standard definition for the smart cities. [52] This model has the individual scale for calculation.so the optimal comparison within indicators has not done.

"ITU FG-SSC"[53] having the six dimensions with eighty-eight indicators for ranking the cities. "The International Telecommunication Union (ITU)"[54], "Focus Group on Smart Sustainable Cities (FG-SSS)," calculates the arithmetic means, partial means for dimensions possible 0-100. ITU FG-SSC [55]used the six dimensions and 88 indicators. In the model, the scale 0 to 100 score assign to rank the city.

Shapiro [56] defines the city based on the Neoclassical city growth model with three dimensions and nine indicators[57]. It considers only Employment growth sources, productivity, quality of life.

Convocation of Biological Diversity[58] having three dimensions, "Native Biodiversity in the City," "Ecosystem Services provided by Biodiversity," "Governance and Management of Biodiversity." it is about the diversity of the cities with 23 indicators.

The Indians are becoming quicker in urban territories over the most recent couple of decades. Along these lines, the ranking model for "Indian smart cities" needs to create. " Shakti Sustainable Energy Foundation with ISB"[11] created a ranking model for "Indian smart cities" with six dimensions and fifty-eight indicators as an act in the "European ranking model"[42]. This model keeps on the dimension and indicators of the "European" [59]cities. So, the model, matching to the smart Indian city, is proposed in this thesis with eight dimensions and eighty indicators corresponding to "Indian smart cities." The all the model assigns a numeric value for calculation. The model does not work on the actual value of indicators.

All these models used the traditional techniques for calculation like average and giving score. The models are given the ranking of cities not based on the actual quotative value of the indicators.

The literature analysis gives detail of models proposed by researchers that all the defined model ranks the cities based on different indicators of related cities. The all the smart city model is based on their area related indicators. The calculation method of ranking is different in several models. The model is not considered the normalization of indicators to avoid the influence of measurement units. The collection of information is difficult for all indicators of cities. The DBA and TDBA solve all these issues to calculate optimal solution ranking.

2.2 Big Data

The bid data term introduces since the 1990s. The big data is to manage and store structured and unstructured data together. The ability to manage such a huge variety of data makes it popular in the market and towards the researchers. The information stored in big data collected from different sources. The concept of big data is providing a solution to smart cities to manage data smartly.

Big data [60]is a solution to manage the huge volume of both structured data and unstructured data. The processing of a large data set with verity and volume is not possible on traditional software of database software and data techniques. The generation of data is too large and fast changes, so the database is becoming too heavy to analyze the

information from traditional databases. The challenges include search, storage, sharing, capture,

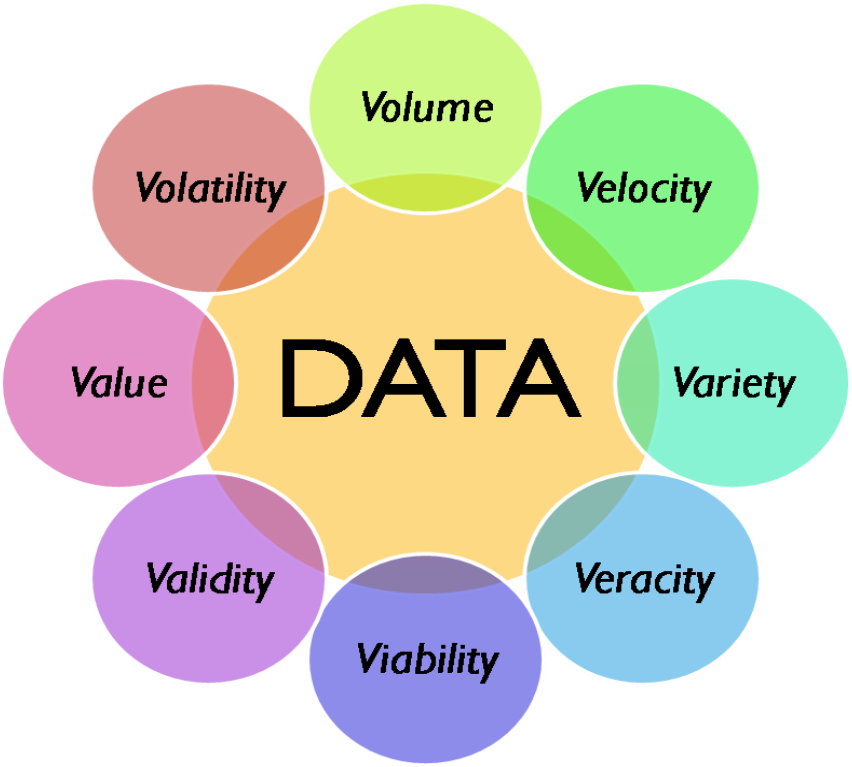
In the present scenario, data has acquired 10V characteristics Figure 2.1, namely: Volume, Velocity, Variety, Variability, Veracity, Validity, Vulnerability, Visualization, Value, and Volatility. It is the magic of 10Vs that data is growing every second. So, this information managed by big data.[61]

Big Data is a large gathering of data sets[62], so large in variety, volume, and velocity that it is impossible to evaluate and handle them using on-hand managing tools or traditional methods. To pick up an incentive from this information, you should pick an elective method to process it. Big Data has turned out to be reasonable as practical methodologies have developed to tame the volume, speed, and assortment of monstrous information.

The trend of urbanization is rising word wide. All the cities were looking forward to smart work or becoming smart cities. But the city has a complicated ecosystem with many subsystems to make it work smatter such as restaurants, residences, offices, shopping malls, entertainment, transport, water, energy, education, health care, communication, politics, business, waste, etc.[63] A smart city gets the benefit of digital tools to boost functioning and maintenance, to minimize pricing and resource ingesting, and to occupy more essentially and vigorously with its civilians.[64] A smart city should be capable to react faster to the city and challenges in the global scenario than a traditional relationship with its citizens. Other words used for the similar concept include “cyber Ville,” “digital

city’’, “electronic communities,” “Flexi city,” “intelligent city,” “knowledge-based city,” “MESH city,” “telicity,” “telecopies’’, “Ubiquitous City.”[65]

Big data [60]is a solution to manage the huge volume of both structured data and unstructured data. The processing of a large data set with verity and volume is not possible on traditional software of database software and data techniques. The generation of data is too large and fast changes, so the database is becoming too heavy to analyze the information from traditional databases. The challenges include search, storage, sharing,



capture, analysis, and visualization. [66]. The data processing to allow improved decision making, insight discovery, and process optimization. [67]

Figure 2-1 10 v' of Big Data

The Smart process needs to process a big volume of data that needs a Big Data solution[68]. Big Data provides cities to store, open, link, and analyze a large number of structured and unstructured data on top of which they are sitting. Big data provided predictive analysis and actionable insight. Smart Cities are all about sharing information, collaboration, and open dialogue. Smart City application[20] provides a need to view Open Data in a much larger context than it is usually defined. It is not just city administration disclosing their data in open platforms, but utility companies, citizens, healthcare providers, and all other stakeholders sharing their information with governments and each other for more development and utilization of information. Organizations using Big Data technology to collect all of the information digitally for greater intelligence[69].

2.3 Role of Big data in Smart city

Gartner defines "Big data" as a “high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making” [14]

IBM explains[70]that the idea of Big Data concerns to all data that can't be prepared or investigated utilizing customary apparatuses and procedures. In any case, it is essential to comprehend that traditional databases are significant and important logical solutions. [71]Furthermore, this huge volume of information exists in an mixture of information, which can be vocalized to in different manners around the globe, including versatile devices[72] GPS frameworks, audio, video, advanced sensors in incalculable modern hardware, autos, electric meters, vanes, and so on.., which can gauge and convey the position, development, vibration, temperature, stickiness and to concoction changes

practiced by the air, with the goal that these information examination applications require an exceptionally fast speed reaction to acquire the correct data at the perfect time. [73] Our urban areas loaded with data created by heterogeneous sources in various organizations, granularity, dynamism, and quality. Learning of this mind-boggling data space is fundamental to make Smart City administrations, and it is connected not exclusively to innovative issues basic centralization, stockpiling handling, and investigation of data, yet besides reaches out to issues, for example, wellbeing and property of the information created in the city, interoperability, etc.[74]

- It helps in reducing emissions. Using Big, Data companies can keep a check on the emissions and can apply techniques to bring down the emissions through analysis. It will also help the government in setting the emission parameters analyzing data of each industry[75].
- Sensors can be fitted on roads to collect traffic information and emissions, the technique implemented through big data. The data through sensors can be stored through big data techniques and can be analyzed[76].
- It can reduce parking problems. Cars can have sensors attached to them, which can take them to the nearest parking spot.[77]
- It can help in keeping a check on the garbage generated in the entire city and how disposed of it. This technique already implemented in Songdo, a city in South Korea.[78]

- It can help in saving electricity. Smart energy grids developed which can sense the presence of people in an area thus can help regulate the street lights in an area [79]
- It expands the exactness, quality, and revenue of the medicinal creation organizations. There are around 200 factors that should be dealt with to guarantee the immaculateness of fixings. The revenue can fluctuate from 50 to 100%. In light of Big Data investigation, the yield was expanded by around 50 %.[14]
- It coordinates examination over the Six Sigma system. It helps in improving the generation work processes by examining the different parts of creation and aides in conveying ideal quality items. [80]
- It helps in breaking down provider execution and quality after some time. By methods for Big Data investigation, makers can keep a view on the item quality and the exactness of conveyance continuously, which enables providers to get the items on the schedule. It likewise helps in looking after quality [77].
- It helps in estimating the moment subtleties of hardware and procedures. With the assistance of sensors, the data of each procedure given to the activity's chiefs. Along these lines, it demonstrates the quality and effectiveness of each machine and its administrators.[81]
- It enables organizations to comprehend and sell the most beneficial item arrangement, which has minimal impact on the creation. Makers are finding

which work to-arrange designs they should sell, which has an insignificant effect on the current creation plans. [66]

- It helps in dealing with the production network hazard. A production network is the key hazard zone where organizations are very concerned. Enormous Data investigation may help organizations conquer store network difficulties. For instance, the prescient examination may help organizations discover the probabilities of postponement, and consequently, organizations may work in like manner.[82]
- It helps in evaluating how creation may impact monetary execution. It has consistently been an issue for firms to set an association between day by day generation to money related execution. Be that as it may, with the assistance of Big Data, the situation is evolving. Enormous Data organizations may have live data about a plant floor working and can enable them to scale the tasks better [83].
- It enables organizations to accomplish the objective of Customer Fulfilment. With the assistance of Big, Data organizations can monitor every one of the issues looked by their clients while utilizing their items. Consequently, organizations may work to discover answers to the issues. It likewise enables organizations to break down the advantages given by different associations, and accordingly, they can improve their client administrations, which at last prompts an expansion in a benefit[84].

As indicated by the previously mentioned overviews and the examinations. The information about Big Data and its functions, it demonstrates a noteworthy job Big Data can play in “Make in India” brilliant Smart City. The outcomes have been very positive, and they show that it is so advantageous to execute Big Data. There has been a continuous increment in the generation of the organizations in the wake of executing Big Data in their organizations. Further factors like considerate plant execution over numerous frameworks and continuous cautions dependent on dissecting fabricating information have gone up altogether.[85] It is probably the greatest effect is on client administrations and bolster, which is one of the most significant elements for organizations. Clients center a great deal around client administrations, and along these lines helps in structure the organization's notoriety. Numerous different components have additionally demonstrated a consistent increment and their different variables that will demonstrate a progressive increment with time. Additionally, makers have given great evaluations for the advantages of enormous information in different parts of assembling.[86] The advantages are like item quality surrenders appraisals, supply arranging, and so forth, have been given great evaluations by different organizations. Along these lines, it demonstrates that producers have comprehended the estimation of Big Data and are agreeable to utilizing it. The time is close to when every single organization would utilize Big Data innovation.[87] Numerous worldwide associations have upheld the utilization of Big Data. Enormous Data and “Make in India” together can help in structure another India. It will profit every single organization which puts resources into “Make in India.” Enormous Data has worked in different nations, and it will work in India as well. It will help in understanding the need for individuals, and

the kind of items they should assemble. [88] Many world levels survey encouragement Big Data in manufacturing. The best objective of Big data is that it is not for an explicit industry. Thus, it can cover overall industrial evolution in the country with benefits for everybody reaching from consumers to industrialists.[89] Big Data is another idea, and it will require some investment for organizations to adjust to it. In any case, results show what Big Data can do to Manufacturing. The more industrialized a country is, the more created it is. Thus, Big Data and "Make in India" ds to a brilliant fate of India. Time is close to when we plan to see India in the top countries of the world in perspective.[90]

2.4 Smart City Models

Table 1 shows the smart city model from 2006 to 2018.

Table 1 Smart City Models

S.NO	YEAR	TYPE	GIVEN BY	MODEL
1	2017 2016 2015 2014	CITIES IN MOTION. CONCEPTUAL FRAMEWORK, DEFINITIONS, AND INDICATORS	IESE BUSINESS SCHOOL[91]	CITIES IN MOTION INDEX
2	2007-2018	BASIS OF THE INNOVATION CITIES™ INDEX	2THINKNOW SITES[46]	STRUCTURED DATA APPROACH
3	2018	CITIES IN MOTION. CONCEPTUAL FRAMEWORK, DEFINITIONS, AND INDICATORS	IESE BUSINESS SCHOOL[92]	CITIES IN MOTION INDEX
4	2017	THE BRUSSELS-CAPITAL REGION AS A SMART CITY	SMART CITY BRUSSELS[6]	THE BRUSSELS-CAPITAL REGION AS A SMART CITY
5	2017	KPMG SMART CITY MATURITY MODEL	KPMG[93]	KPMG SMART CITY MATURITY MODEL
6	2017	ISB	SHAKTI SUSTAINABLE ENERGY FOUNDATION WITH ISB[1]	ISB
7	2016	FACILITIES	CALVILLO ET, AL. (2016)[94]	SMART CITY ENERGY INTERVENTIONS AND ENERGY SYSTEM DESIGN MODEL

8	2016	SERVICES	FAN ET AL. (2016)[95]	SMART HEALTH ORGANIZATION MODEL
9	2016	POLICY EVALUATION BENCHMARKING TOOL	GOUVEIA ET, AL. (2016)[96]	GENERAL FRAMEWORK CONCEPT FOR AN INTEGRATIVE ENERGY CITY PLANNING
10	2016	GARUNDA SMART CITY MODEL	SMART INDONESIA INITIATIVE[97]	GARUNDA SMART CITY MODEL
11	2015	GOVERNANCE/SMART CITY PROGRESS BENCHMARKING TOOL	ALBINO AND ANGELICO (2015)[95]	SMART CITY DIMENSIONS
12	2015	ARCHITECTURE	ANTHOPOULOS (2015)[98]	SMART CITY DIMENSIONS
13	2015	CITY CAPACITY BENCHMARK TOOL	DE MARCO ET AL. (2015)[97]	SAFETY MEASUREMENT INDEXES
14	2015	SMART CITY MONITORING BENCHMARKING TOOL	MARSAL-LACUNA ET, AL. (2015)[99]	SMART CITY MONITORING INDEXES
15	2015	SUSTAINABILITY AND RESILIENCE BENCHMARKING TOOL	STRATEGIC ENERGY TECHNOLOGIES INFORMATION SYSTEM (SETIS)[42]	SMART ENERGY EFFICIENCY'S KPIs
16	2015	ENVIRONMENT	TSOLAKIS AND ANTHOPOULOS (2015)[100]	ECO-CITY SYSTEM DYNAMICS MODEL
17	2014	DATA AND KNOWLEDGE	BELLINI ET, AL. (2014)[101]	KNOWLEDGE MODEL FOR SMART CITY DATA (KM4CITY ONTOLOGY)
18	2014	SUSTAINABILITY AND RESILIENCE BENCHMARKING TOOL	DA CRUZ AND MARQUES (2014)[102]	SUSTAINABLE LOCAL GOVERNMENT SCORECARD
19	2014	SMART CITY PROGRESS BENCHMARK TOOL	DUARTE ET, AL. (2014)[103]	DIGITAL CITY ASSESSMENT FRAMEWORK
20	2014	ARCHITECTURE	GLEBOVA ET AL. (2014)[6]	SMART CITY CONCEPTUAL ELEMENTS
21	2014	SMART CITY PROGRESS BENCHMARK TOOL	GLEBOVA ET, AL. (2014)[104]	INDEXES FOR ICT ELEMENT EVALUATION
22	2014	ARCHITECTURE	IBM (SÖDERSTRÖM ET AL., 2014)[105]	NINE PILLAR MODELS SMARTER CITY EQUATION
23	2014	SUSTAINABILITY AND RESILIENCE BENCHMARKING TOOL	ISO (2014A)[106]	ISO 37120 SUSTAINABLE DEVELOPMENT OF COMMUNITIES INDICATORS FOR CITY SERVICES AND QUALITY OF LIFE
24	2014	GOVERNANCE	ISO (2014B)[95]	A TABLE OF CITY CHARACTERISTICS WHERE SMARTNESS IS APPLIED
25	2014	GOVERNANCE	ITU (2014A)[6]	ATTRIBUTES AND CORE THEMES

26	2014	SUSTAINABILITY AND RESILIENCE BENCHMARKING TOOL	ITU (2014B)[107]	SMART SUSTAINABLE CITY KEY PERFORMANCE INDICATORS (KPIS)
27	2014	POLICY EVALUATION BENCHMARKING TOOL	KII ET, AL. (2014)[42]	LAND-USE-TRANSPORT (LUT) MODEL FOR POLICY EVALUATION IN SMART CITY
28	2014	CITY CAPACITY BENCHMARK TOOL	KOURT ET, AL. (2014)[107]	GLOBAL CITY PERFORMANCE MEASUREMENT INDEXES
29	2014	GOVERNANCE	LEE ET, AL. (2014)[80]	FRAMEWORK FOR SMART CITY ANALYSIS
30	2014	GOVERNANCE/SMART CITY PROGRESS BENCHMARKING TOOL	LIU ET, AL. (2014)[42]	SMART CITY VALUE CHAIN (SCVC) MODEL
31	2014	ARCHITECTURE	NEIROTTI ET, AL. (2014)[98]	SMART CITY DOMAINS
32	2014	SUSTAINABILITY AND RESILIENCE BENCHMARKING TOOL	PIRES ET, AL. (2014)[108]	LOCAL SUSTAINABLE DEVELOPMENT INDICATORS
33	2014	CITY CAPACITY BENCHMARK TOOL	UN-HABITAT (2014)[53]	GOOD URBAN GOVERNANCE INDICATORS
34	2014	GOVERNANCE	UNITED NATIONS HABITAT (UNITED NATIONS, 2014)[109]	DIMENSIONS OF CITY PROSPERITY
35	2013	PLANNING AND MANAGEMENT	ANTHOPOULOS AND FITSILIS (2013)[98]	TECHNOLOGY ROADMAPPING FOR SMART CITY DEVELOPMENT
36	2013	SUSTAINABILITY AND RESILIENCE BENCHMARKING TOOL	DESOUZA AND FLANERY (2013)[110]	RESILIENCE CITY EVALUATION AND IMPLEMENTATION FRAMEWORK
37	2013	ARCHITECTURE	HANCKE ET, AL. (2013)[95]	SENSOR AREAS IN SMART CITY
38	2013	PLANNING AND MANAGEMENT	LEE ET, AL. (2013)[80]	TECHNOLOGY ROADMAPPING FOR SMART CITY DEVELOPMENT
39	2013	SMART CITY PROGRESS BENCHMARK TOOL	LEE ET, AL. (2013)[42]	SMART SERVICE ASSESSMENT INDEXES SMART DEVICE ASSESSMENT INDEXES TECHNOLOGY ASSESSMENT INDEXES
40	2013	ENVIRONMENT	SHWAYRI (2013)[78]	U-ECO-CITY MODEL
41	2013	CITY CAPACITY BENCHMARK TOOL	SINGHAL ET, AL. (2013)[111]	COMPETITIVENESS PARAMETERS
42	2012	GOVERNANCE	BARON (2012)[6]	THREE LEVEL-MODEL FOR CITY INTELLIGENCE FOR RESILIENCE CONCEPTUALIZATION
43	2012	DATA AND KNOWLEDGE	BATTY ET, AL. (2012)[112]	STRUCTURE OF FUTURISTS SMART CITY PROGRAMME

44	2012	SMART CITY PROGRESS BENCHMARK TOOL	FEI (2012)[42]	LEVEL OF SMART CITY PROGRESS
45	2012	CITY CAPACITY BENCHMARK TOOL	KOURTIT ET, AL. (2012)[58]	CITY INNOVATION POTENTIAL MEASUREMENT
46	2012	SMART CITY PROGRESS BENCHMARK TOOL	LAZAROIU AND ROSCIA (2012)[113]	MODEL FOR COMPUTING "THE SMART CITY" INDICES
47	2012	GOVERNANCE/SMART CITY PROGRESS BENCHMARKING TOOL	LOMBARDI ET, AL. (2012)[97]	TRIPLE HELIX MODEL FOR SMART CITY ANALYSIS AND PERFORMANCE MEASUREMENT
48	2012	SUSTAINABILITY AND RESILIENCE BENCHMARKING TOOL	MORI AND CHRISTODOULOU (2012)[114]	CITY SUSTAINABILITY INDEXES
49	2012	ARCHITECTURE	ZYGIARIS (2012)[65]	SMART CITY REFERENCE MODEL
50	2012	THE SMART CITY WHEEL	BOYD COHEN [11]	SMART CITY WHEEL
51	2011	SMART CITY PROGRESS BENCHMARK TOOL	CARAGLIU ET, AL. (2011)[115]	6 SMART CITY INDICATORS (DATA ANALYSIS FROM URBAN AUDIT DATASET)
52	2011	GOVERNANCE	LEYDESDORFF AND DEAKIN (2011)[116]	TRIPLE-HELIX MODEL OF SMART CITIES
53	2011	ARCHITECTURE	NAPHADE ET AL. (2011)[6]	SMART CITY MODEL
54	2011	PEOPLE	THITE (2011) [117]	URBAN FACTORS FOR HUMAN CAPITAL ATTRACTIVENESS
55	2011	CITY CAPACITY BENCHMARK TOOL	WINTERS (2011)[118]	FORMULAS FOR SMART CITY GROWTH MEASUREMENT
56	2010	SMART CITY MONITORING BENCHMARKING TOOL	MALEK (2010)[118]	SMART CITY MONITORING INDEX
57	2010	SINGAPORE INDEX ON CITIES' BIODIVERSITY	CONVOCATION OF BIOLOGICAL DIVERSITY[119]	FRAMEWORK OF THE SINGAPORE INDEX ON CITIES' BIODIVERSITY
58	2009	ARCHITECTURE	YOVANOF AND HAZAPIS (2009)[120]	DIGITAL CITY ARCHITECTURAL FRAMEWORK FOR SMART SERVICE PROVISION
59	2008	ARCHITECTURE	HOLLANDS (2008)[121]	SMART CITY MODEL
60	2007	ARCHITECTURE	GIFFINGER ET AL. (2007)[59]	SMART CITY COMPONENTS
61	2007	SMART CITIES RANKING OF EUROPEAN MEDIUM-SIZED CITIES	RUDOLF GIFFINGER[122]	RANKING OF EUROPEAN MEDIUM-SIZED CITIES
62	2006	DATA AND KNOWLEDGE	EDVINSSON (2006)[43]	CITY AS A KNOWLEDGE TOOL MODEL
63	2006	PEOPLE	SHAPIRO (2006)[34]	NEOCLASSICAL CITY GROWTH MODEL

2.5 Smart City Ranking Models

The smart city ranking model explains in Table 2 gives the detail of dimensions used by the different models.

Table 2 Smart City Ranking Models

Year	Model	Type	Given By	Categories	Indicators	1	2	3	4	5	6	7	8	9	10
2018	Cities in Motion Index	Cities in Motion. Conceptual Framework, Definitions, And	Iese Business School[37]	9	49	Economy	Human Capital	Social Cohesion	Environment	Governance	Urban Planning	International Outreach	Technology	Mobility and Transportation	
2017 2016 2015 2014	Cities in Motion Index	Cities in Motion. Conceptual Framework, Definitions, And Indicators	Iese Business School[91]	10	49	Economy	Human Capital	Social Cohesion	Environment	Public Management	Governance	Urban Planning	International Outreach	Technology	Mobility and Transportation

2012	**2007-2018**	**2010**	**2007**
Smart City Wheel	Structured Data Approach	The Framework of The Singapore Index on Cities'	Ranking of European Medium-Sized Cities
The Smart City Wheel	Basis of The Innovation Cities™ Index	Singapore Index on Cities' Biodiversity	Smart Cities Ranking of European Medium-Sized Cities
Boyd Cohen [11]	2thinknow Sites[46]	Convocation Of Biological	Rudolf Giffinger[59]
6	3	3	6
24	162	23	74
Smart Economy	Cultural Assets	Native Biodiversity in The	Smart Economy
Smart People	Human Infrastructure	Ecosystem Services Provided	Smart People
Smart Governance	Networked Markets	Governance and Management	Smart Governance
Smart Mobility			Smart Mobility
Smart Environment			Smart Environment
Smart Living			Smart Living

2017	**2017**	**2016**	**2017**
ISB	Kpmg Smart City Maturity Model	Garunda Smart City Model	The Brussels-Capital Region as A Smart City
ISB	Kpmg Smart City Maturity Model	Garunda Smart City Model	The Brussels-Capital Region as A Smart City
Shakti Sustainable Energy	Kpmg[126]	Smart Indonesia Initiative[125]	Smart City Brussels[124]
6	5	5	6
58	Na	13	19
Living	Leadership and Governance	Economy	Economy
Governance	Stakeholder Engagement and	Society	Government
People	Effective Use of Data	Governance	Environment
Economy	Integrated ICT Infrastructure	Environment	Mobility
Mobility	The Existing Level of Smartness	ICT And People	Society
Environment			Quality of Life

2006	City as A Knowledge Tool Model Data and Knowledge	Edvinsson (2006)[43]	6 11	ICT And Multimedia University Society and	Knowledge Cafes/Cathedrals Diversity Strange Attractors				
2006	A Neoclassical City Growth Model People	Shapiro (2006)[56]	3 9	Employment Growth Sources Productivity Quality of Life					

2.6 Smart City Ranking Method

- Taking an average of indicators within dimensions, average dimensions to receive a composite value. Arcadis Sustainable Cities Index model, rank is calculated by averaging of indicators for all dimensions and averaging dimensions to receive a combination rating uppermost indicator value with 100%, lowermost with 0% of the cities.

- Ruddolf Giffinger developed this model for ranking and monitoring the growth of the cities. This model calculates the rank by taking Arithmetic mean $zi = (xi - \overline{x}) / s$, average 0, standard deviation 1 of all dimensions.

- Shakti Sustainable Energy Foundation with ISB model is calculating the rank as the scale of 1 to 10 to each indicator where 1 = Very Poor and 10 = Excellent. zi = (xi –

$\bar{x}$) / s, average 0, standard deviation 1. The model gives the rank on the base of the scaled mean value.

- IUME Integrated Urban Monitoring in Europe has a Selection of 15 headline indicators and model average normalization in the variety 0-3, x<1 under average, x>1 above average.
- Networked Society City Index is a “Hierarchic structure” using a “geometric mean” as the changed weight of groups 1-100, the adding of 20% to the max/min value, or by using hypothetical max/min values.
- Siemens Green City Index has calculated for each subtheme average score equal to weight 0-10 per indicator, 0-100 overall. By this, it calculates the score from 100 to rank the city.

The all the ranking methods are not normalizing the value of the parameter of influencing measuring units. The criteria attribute either maximum or minimum is optimal, which is not considered by the ranking calculation.

2.7 Key Points from Literature Review

- The All model was having a different number of dimensions and indicators.
- All models used the techniques for calculation like average and giving the score.
- The models are given the ranking of cities not based on the actual quotative value of the indicators.
- The literature analysis gives detail of models proposed by researchers that all the defined model ranks cities based on different indicators related to their cities.

- The all the smart city model is not based on Indian cities indicators.
- The model is not considered the normalization of indicators to avoid the influence of measurement units.
- All cities may not gather all the indicator values; hence, it doesn't serve the purpose of a standardized indicator set used for ranking purposes.
- The models did not cover the optimal value for the ranking.
- The models do not have “Progress Tracking” and “Security.”

2.8 Summary

This chapter has given a detail literature review of the smart city, big data. The relation and role of big data have reviewed in detail. The existing smart city model has explained in detail with their dimensions and the ranking methods used by the existing models. The literature review finds the key points from the literature review.

The next chapter is about the proposed model for the smart city. This chapter is about the Indian smart city model proposed in this thesis with detail of dimensions and indicators. The Indian smart city data set is given in detail. The chapter also covers the two more models proposed based on the dimension of the Indian smart city model. One is the recommender systems model for smart cities. The second is based on the adaptive whale optimization algorithm to give us an air quality to think about ten urban areas, among that we discover the risk level of every city.

Chapter Three: Proposed Model for Smart City

This chapter is about the Indian Smart city Proposed Model. It defines all the dimensions of the proposed model with its indicators. The data set of Indian smart city defined in detail.

3.1 Introduction

The smart city is more popular in the last decade; more than 100 definitions proposed by the researcher, academician, corporate, government. Some of the literature reports define the smart city for making some standard definition the “ISO 37120:2014”, TMB smart city definitions defined in “ISO/IEC JTC 1 report” published in 2014[127]. The ISO also update the definition in “ISO 37120:2018” updated in the “ISO/IEC JTC 2 report” published in 2018[106]. The ITU-T Group also defines the smart city in the report FG-SSC report 2014[128] and again updates in 2015[104]. The BSI Group is also defining a smart city in 2014 with the vocabulary in PAS 1802014[129].

All the definitions focused on the six basic dimensions like governance, mobility, economy, people, living, and environment. The researcher does improve with the small changes in the definitions. The currently exiting model focused on these six dimensions of the definitions. All models designed to consider the indicators of the European and American cities. [130]. The mid-size European model had the six basic dimensions with a smart word like “Smart Governance,” “Smart Economy,” “Smart Mobility,” “Smart Environment,” “Smart Living”[41]. India is growing towards the smart city concept the

urban area growth increased in India, so the Indian prime minister has been imitating the 100 smart cities made in the project in 2014. [131].

The current model is not focused on the Indian urban area. So, the model for the Indian smarty city needs to developed, which has the dimensions and indicators according to Indian cites. The "Indian Smart City Model" in Figure 3.1 is proposed with eight dimensions, "Demographics Profile", "Economic Profile", "Infrastructure Profile", "E-Governance & computerization", "Finance", "Environmental", "Progress Track", "Security", and with 80 indicators according to Indian cites.

The data collection, according to the dimensions, is done with the help of big data techniques. The information in each dimension has its indicators.

Demographic Profile has information about the population growth, social structure of the society, the one important indicator about the youth and young working-age population. The profile has 13 such important indicators, all shown in Table 5.

The "Economic Profile" is very important to develop the smart city for all the existing model having this dimension, the indicators in focused in this dimension based on the income of cities with the available working sectors and its also considered the poverty rate and unemployment rate for the proper assumption of the economic growth. This profile has 21 indicators.

Infrastructure is playing a vital role in the growth of the city. This profile is about the number of schools, colleges, hospitals, etc. The infrastructure development calculated on the based of different 15 indicators shown in Table 5.

"E-Governance & computerization" is measuring development in the field of technology and smart services. The city should provide e-governance services for the smart management and living of the citizens. This profile has 12 indicators.

The "Finance" is describing the sustainability of the city, the financial indicators like expenditure, income sources, allocation budget to manage the city. This profile has 11 indicators to show the status of "Finance" of a city.

The "Environmental" dimension is to make a clean city. It shows the green and clean city parameter, which manage by Swachh Sarvekshan to conduct the environment Assessment of Indian cities.

"Progress Track" dimension is kept for keeping an eye on the growth of work done for the smart city. The progress of work is tracked by the percentage of work completion and utilization of funds.

"Security" is important with smart living as the smart crime rate would increase, so the safety index of the cities with crime index could be an indicator in this profile.

The all-embracing of "Indian smart cities"[2] is defined using these eight dimensions.

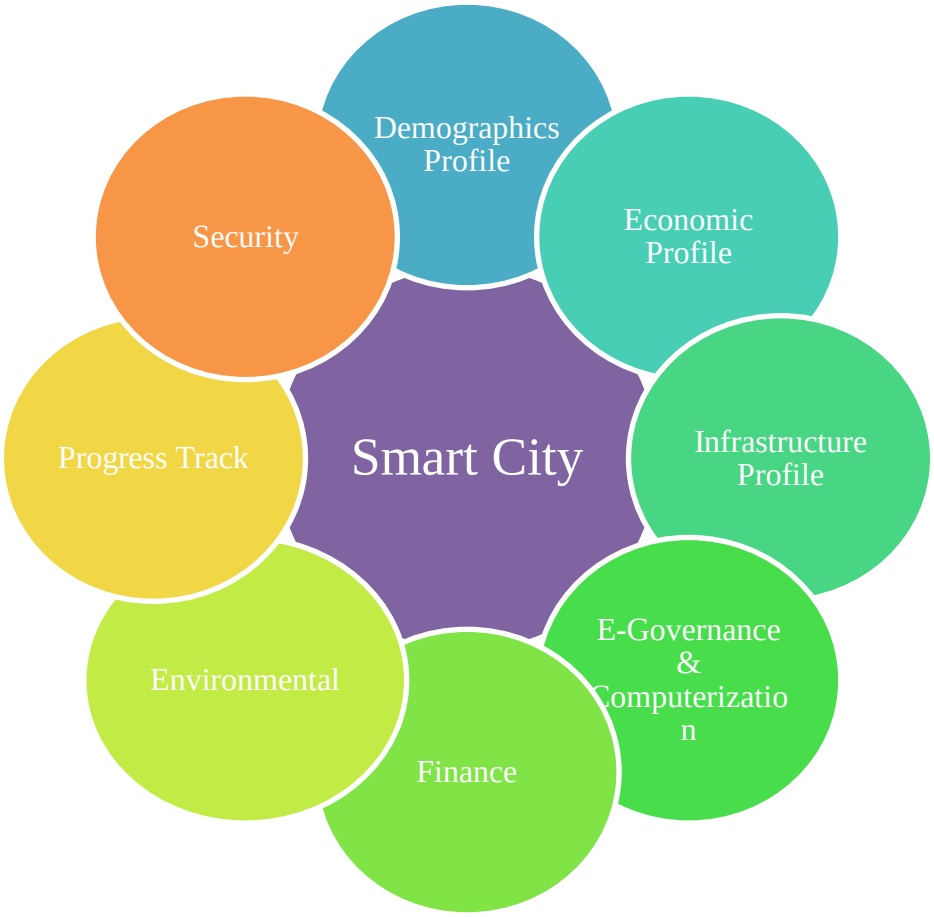


Figure 3-1 Indian Smart City Proposed Model

3.2 Indian Smart City Proposed Model

3.2.1 Demographics Profile

Demographic Profile is a collection of information about the population. This information is gathering through the surveys and the District Census Handbook, Census of India. The information contains the urban area total population, district urban area population, the growth rate of population, the area in a kilometer, percentage area of the district, the population density per square kilometer, the literacy rate, Percentage of social elements.

3.2.2 Economic Profile

The economic profile is a collection of economic data from various sources. This profile shows the per capita income of the city, the rate of unemployment, the ratio of urban poverty, rate of work participation, self-employment, salaried employees, labor, percentage of professionals, technicians, elementary occupations, clerks, workers not classified by occupation.

3.2.3 Infrastructure Profile

The infrastructure profile is the facility pillar for the smart city. It is about basic infrastructure like tap water, electricity, toilet, internet, mobile phones in households. The percentage of owned houses and rented houses. The main facility in number like hospitals, schools, colleges. It is a very important profile for "Make in India" smart city.

3.2.4 E-Governance & Computerization

E-government & Computerization profile is about the e-services available in the city. The smart city is to provide smart governance for smart living. It is having the e-governance facility of Property Tax, utility payments, Citizens' Grievance Monitoring, LAN, WAN, State Data Centre, and e-Procurement. The technology used for smart monitoring of life.

3.2.5 Finance

Finance is about the income source and expenditure of the smart city. It is the source of income from taxes, municipal, and budget with banking facilities. This information is very useful to calculate the sustainability of a city. The money expenditure for the development of the city.

3.2.6 Environmental

The Environmental dimension for the smart city is for green living. The development of a smart city with a green environment is important. This dimension covers the indicators for the clean and green city. The information collected through the surveys.

3.2.7 Progress Track

The tracking dimension is keeping the control and progress of smart city work. The progress tracking was done by monitoring the percentage of complete project and allocation and utilization of the project cost.

3.2.8 Security

The security is another dimension about the safety index and crime index of the city; the data is important for the safe and smart living. The smart city provides a clean, safe with smart work.

3.3 Data Set According to Indian Smart City Proposed Model

The data set of 20 Indian cities which name with symbols given in Table 3. The criteria of the smart Indian city for ranking given in Table 4. Indian Smart City Indicators Symbols are given in Table 5, the 80 indicators of 8 dimensions. The data value of 20 cities given in Tables 6 to 9. The data is an input to the ranking algorithms as 20 X 80 matrix of city-data.

Table 3 Indian Smart City Name Symbols

S. No	City	Symbol
1	Kakinada	C1

2	Vishakhapatnam	C2
3	Guwahati	C3
4	Delhi	C4
5	Ahmadabad	C5
6	Surat	C6
7	Belgaum	C7
8	Davanagere	C8
9	Kochi	C9
10	Bhopal	C10
11	Indore	C11
12	Jabalpur	C12
13	Solapur	C13
14	Pune	C14
15	Bhubaneswar	C15
16	Ludhiana	C16
17	Jaipur	C17
18	Udaipur	C18
19	Chennai (M. Corp)	C19
20	Coimbatore	C20

Table 4 Indian Smart City Criteria Symbols

S.No.	Criteria	Symbol
1	Demographics Profile	A
2	Economic Profile	B
3	Infrastructure Profile	C
4	E-Governance & Computerization	D
5	Finance	E

6	Environmental	F
7	Progress Track	G
8	Security	H

Table 5 Indian Smart City Indicators Symbols

S. No	Indicators	SYMBOL
1	"Total Population"	A1
2	"Total Population of UA" (if)	A2
3	"Share of ULB population in District Urban population" (%)	A3
4	"Population Growth Rate" (AEGR) 2001-11	A4
5	"Area" (sq. km)	A5
6	"Share of ULB area in the district" (%)	A6
7	"The density of population" (person per sq. km)	A7
8	"Literacy Rate" (%)	A8
9	"Scheduled Caste" (%)	A9
10	"Scheduled Tribes" (%)	A10
11	"Youth, 15 - 24 years" (%)	A11
12	"Slum Population" (%)	A12
13	"Working Age Group, 15-59 years (%)"	A13
14	"Per Capita Income (Rs.) at 2004-05 constant price "	B1
15	"Urban Poverty Ratio" (% of urban population)	B2
16	"Unemployment Rate," 2011-12	B3
17	"Work Participation Rate," 2011-12	B4
18	"Work Status, 2011-12 (%) Self-employed."	B5

19	"Work Status, 2011-12 (%) Regular/wage salaried employees"	B6
20	"Work Status, 2011-12 (%)" Casual labor:	B7
21	"Sectoral Distribution of Workers, 2011-12 (%) Primary."	B8
22	"Sectoral Distribution of Workers, 2011-12 (%) Secondary."	B9
23	"Sectoral Distribution of Workers, 2011-12 (%) Tertiary."	B10
24	"Legislators, senior officials, and managers."	B11
25	"Professionals"	B12
26	"Technicians and associate professionals"	B13
27	"Clerks"	B14
28	"Service workers and shop and market sales workers."	B15
29	"Skilled agricultural and fishery workers."	B16
30	"Craft and related trades workers."	B17
31	"Plant and machine operators and assemblers."	B18
32	"Elementary occupations."	B19
33	"Workers not classified by occupation."	B20
34	"no. of sanctioned SEZ."	B21
35	"% of households with access to tap water (from treated source) within Premises."	C1
36	"% of households with access to electricity."	C2
37	"% of households having toilet facilities within premises."	C3
38	"% of household Waste water outlet connected to the drainage."	C4
39	"% of households with access to computer/laptop with internet."	C5
40	"% of households with access to computer/laptop without internet."	C6
41	"% of households with access to mobile phones."	C7
42	"Ownership Pattern of Housing (%) Owned."	C8

43	"Ownership Pattern of Housing (%) Rented."	C9
44	"% of households living in congested houses."	C10
45	"No. of Hospitals per 1,00,000 people "	C11
46	"No of Schools per 1,00,000 people Primary."	C12
47	"No of Schools per 1,00,000 people Middle."	C13
48	"No of Schools per 1,00,000 people Secondary."	C14
49	"No of Schools per 1,00,000 people College"	C15
50	"Property Tax"	D1
51	"Accounting"	D2
52	"Water Supply & Other Utilities"	D3
53	"Birth & Death Registration and Health programs."	D4
54	"Citizens' Grievance Monitoring"	D5
55	"Personnel Management System"	D6
56	"Building Plan Approval"	D7
57	"e-Procurement"	D8
58	"Does ULB has the facility to Accept Online Payments."	D9
59	"Are ULB offices connected through local area network (LAN)/ wide area network (WAN) "	D10
60	"Do you have access to the State Data Centre (SDC)? "	D11
61	"Does the ULB have their website#."	D12
62	"% of households with access to banking facilities "	E1
63	'Property Tax Coverage (%)"	E2
64	"Property Tax Collection Efficiency (%) "	E3
65	"Property Tax Amount (Rs.)"	E4
66	"Details of municipal income (Rs. Lakhs) 2009-10"	E5

67	“Details of municipal income (Rs. Lakhs) 2010-11”	E6
68	“Details of municipal income (Rs. Lakhs) 2011-12”	E7
69	“Details of municipal expenditure (Rs. Lakhs)2009-10”	E8
70	“Details of municipal expenditure (Rs. Lakhs)2010-11”	E9
71	“Details of municipal expenditure (Rs. Lakhs)2011-12”	E10
72	“% of municipal Budget reserved for urban poor.”	E11
73	“Swachh Sarvekshan was conducted between 5 January 2016.”	F1
74	“Comprehensive Environmental Assessment for available cities”	F2
75	“Total Approved Cost of projects (Rs. Lakhs)”	G1
76	“Share of Central Assistance released (%)”	G2
77	“% of work completed (Physical Progress)”	G3
78	“Funds Utilised (%)”	G4
79	crime index	H1
80	safety index	H2

Table 6 Indian Smart City Data Set Part 1

City	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	B1	B2	B3	B4	B5
C1	312538	443028	24	1	57	1	5449	81	9	1	20	36	68	37712	5	5	37	45
C2	1728128	1728128	85	6	514	5	3365	82	8	1	19	45	68	50580	6	4	37	26
C3	957352	962334	92	2	219	23	4370	91	6	5	19	3	71	43278	9	4	32	47
C4	257803	16349831	0	2	43	0	6032	90	19	0	22	8	74	112510	1	2	37	19
C5	5577940	6357693	92	5	469	6	11895	88	11	1	19	4	67	0	7	1	37	51
C6	4467797	4591246	92	6	336	7	13304	88	2	3	22	10	68	0	6	0	42	37
C7	488157	610350	40	2	100	1	4901	90	8	3	18	11	66	28856	32	7	35	48
C8	434971	0	69	2	77	1	5640	85	12	6	21	14	67	30219	24	4	32	47
C9	602046	2119724	27	0	107	0	5620	97	3	0	15	1	66	63599	4	4	42	32
C10	1798218	1886100	94	2	286	10	6290	83	13	3	21	27	65	47214	10	2	36	33
C11	1964086	2170295	81	3	172	4	11393	86	15	3	20	30	65	52501	3	2	29	56
C12	1055525	1268848	73	1	153	3	6920	87	14	4	20	46	67	38968	17	0	33	40
C13	951558	0	68	1	179	1	5329	83	15	2	21	18	64	45859	19	2	41	45
C14	3124458	5057709	54	2	276	2	11304	90	13	1	19	22	67	88341	3	3	35	28
C15	840834	885363	78	3	135	5	6228	92	8	5	20	20	70	33312	5	4	36	40
C16	1618879	0	78	1	159	4	10158	86	14	0	20	15	67	51633	10	0	38	34
C17	3046163	0	88	3	485	4	6285	83	13	4	21	11	65	3046163	6	4	35	42
C18	451100	474531	74	1	57	0	7925	90	10	5	19	14	67	24135	16	0	39	46
C19	4646732	8653521	100	1	175	100	26553	90	17	0	17	29	69	57706	2	2	41	26
C20	1050721	2136916	40	1	106	2	9950	91	10	0	16	12	68	65781	4	3	41	33
Optimal	257803	16349831	100	0	514	100	3365	97	2	0	22	1	74	3046163	32	0	42	56

Table 7 Indian Smart City Data Set Part 2

city	B6	B7	B8	B9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	B21	C1	C2	C3	C4	C5	C6	C7
C1	44	11	10	28	62	18	10	6	6	10	5	18	10	17	0	4	83	97	83	84	6	9	57
C2	55	19	2	33	65	11	7	10	4	17	2	24	11	15	0	11	71	97	84	92	9	11	67
C3	48	4	16	19	65	1	15	5	4	19	6	28	3	15	4	0	33	93	85	77	12	22	79
C4	77	3	0	41	59	16	0	6	12	9	0	17	25	15	0	1	98	99	85	97	28	16	63
C5	46	4	2	48	50	16	5	6	6	17	1	22	19	7	0	7	80	99	92	93	10	11	63
C6	57	6	0	66	34	28	2	6	1	6	0	14	34	8	0	1	81	99	94	96	5	8	71
C7	34	18	26	24	50	19	5	9	10	7	20	13	4	13	0	1	90	98	91	89	9	13	62
C8	40	12	15	22	63	28	8	3	4	15	14	13	4	11	0	0	82	97	82	95	5	13	65
C9	45	23	5	31	64	16	12	9	9	14	3	14	7	15	0	14	96	99	95	87	21	15	54
C10	53	14	5	25	70	30	10	9	15	7	0	9	3	17	0	1	79	97	82	92	10	13	61
C11	40	4	1	26	73	39	9	3	9	9	1	15	11	4	0	7	74	98	90	92	11	13	63
C12	49	11	1	25	74	17	3	7	2	28	0	18	16	9	0	3	70	96	81	91	8	12	57
C13	28	27	21	27	51	19	5	5	2	8	14	23	11	13	0	0	97	94	59	91	5	9	61
C14	66	6	0	32	68	21	19	6	8	12	0	13	12	10	0	20	99	98	76	98	22	17	64
C15	45	15	1	28	71	15	10	12	4	25	0	29	3	2	0	3	63	87	72	69	16	15	65
C16	64	3	1	51	48	8	5	9	1	15	1	20	18	22	0	0	81	99	93	89	12	11	65
C17	47	11	6	45	49	13	12	6	3	13	7	22	9	15	0	8	89	98	89	91	12	14	66
C18	47	7	8	29	64	13	10	10	5	21	2	19	8	11	0	1	83	97	83	84	6	9	57
C19	66	9	1	27	71	19	13	10	3	11	1	16	13	14	0	12	71	97	84	92	9	11	67
C20	55	11	9	45	46	24	7	6	5	10	4	26	8	10	0	4	33	93	85	77	12	22	79
Optimal	77	3	26	19	34	39	19	12	15	28	20	29	34	22	0	20	99	99	95	98	28	8	79

Table 8 Indian Smart City Data Set Part 3

City	C8	C9	C10	C11	C12	C13	C14	C15	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	E1
C1	53	46	31	0	65	42	27	9	0	0	0	0	1	0	0	0	0	1	0	1	54
C2	51	47	33	3	8	7	12	2	1	1	1	1	1	1	1	1	1	1	1	0	64
C3	54	41	23	1	27	11	15	14	1	0	0	0	0	0	0	0	0	0	0	1	84
C4	12	68	37	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	1	89
C5	77	21	36	0	24	0	9	0	1	1	1	1	1	1	1	1	1	1	0	1	71
C6	55	43	49	0	19	0	10	2	1	1	1	1	1	1	1	1	1	1	0	1	54
C7	62	35	28	0	41	33	23	13	0	1	0	1	1	0	0	0	0	1	0	1	71
C8	52	45	34	0	51	45	37	16	1	1	1	1	1	1	1	1	0	1	0	1	53
C9	75	22	6	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	0	1	83
C10	68	27	31	1	43	8	15	12	1	1	1	1	1	1	1	1	0	1	0	1	68
C11	69	28	25	1	25	14	12	5	1	1	1	1	1	1	1	1	1	1	0	1	68
C12	75	22	19	0	21	19	14	4	1	1	1	1	1	1	1	1	1	1	0	1	65
C13	70	29	45	2	29	25	20	15	1	0	1	1	1	0	0	1	0	1	0	1	63
C14	67	31	40	0	32	28	21	9	1	1	1	1	1	1	1	1	0	1	0	1	86
C15	47	43	33	1	31	15	11	9	1	1	1	1	1	0	1	1	1	1	0	1	73
C16	69	28	31	0	15	3	10	5	1	1	1	1	1	1	1	1	0	0	0	1	65
C17	78	20	25	1	59	57	49	12	1	1	1	1	1	1	1	1	1	1	0	1	73
C18	53	46	31	0	65	42	27	9	1	1	0	1	1	1	1	1	0	1	0	1	83
C19	51	47	33	3	8	7	12	2	1	1	1	1	1	1	1	1	1	1	0	1	71
C20	54	41	23	1	27	11	15	14	1	1	1	1	1	1	1	1	1	1	0	1	70
Optimal	78	20	6	3	75	59	56	16	1	1	1	1	1	1	1	1	1	1	1	1	89

Table 9 Indian Smart City Data Set Part 4

city	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	F1	F2	G1	G2	G3	G4	H1	H2
C1	0	1	34	6932	7239	0	6530	6787	0	0	0	0	0	0	0	0	48	52
C2	1	1	0	0	0	131899	0	0	129990	0	5	71	154677	87	69	80	42	58
C3	95	79	0	0	0	0	0	0	0	0	50	0	31611	87	76	107	56	44
C4	85	82	300	194461	169119	137720	193570	152336	130056	0	16	0	0	0	0	0	60	40
C5	1	1	322	161600	189900	223900	115500	132200	160600	0	14	75	257380	85	27	90	40	60
C6	1	1	356	147951	159532	187027	173618	172113	163676	0	6	58	200533	86	55	90	42	58
C7	0	0	16	7467	7702	0	6401	5884	0	0	0	0	0	0	0	0	44	56
C8	0	0	0	44083	35087	33120	47633	36308	33841	33	0	0	3464	98	81	104	42	58
C9	90	86	57	1652	1235	1093	1435	966	809	25	55	75	48712	49	49	0	34	66
C10	1	1	23	100033	0	0	114247	0	0	0	21	0	119431	74	37	63	38	62
C11	1	1	40	116187	151709	134054	117762	151714	134340	0	25	71	68981	77	56	111	60	40
C12	1	1	13	184	272	275	163	236	320	0	63	0	48937	82	58	70	25	75
C13	0	0	0	1626	2004	1584	1626	2004	2584	0	0	0	7199	100	82	0	44	56
C14	0	0	0	324700	233521	203164	324700	184143	219085	25	11	0	390667	81	52	90	44	56
C15	1	1	20	33	37	38	33	38	37	0	24	0	57326	62	17	47	29	71
C16	57	68	61	41364	55512	58078	35803	56818	59743	0	34	82	33924	48	26	50	58	42
C17	17	0	27	39227	36929	42441	0	34224	43193	0	29	67	57808	64	59	104	31	69
C18	0	0	0	7246	5409	7069	6992	5090	5399	0	0	0	0	0	0	0	40	60
C19	1	1	495	94004	102258	136838	82219	99738	85623	0	36	0	388302	79	40	62	43	57
C20	1	1	103	34669	35983	39551	26176	27985	30811	0	18	72	127787	67	63	49	42	58
Optimal	95	86	0	324700	233521	223900	0	0	0	33	63	82	390667	100	82	111	25	75

3.4 Recommender Systems Model for Smart Cities

3.4.1 Introduction

Recommender Systems[132] gained popularity since 2005 when people understood that the Internet meant much more than surfing and socialization. The growth of E-commerce accelerated, and by the end of 2007 itself, its sales accounted for 3.4% of total sales. Customers started finding the world of online shopping, vast and reliable, and organizations saw this as an opportunity to boost their business by getting to know their customers better. The user's feedback collection and analysis were no longer limited to some companies now. Feedbacks marked the valued opinion of the user and could very well use by generalizing users on demographic fronts. Thus, they not only helped organizations in better responsiveness to the same user in the future but also the users with the same hidden interests. It led to the nationwide deployment of collaborative filter out. Content-supported and Knowledge-based testimonial near quickly also became popular for bestowal with available data differently.

The recommendations based on user records should be as personalized as possible so that users feel satisfied and connected. Moreover, their implementation should not have a major impact on responsiveness to the user. Consequently, new techniques are being developed to progress the presence of the recommender system handling large datasets. Various datasets have been made publicly available by companies to find out the best possible techniques to optimize the process. Million Song Dataset by Mismatch with lyrics, large movie rating dataset by Fluster and Movie Lens, product rating dataset by Amazon being some of them. Every year there is a dedicated annual conference (ACM

Recommender System Conference) held on the topic that values new research topics on the topic and sought to find more areas to deploy a recommender system.

As of now, recommender systems may majorly categorize into four classes designed on the approach used for the method, specifically the cooperative filtering system, content-based system, knowledge-based, and, therefore, the hybrid technique system.

1. Collaborative recommendation System: It provides a recommendation of products or facilities to users based on their likeness in demographics or interest with other people whose preferences and personalized data has been present from past interactions.

2. Content-based recommendation System: It provides a recommendation of objects based on their specifications to users. Objects had a similar specification to the ones earlier preferred by a user recommended.

3. Knowledge-based recommendation System: It provides recommendations based on clear knowledge near the item classification and recommendation criteria, i.e., which item should advise in which context.

4. Hybrid Method System: It renews content-based filtering & collaborative-filtering approximate to refer most pursuit recommendations to the user.

3.4.2 Collaborative Filtering

Collaborative Filtering is an initial technique of recommendation and maximum used in E-commerce systems since its creation. It might be separated further by Memory-based Collaborative filtering and model-based collaborative filtering.

Memory-based Collaborative filtering is the idea that users have similarities, and their interests can be predicted by deducing patterns from their earlier interactions and comparing those preferences with other users or inducing possible interests based on their demographic likeness with other users. K-nearest-neighbor prediction algorithm was the first algorithm to be deployed widely in the recommender system. It performs the task, i.e., estimate likeliness of user to buy a particular item by neighborhood selection and computation of weighted average. This approach is then scaled to all items in the domain to begin recommendations to the target user, thus accomplishing the recommended task.

Several extensions to the Pearson coefficient and cosine likely approach suggested and widely calculated over time, e.g., the default voting, case amplification, weighted majority prediction, and imputation-boosted CF, Inverse User frequency, etc. The default voting and significance weighting deal with improving accuracy in prediction if there are only a few numbers of co-rated objects. Default Voting is because it is safe to give an explicit rating to the items that haven't been rated to overcome the problem of sparse rating matrix. This change improves the accuracy of likeness techniques by a large factor. On the other hand, Significance weighting gets rid of the problem by multiplying likeness weight through a Significance Weighting factor, that undervalues the correlations base on some common rated items during recommending task. Case amplification improves performance by giving more importance to neighbors with higher likeness by introducing the concept of amplification factor and transforming original likeness weight equations.

Error procedures such as mean squared error and mean absolute error accesses the accuracy of the forecast ratings by comparing them with actual ratings. Precision (ratio of

relevant and retrieved to retrieved) and recall (ratio of relevant and retrieved to relevant) matrices are other evaluation matrices that widely used.

There are several other likenesses measures available in the literature, with Spearman rank correlation, Kendall's τ mean squared differences, correlation, etc.[133]. Spearman rank correlation is the next most used likeness function after Pearson. It modifies Pearson's likeness correlation by advocating the use of item ranks rather than ratings. In Spearman rank correlation, items ranked according to their ratings with highest rated as lowest-ranked. Items have a similar rating given the mean rank of position, and then similar computation to Pearson correlation is followed. [132] draws light on different variants of neighborhood formation, weighting, and normalization processes, along with their comparisons with other alternatives based on experimental data.

With the advent of big data, the Collaborative Filtering approach expanded to the application of the clustering algorithm on the user-item ratings. It has been done to ensure good results when datasets are very large, and parallel processing effectively exploited by making use of advanced frameworks. The algorithm specifies the clustering of datasets into partitions of a fixed or variable size such that similar users belong to the same partition.

All these techniques based on user-user based correlation, and despite many improvements in algorithm and strategy, it still has got some cons. One of the major upsurges for Collaborative filtering came up when item-item based collaborative filtering came into the picture proposed this approach where the user's valued items matched to parallel items more exactly than matching the target user to corresponding users. Provided

the first major contribution by specifying the algorithm that taken in item-item built collaborative filtering. The strategy was to combine the cosine formula and Pearson coefficient to calculate the object likenesses. Several variants of item-based collaboration have also come up for optimization. The results display that the recommendation algorithms built on similarities of the object perform considerably better than on item-based in most cases.

Model-based Collaborative filtering technique provides recommendations by building and analyzing statistical models for user ratings along with several well-defined parameters. The earliest model-filtering technique systems compared prediction problem to classification problem casting items as characteristic vectors above users and existing ratings as labels to create a multidimensional structure to better relate users along with these dimensions. Dimensionality reduction techniques were used to overcome situations of data scarcity.

To support recommendation systems along with the model-based collaboration, many distinct models came up with the passing time. The techniques which were seen most commonly in this proposition was Latent factor and matrix factorization models. The latest model was a good model as it takes into consideration that the likeness induced parallel between the users and items depending on the lower-dimensional structure, which was not visible in the data. And not like other models like neighborhood-based methods which supported enumerated notions between user's likeness, or between items likeness. For example, a user not only based on likeness to a group of songs gives a rating to it, but other suggested factors are also there like his taste of music. In an approach of choice ratings of

the user concerning some items, vectors are very useful. And these vectors have hidden dimensions that emerged with users and items therein matrix factorization techniques.

The second most common class of technique that was revealed by Lee and Seung was Non-negative matrix factorization. It was given this name because non-negative constraints included. When a user gives a rating, it is seen by the manifestation of different roles. For example, the target user presence was sure by the community in different clusters. Therefore, each user gives ratings which are traced along all the various dimensions to get the sum of users rating to that item.

To summarize, though Collaborative Filtering Technique has its problems and still a lot needs to be evolved, but its advantages are more profound and better express the approach. An efficient recommender system can very well be used to recommend new information to the users citing his/her hidden interests. It can be used to recommend non-structured data such as art, music, etc. These are some of the favorable features that make Collaborate Filtering techniques such as an important class of recommender systems.

3.4.3 Content-Based Recommendation

Content-Based Recommendation System gives a recommendation of an object to the user-built on the content information of objects instead of correlating different items or users based on statistical comments or feedbacks as in the Collaborative filtering approach. The recommendation to a particular user does not have anything to do with the interests of other users of the system. In this approach, user and object profiles are made and continued based on the analysis of these profiles' recommendations are suggested. In this way, these

systems learn function for modeling each user's interests. In music recommendation systems, the common characteristics (tone, beats, lyrics, etc.) of the songs existing in the profile of the target users analyzed and then similar songs recommended.

The advanced algorithms influence be meant higher than the classical data retrieval approaches. They will be adjusted additional swimmingly by reducing weights of corresponding terms or nodes. Naive Thomas Bayes is one such approach to inductive learning and makes use of theorem classifiers for its analysis. It generates a probabilistic model supported applied mathematics knowledge and performs a cluster of the offered dataset. It's two frequently used formulations, particularly the variable Bernoulli and also the multinomial model. Each of these models allocates common values within which text documents are assumed to be generated by associate degree primary generative model, especially a parameterized mixture model. The vital alternative approach makes use of call trees within which internal nodes, branches, and leaves labeled for various classes. During this technique, coaching knowledge is partitioned off into text documents into subgroups until solely instances of one category are a gift in those subgroups. Partitioning majorly involves text classification that supported some features like the presence or absence of a personal word or phrase within the known word. It creates associate degree organized structural read for a far better understanding of user and object profiles. This system is generally used for tiny trees with few tests to avoid a performance hit. Alternative adept ways are relevancy technique cluster analysis, neural networks, etc.

Content-Based Recommendation System also faces the problem of increased complexity when scaled to large datasets. The optimal threshold algorithm overcomes this

problem. The algorithm is an improved form of adaptive filtering technique which analyses existing data and likeness distribution of user profile to decide threshold value and files having likeness over this threshold are only familiar inform user profiles. This technique was greatly growing the system's correctness and efficiency. These systems, like Collaborative filtering systems, also face the problem of accuracy-speed trade-off. Effective human-machine interaction is one remedy to this problem. They also face linguistic problems as user profiles with words from different languages need to handle. Offered an adaptable language method to improve user profile compatibility.

To summarize Content-Based Recommendation is a relatively new recommendation approach but has well-defined areas of application. Unlike the Collaborative filtering technique, it could not face the cold-start problem and is more effective in the case of sparse datasets. Moreover, it can even recommend new objects to the target user by exploiting likeness in the user’s profiles and object specification.

3.4.4 Hybrid System

Both Collaborative Filtering techniques and Content-Based Techniques are having some cons. The former face problems like early rater and sparsity of datasets while subjective domain and expressive content description are some of the weaknesses of the latter. So, most real systems combine these approaches to overcome the shortcomings of each and provide the best results. There are many methods employed in this process, like weight, switched, feature combination, and cascade.

In the weighted technique, predictions by different recommendations approach available in the system are combined linearly after independent computation to provide optimal suggestion in switching technique, and many recommendation techniques deployed in the system. Hybridization strategy is made sensitive to a confidence level of results given by one technique, and accordingly, the decision is made to switch to other techniques. In the feature combination technique, collaborative information is treated as basically extra information data related through each record and content-based methods used over these augmented data sets. In the cascade technique, hybridization done in stages. Firstly, one method is working to produce coarse dataset speciation, and then the second method is employed optionally depending upon the differentiation level of previous results. Hybrid systems can also deploy any of these methods to merge different recommendations and data analysis techniques. There are many models based on this approach.

To summarize, a hybrid recommender system is most reliable and fast as they created for specific purposes based on the situation by combining the best possible choice of available approaches or techniques. They generally have specialized algorithms that deployed in these systems for improving accuracy as well as speed.

3.4.5 Implementation

Different kinds of recommender systems are used extensively for specialized purposes. Collaborative filtering can be put to use in many applications. We are at this moment discussing how it can be used to make a movie recommender system. We are using Movie Lens dataset for this purpose, which downloaded from their official website. The dataset comprises of many files of which we have used “movies.csv” and “ratings.csv”

majorly to create a simple application that suggests similar movies to a user based upon the historic rating records. Recommending movies similar to a particular movie was done using Item-based Collaborative filtering. The likeness between movies was found out by using the Pearson Coefficient relation. Map Reduce programming technique was used to create the application. Movies were input from the user and then likeness of input movie with every other movie in the dataset was calculated and the top 25 results were filtered and displayed to the user. Internally the process performed the mapping from movie name to movie Id or vice versa. Four maps reduce tasks that were performed by application for the purpose. These are:

Mapper 1

Scan the records of the dataset ratings.csv and divide it into key-value pairs.

Input- Record(uid,mid,rating,timestamp) where uid=unique Id of user, mid=Id of movie

Output- key(uid) value(mid,rating)

Reducer 1:

Input the sorted keys and list of corresponding values and output each key-value pair.

Input -key(used) value (List<value (mid, rating)>) Output-key(uid) value (mid, rating)

Mapper 2:

Scan the records from output file of the first job and divide value into a pair of items, each having the desired target item (whose likeness needs to be found out) and every other item rated by the given user.

Input- Record(uid,List<mid,rating>)

Output- key(midn,targetmid) value(ratingn,targetrating) where midn=movie id of nth movie, targetmid=movie id of target movie rating and target rating-rating given by user to nth movie and target movie respectively.

Reducer 2:

Input all the movie pairs and their corresponding ratings and output the Pearson correlation between these movies.

Input- key (midn, targetmid) List<value (rating, target rating)>

Output-key (mind, targetmid) value (Pearson coefficient)

Mapper 3:

Scan the records from the output file of the first job and divide it into key-value pairs with keys being the Pearson correlation and output being an nth movie.

Input- Record (midn, targetmid, Pearson coefficient)

Output-key (Pearson coefficient) value(midn)

Reducer 3:

Input all the key-value pairs from preceding mapper sorted concerning key and obtained the value of the last 25 pairs and prepare the corresponding list for reducer 4.

Input- key (Pearson coefficient) value(midn)

Output-key(midn) value (Pearson coefficient) of the last 25 pair

Mapper 4:

Scan the records of the dataset movies.csv and divide it into key-value pairs.

Input- Record(id,name,genre) where id=id of movie,

name= name of the movie, genre=genre of movie

Output-key(id) value(name)

Reducer 4:

Input key value from Mapper 4 and a list of 25 items from mapper 3. Find the name of the corresponding movie with its presence on that list.

Input- key(id) value(name)

Output-key(index) value(name) for items with its present in list where index

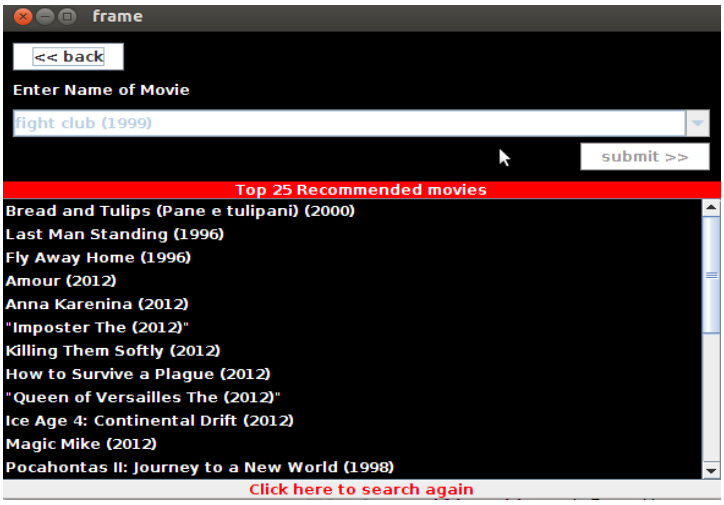


Figure 3-2 Recommender Systems Model for Smart Cities

This simple application can, therefore, be used by movie enthusiasts to know the movies that would like if they enjoy a movie. It may also be worked to predict movie ratings by a target user. The output of the application for the popular movie “Fight Club” shown in Figure 3.2.

In this way, this application can be used by movie enthusiasts to make smart decisions about whether to watch a movie or not.

3.4.6 Conclusion

This topic analyzed different approaches to a recommendation system model for smart cities, each having its pros and cons. The best recommendation approach is found out to be the use of a Hybrid System that can be customized and scaled according to the problem. They programmed for better functioning concerning speed-accuracy trade-off.

The domain of the recommender system is expanding at a fast pace with new possible optimizations getting suggested every day. The topic has been written with the utmost care to mention all the presently deployed techniques regarding the subject but may need to be reviewed in the future to mention further advancements in this field.

3.5 Adaptive Whale Optimization Algorithm

3.5.1 Introduction

As another kind of useful improvement, the idea "Smart City" knows a brilliant development among the current years [134]. The slight and extensive districts are proposing another city show called "Smart City," which addresses a get-together of typical advancement gauge, appealing and secure, pleasant, interconnected, and doable[135]The imaginative Internet of Things arrangements is engaging Smart City rehearses in the general world. It engages to control contraptions and oversee, remotely screen, and to make novel experiences and significant data from gigantic floods of relentless information [65]The essential attributes of a splendid city join an unpredictable state of data improvement bargain and a far-reaching utilization of data assets [113]. The crucial pieces

of urban change for an astonishing city should meld the insightful business, sharp advancement, sharp association, astute life, and awe-inspiring associations [136].

Each piece of the economy of countries changed by utilizing smart city huge information. Such change connects with urban networks to complete the learning standards and necessities of the jobs of the sharp city by understanding the basic sharp condition characteristics[137]The smart city misuses favored point of view of rising headways, for example, remote sensor arrange, to constrain resource usage and cost [138]. The rising progressions among one with the gigantic plausible to improve keen city organizations are tremendous data examination [139]. At the show, a lot of data is being produced using distinctive data sources like PCs, propelled cell phones, cameras, sensors, relational cooperation goals, overall arranging systems, preoccupations, and business trades [140]. The dataset made in the present universe of digitizing makes amazing data storing and dealing with workplaces has exhibited difficulties in the consistent stages and conventional data mining [141]. The utilization of gigantic data in a keen city has various focal points and troubles, together with the transparency of generous numerical and storerooms to practice surges of data conveyed inside a splendid city condition [142].

Information Computation and Processing Phase (DCPP) [66] takes after nine component highlight methods for fruitful part extraction. They are DIA association factor, Chi estimation, Information Gain, Document Frequency, OCFS, DPFS, Comprehensive Measurement Feature Selection, Mutual Information, and improved Gini record [143][144]. The tremendous test here is we have to recognize which procedure is appropriate [66] for especially perceiving the highlights.[145] To overcome this test, we

have required a suitable methodology in light of perceived features for the perfect utilization of assets.[146]

3.5.2 Feature Selection Based Whale Optimization Using Wrapper Based Method

The way toward disposing of the insignificant features and excess features from the database is known as feature choice, which is used to improve the learning algorithm. The two criteria, named as assessment and search criteria feature collection approaches arranged. By utilizing the wrapper and channel approaches, the chose feature subsets assessed. The algorithm of Learning utilized as a part of the wrapper strategy in the determination procedure [147].

Big data smart cities[148] sensible utilization of benefits is the crucial objective of our proposed methodology. For this movement, various improvement procedures used. In our proposed article, the whale upgrade procedure is utilized to get a sensible utilized perfect rundown of capabilities [149] By using this perfect skill rundown of capacities to be picked and decline the computational period; furthermore, the estimate organization of measurements has the favored characteristics over the present techniques.

3.5.3 Wrapper Method:

The component choice based on the novel Whale Optimization Algorithm is working dependent on wrapper-based technique. The use of the classifier as the controller of the component choice framework is the preeminent element of the wrapper procedures. The ensuing three principle plans are utilized in the component determination in wrapper [150]

Classification Scheme

Criteria of Feature Evaluation

Search Scheme

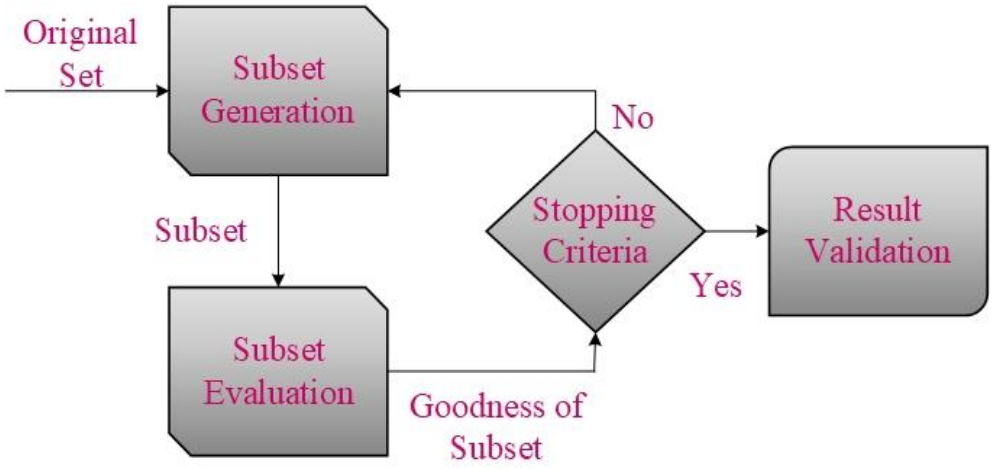


Figure 3-3 Process of Feature Subset Selection

A Subset Generation is a heuristic interest process Figure 3.3 where a search for space contains states, every last one of which exhibits a bright subset for assessment. Two things must be agreed to subset age, Search beginning stage, and Search structure. Chase's beginning stage can be forward, in reverse, bi-directional, and capricious. An intriguing system must pick the contender subsets [151].

An Evaluation Display is utilized to overview each starting late made applicant. In the context of the reliance on learning counts that will be related to the picklist of capabilities, an assessment show organized into two get-togethers, one is destitute criteria. The second one is self-decision criteria. Wrapper shows use subordinate criteria, and for highlight decisions, it needs a learning calculation. Reasonableness of highlight or its

subset is surveyed with the assistance of basic highlights of the course of action information without partner any learning calculation.

In incorporate subset choice, a halting model addresses when the component choice procedure will stop. A touch of the routinely utilized ending criteria are as indicated by the going with:

Exhaustive search finishes.

If a dynamic extension or departure of any component doesn't impact comes about part assurance procedure could be stopped.

If agreeably decent subset chosen.

Toward the end, it comes about are approved by utilizing grouping blunder rate of classifiers as an execution marker. Examinations are led to compare the characterization error rate on the full arrangement of the classifier learned on highlights, and that prepared on the chose include subset [152].

Wrapper based part subset evaluation methodologies brief learning estimations during the appraisal to Figure 3.4 the conventionality of a picked incorporate subset given the count's accuracy so are computationally exorbitant when appeared differently about channels. To the extent farsighted or course of action, precision wrapper procedures saw as superior to the channel.

The used component request criteria or target work in the wrapper feature assurance is consistently reflecting the gathering execution and the number of features. A vague

depiction of the health work addressing for both request execution and the quantity of picked incorporates as depicted in condition [153].

In wrapper-based techniques, the single assessment of guaranteed arrangement is exorbitant as it generally applies preparing and testing of the given more tasteful. Along these lines, the proficient choice of the look technique is fundamental.

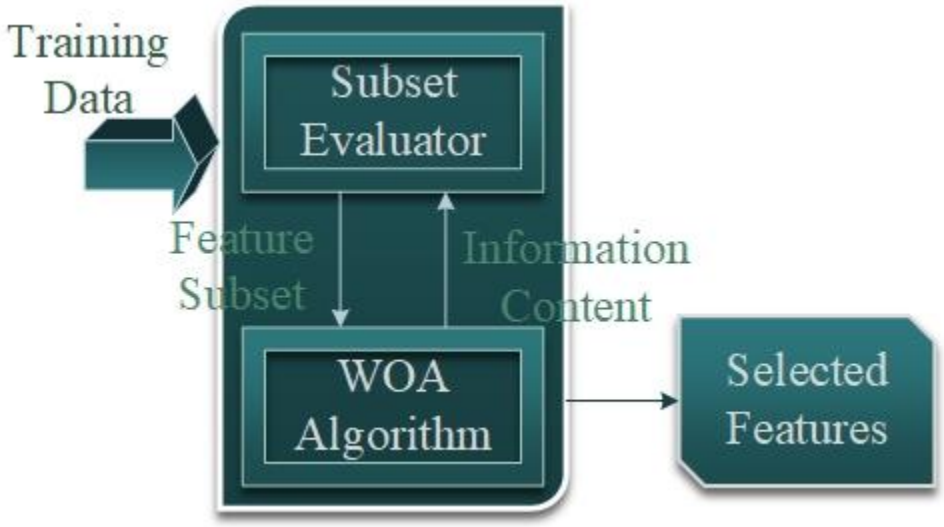


Figure 3-4 Wrapper Approach for Feature Selection

The wrapper is a moderate segment assurance procedure. According to the dependence standard grasped by the wrapper show, it requires a predestined acknowledgment estimation and usages its execution measure to be associated with the picked subset to make sense of which features picked. It all around gives better execution as it picks incorporates progressively able to the destined acknowledgment figuring. Be that as it may, it is computationally progressively exorbitant and moderate, which may not be fitting for other acknowledgment estimations now and again.

The features picked by the classifier are used to envision the class characteristics of concealed events; precision is regularly the evaluator of the created subsets. Nonetheless, it cost high to figure accuracy for every component subset. Subsequently, in the wrapper method, it is basic to separate the decision technique most proper for the given dataset and portrayal procedure. Besides, the subset of features picked is the irrelevant containing all the relevant features for the goal class.

Figure 3.5 demonstrates the wrapper-based element choice methodology that the preparation set delivered to the subset evaluator. The dataset is too huge, so the set is given to the streamlining calculation to assess the ideal dataset. The decreased dataset worth is prepared under the connection procedure to get the ideal list of capabilities. The most important information considered as the ideal setting and different qualities offered back to the subset evaluator.

3.5.4 WOA

Here the WOA is said to be a fresh optimization algorithm. The WOP algorithm used to mimic the Humpback whales' natural behavior. Hunting behavior is the main character of these whales in the path of survival. This technique is to address the optimization problems, and the hunting strategy has been introducing and to chase the prey in the search space; the capability to employ the best agent or random is the whale algorithm's uniqueness. By using spirals, humpback whales bubble-net attaching [154]

The exhibiting of this computation consolidates three executives to emulate the output for prey (examination organize), the encompassing prey, and the air pocket net

looking (misuse arrange) lead of humpback whales. The numerical definition is shown and explained as takes after:

Prey Encircling: By utilizing the best chase administrator, whale estimation starts here. The present exercises are the best, and whether it is, the zone of the prey or close to this anticipates it. Whatever remaining parts of the administrators along these lines invigorate their regions toward the best interest master.

$$\vec{P} = \left| \vec{B} . \vec{x}^{*}(t) - \vec{x}(t) \right| \qquad \mathbf{1}$$

$$\vec{x}(t+1) = \vec{x}^{*}(t) - \vec{A} . \vec{P} \qquad \mathbf{2}$$

Here, a recent recurrence of the method signified as *t*; coefficient vectors signified as $\vec{A} and \vec{B}$. The best resolution in the position vector signified as $\vec{x}^{*}$, and the position vector is $\vec{x}$. x^{*} must be updated iteratively for the actuality of a better result. $\vec{A} and \vec{B}$ Vectors measured in the following expression:

$$\vec{A} = 2\vec{a} . \vec{r} - \vec{a} \qquad \mathbf{3}$$

$$\vec{B} = 2 . \vec{r} \qquad \mathbf{4}$$

where, above the numeral of iterations from 2 to 0, $\vec{a}$ is linearly reduced and random vector in [0, 1] signified as *r*.

The above showing gives some administrator to invigorating that one region in the region of the present best course of action likewise copies encasing the person in question. For n estimations, it also energizes the look space. Additionally, the experts empower the hypercubes improvement utilizing the about the best course of action achieved.

Exploitation Phase: It is otherwise called bubble-net assaulting. Likewise, two strategies are utilized to work the errand:

Mechanism of Shrinking encircling: here, based on the eq 3, $\vec{a}$ value is reduced, and subsequently, the changeability range $\vec{A}$ is also reduced by $\vec{a}$. it indicates that $\vec{a}$ arbitrarily positioned in [-$\vec{a}$ $\vec{a}$]. Here over the optimization time from 2 to 0, the value of a is reduced. Where, the unpredictability of $\vec{A}$ in [-1, 1], determine the search agent's new location next to the best current location and past agent location.

Spiral updating position: The distance amongst the locations of whale and its prey is estimated. By using humpback whales of helix shape, the spiral equation is formed among prey locations, also whale to enhance the movement. It exposed in the following equation:

$$\vec{x}(t+1) = \vec{P}' . e^{bl} . \cos(2\Pi l) + \vec{x}^{*}(t) \qquad \textbf{5}$$

$$\vec{P}' = \left| \vec{x}^{*}(t) - \vec{x}(t) \right| \qquad \textbf{6}$$

The i^{th} whale distance is given in Eqn 6, and the logarithmic spiral shape constant denoted as *b,* and in the region of [-1, 1] random number is denoted as *l*. It has a whale movement in the direction of its prey is simultaneously spread over the shrinking circling, and it placed in a spiral-shaped path. Consequently, the whale's next position updated by the *50%* postulation of the opportunity to control between the binary modes.

$$\vec{x}(t+1) = \begin{cases} \vec{x^*}(t) - \vec{A}.\vec{P} & if \;\; p < 0.5 \\ \vec{P'}.e^{bl}.\cos(2\Pi l) + \vec{x^*}(t) & if \;\; p \geq 0.5 \end{cases} \qquad \textbf{7}$$

where *p* is a random number in [0, 1].

Algorithm: 1

Input:

$D(f_1, f_2, ..., f_{|M|})$ // *a dataset D with* $|M|$ *features*

S_0 // *an initial subset of features*

δ // *a stopping criterion+*

Output:

S_p^* // *a* $(sub) - optimal set of$ *features*

Begin

Initialize $S_p^* = S_0$;

$J(S_p^*) = Evaluate(S_0, D, L)$; // *Evaluate goodness of fit of* S_0 *on* D *w.r.t learning task* L ***Step 1:***

Start the Whole Population $X_i (i = 1,2,3,...n)$**.**

Step 2: **Calculate the fitness of each whale.**

Step 3: **Set the best whale is X*.**

Step 4: **While (*t< maximumnumberofiterations*) do**

for (*eachsearchwhale*) do

Update *a*, A, C, *l* and *p*.

if (*p < 0.5*) then

if (|A| < 1) then

Updating whale position is by Eq. (1).

else

if (|A| ≥ 1) then

Choose the random whale X_{rand}.

Updating whale position is by Eq. (9).

end

end

else

if (*p* ≥ *0.5*) then

The whale position is modifying by the Eq. (5).

end

end

end

Verify if any search agent goes beyond the search space and amend it.

Calculate each search agent's fitness.

If there is a better solution, update X^*.

t* = *t* + *1

end

while δ not reached

$S_i = Generate(D)$ **;** **//Generate a new feature subset** $S_i \subseteq D$

$J(S_i) = Evaluate(S_i, D, L)$; // $Evaluate new subset S_i$

if $J(S_i) is better than J(S_p^*)$ // $Update optimal feature subset$

$$S_p^* = S_i$$

$J(S_p^*) = J(S_i);$

end while

return S_p^*

Validate

$(S_p^*, V, L);$ // *validate* S_p^* *on dataset* V *w.r.t learning task* $L;$

end

Figure 3-5 Algorithm of WAO with Feature Subset Selection

Exploration Phase: A global optimization is attaining in WOA in this phase. To accommodate the search agent $\vec{A}$ is set casually from [-1, 1] and which is used to retreat as of the reference whale. It means the value of $\vec{A}$ ranges from either less than the -1 or higher than 1. Search agent location simplified through choosing an agent randomly, and that permits WOA to execute the global search.

The mathematical expression of the displaying of this assessment appliance given below:

$$\vec{P} = \left| \vec{C} . \overrightarrow{x_{rand}} - \vec{x} \right| \quad \textbf{8}$$

$$\vec{x}(t+1) = \overrightarrow{x_{rand}} - \vec{A} . \vec{P} \quad \textbf{9}$$

Where random location for the arbitrary whale denoted as $\overrightarrow{x_{rand}}$ and that preferred since the present residents. The WOA pseudo-code presented in Algorithm 1.

3.5.4.1 Feature selection of WOA:

The whale advancement calculation approach used for comprises of the ideal in a wrapper-structure system. The critical standard for the wrapper-based procedure is utilized to put on the request for philosophy as a manual for the assurance system incorporates some upgraded parts that picked the list of capabilities. This article applies the WOA adaptively for the perfect component subset to find that growing request execution. With a fundamental best interest administrator, the whales interminably adjust their territories to any point in the space start the appropriate administrator in the WOA. As in conditions (1) and (2) their work to invigorate the circumstances toward the best interest authority from that forward point.

With a comparable estimation in the dataset, the individual course of the action depicted as a consistent vector. The game plan vector regards constrained to [0, 1] and steady. While the estimations of plan the wellness, count are binarily implied. An associated wellness work is regularly organizing a few picked highlights and the gathering execution. It addressed the accompanying condition.

$$f_\theta = \beta . H + (1-\beta)\frac{\sum_i \theta_i}{N} \qquad \textbf{10}$$

Based on the above equation, f_θ it signified as a function of fitness specified a vector θ, N sized with *1/0* components, which is elaborating the unselected or selected

feature subset. In the database, the whole number of features described as N. From the given selected feature subset, H represented as a classification error. β is a constant. Finally, the constant handling of the trade-off amongst classification errors to feature subset quantity. The organization enactment is the main objective, so $\beta=1$ use in this work.

3.5.5 Regression Model Random Forest with Support Vector Machine

3.5.5.1 Random Forest:

This system includes an outfit of different backslides trees in Figure 3.6. Irregular backwoods relapse is a tree-based technique that incorporates stratifying or dividing the indicator space into different clear areas. The mean of the response estimations of the readiness observations in a comparable zone is normally associated [155].

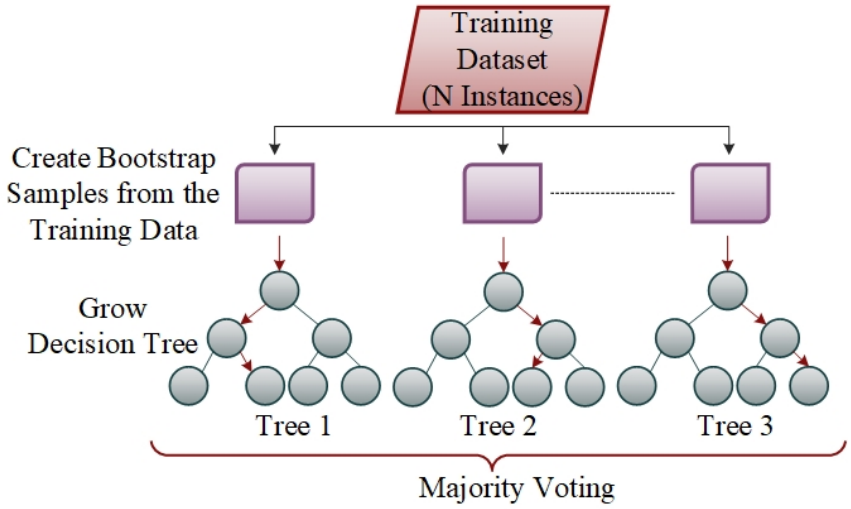


Figure 3-6 Random Forest Pictorial Representation

It is a piece of programmed learning technique: "Bagging" and irregular subspaces ideas consolidated by utilizing this calculation. The choice tree forest calculation prepares on numerous choice trees driven on somewhat unique subsets of information.

The arbitrary backwoods are a piece of the set strategies that takings the choice tree as a separable indicator and depend on the techniques for Random Subspace pardoning boosting, Randomizing Outputs, and Bagging [156]. The calculation of random forest is truly outstanding among characterization calculations - ready to order a lot of information with precision. It is an outfit learning technique for characterization and relapse that builds various tree choice at preparing period and conveys the class that is the method of the programs yield through singular trees.

Algorithm: 2

For b = 1 to B Make

Draw *Z of size *N*.**

Construct a random forest tree R_f to the initial condition, through the repeatedly reconstructing the associated steps for every end child of the tree, till the point that the base hub evaluates n_{min} is makes to a minimum.

From the *p* variables, select the m variables randomly.

Among the*m*, pick the finest variable/split-point.

Divide the node

Output the group of trees $\{T_{b_1}^{B}\}$.

Figure 3-7 Algorithm of Random Forest

The prediction making at the novel point x is denoted by:

Regression:

$$\hat{F}_{rf}^{B}(X) = \frac{1}{B}\sum_{b=1}^{B} T_b(X) \qquad \textbf{11}$$

Classification:

The b[th] random forest tree prediction is assumed $\hat{C}_b(x)$.

$$\hat{C}_{rf}^{B}(X) = majority\ vote\left\{\hat{C}_b(X)\right\}_1^B \qquad \textbf{12}$$

In arbitrary timberland plan procedures, various classifiers delivered from tinier subgroups of the data and late their particular results are gathered in light of a democratic instrument to make the pined for the yield of the educational information file. This outfit learning framework has, starting late, ended up being especially common. Already from the RF, Bagging, Boosting, and were the primary two gathering training methodologies used. RF has been comprehensively associated in various zones together with quiet present-day revelation, sort out interference recognizable proof, show up spread assessment, FICO appraisal examination, remote identifying and quality microarrays data assessment, etc. [157]

Two methodologies are utilized to evaluate the goof extent. They are the preparation and testing parts. The readiness part utilized, and to calculate the bungle rate, the test part-use. An elective way is to use the Out of Bag (OOB) goof evaluate. Since

unpredictable woodlands estimation figures the OOB botch during the planning stage, we don't need to part the readiness data Figure 3.7.

3.5.5.2 SVM:

It is an assessable gathering approach and relies upon the increase in edge among the events, additionally the hyper-plane segment. The method saw as the best substance course of action of execution. It is said to be a non-probabilistic matched direct classifier, which can straightly detach the classes by a considerable edge, it ends up a champion among the most extreme classifier prepared for dealing with boundless dimensional part vectors.

SVM is an organized AI calculation that can be utilized for both falls away from the faith undertakings and get-together. Notwithstanding, it is, generally, used as a couple of depiction issues. In this check, we plot each datum thing in n-dimensional space point with the estimation of each portion being the estimation of a particular genius. By then, we execute collecting through disclosure the hyper-plane and that unmistakable the twofold classes to a mind-boggling degree well.

It has shown compelling in portraying supposition files, including style. The rule for assistance vector machine estimation is to handle plan and backslide issues. It has associated with various fields. SVMs created during the 1990s, given the theoretical examinations of Vladimir Vapnik on the progression of a quantifiable theory learning is said to be Vapnik-Chervonenkis speculation. SVMs were quickly gotten gone for their

ability to work through tremendous data, the humble hyper requirements number, their speculative accreditations, likewise their incredible results eventually.

A linear classifier is measured to make the SVM is less demanding for twofold arrangement issue with highlights x and y names. $y \in \{-1,1\}$ Is used to designate the constraints w, b and class labels:

$$f(x) = w^T x + b \qquad \textbf{13}$$

where,

The normal to the line denoted as w. Bias displayed by b.

It has spoken to via an isolating the hyperplane Figure 3.8 of *f (x)* in the sense of geometrically separates the information space into two assorted districts subsequently bringing about the grouping of the info information space into two classifications [87]:

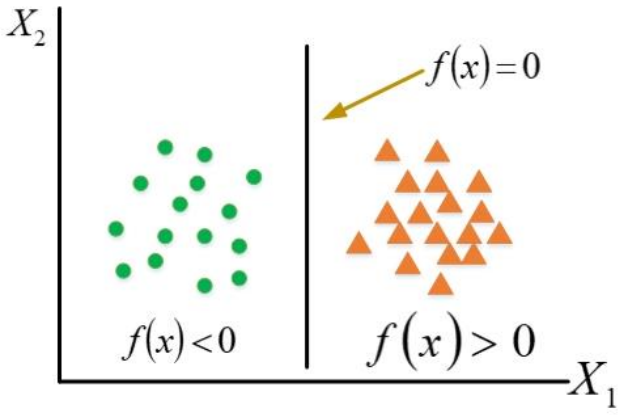


Figure 3-8 Separating the Hyperplane f(x)

The capacity *f(x)* means that the isolation of hyperplane in two locales also encourages in a grouping of the informational index. Hyperplane creates two locales

symmetrically and compares to two classifications of information underneath two class names. An information point "a" has a place with both of the areas relying upon the estimation of *f (a)*. On the off chance that *f (a) > 0*, it has a place with one district, and if *f (a) < 0*, it has a place with another region.

Accept that the information comprises of n information vectors where every datum vector is spoken to by $x_i \in Rn$, where *i=1, 2… n*. The class has given, and a chance to mark that should exist allocated to the information directions to execute managed order stay signified using *yi*, which is *+1* aimed at information vectors of one classification and *-1* aimed at the further classification of information vectors. By using hyperplane, the informational collection can be geometrically isolated. Subsequently, the hyperplane is spoken to by a line it can likewise scientifically speak to the following [156]:

$$w^T x_i + b \geq +1 \quad \textbf{14}$$

$$w^T x_i + b \leq -1 \quad \textbf{15}$$

The hyperplane mathematical formation indicated below:

$$f(x) = \text{sgn}\left(w^T x + b\right) \quad \textbf{16}$$

Where, sign function indicated as *sgn ()*, which scientifically characterized through the equivalence, as shown below :

$$\text{sgn}(X) = \begin{cases} 1 \; if \;\; X > 0 \\ 0 \; if \;\; X = 0 \\ -1 \; if \;\; X < 0 \end{cases} \quad \textbf{17}$$

The following condition gives the separation D. It is from the hyperplane to the point is spoken to scientifically:

$$D = \frac{\left| w^T x + b \right|}{w} \qquad \textbf{18}$$

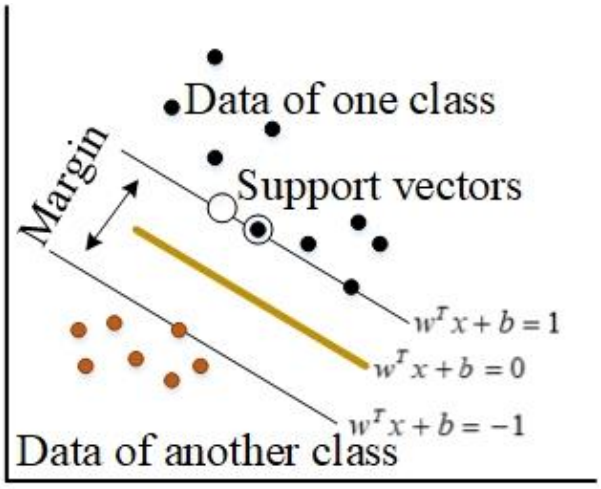


Figure 3-9 Support Vector Machines Hyperplane

The margin is displayed by:

$$\frac{W}{\|W\|} \cdot \left(X_+ - X_- \right) = \frac{W^T \left(X_+ - X_- \right)}{\|W\|} \qquad \textbf{19}$$

$$= \frac{W^T \left(\left(\frac{+1-b}{W^T} \right) - \left(\frac{-1-b}{W^T} \right) \right)}{\|W\|} = \frac{2}{\|W\|} \qquad \textbf{20}$$

Many hyperplanes use in this optimization, and these hyperplanes have split the data into two regions. But the hyperplane is selected vu the SVM Figure 3.9. In the two regions, the hyperplane is the extreme distance from the nearby data points. But few

hyperplanes are there, and that should fulfill this condition. Finally, the accurate classification results are provided by the SVM by ensuring that criterion[158].

The means took after while utilizing SVM in arranging information are specified in the beneath algorithm Figure 3.10:

Algorithm: 3

I/P: I: **Given data**

O/P: **V: Maintenance vectors set**

Start

Separate the given data set into two sets of data items

To support the vector set, V then add them

The divided n data items are then loop

If a data element did not allocate to any of the class labels means then add that to set V.

If inadequate data substances are created then break

end the loop

Train and test using the SVM classifier

End

Figure 3-10 SVM Algorithm

3.5.6 Experiments Results

3.5.6.1 Dataset:

CityPulse of EU FP7 Assignment [159] incorporates a couple of SC applications in light of IoT. Inside the degree of the contamination, task, street traffic, atmosphere, social, stopping, and library data accumulated beginning the urban territories of Brasov and Aarhus in Denmark and Romania independently in the region of 2015 and 2013. In this examination, the CityPulse EU FP7 contamination dataset Task is used to comprehend the novel structure. The dataset has eight highlights, for example, particulate issue, ozone, sulfur dioxide, carbon monoxide, nitrogen dioxide, timestamp, longitude, and scope used for investigating. Seventeen thousand 500 sixty-eight models displayed in the dataset and gathered at five-minute interims. Each test regard set as standard of EPA's AQI. In this assessment, nitrogen dioxide and ozone poisons picked as desires for air quality.

3.5.6.2 Thresholds:

In this part, by thinking about the AQI (Air Quality Index) basic level (100), its levels are isolated into three. Because of the caution shading, the limit esteems characterized, and it portrayed in the beneath the Table 10.

As per these limit esteems, the test datasets of nitrogen dioxide and ozone freely named Healthy, Medium Healthy, and Unhealthy. What's more, it specified in Table 11.

Table 10 Threshold value for AQI

Index value	Level	Alarm Colour
0 - 50	Healthy	Green
51 – 100	Medium Healthy	Yellow
101 - 500	Unhealthy	Red

Table 11 The Alarm based Decision Table

Ozone	Nitrogen Dioxide	Alarm based Colour
Healthy	Healthy	
Healthy	Medium Healthy	
Healthy	Unhealthy	
Medium Healthy	Healthy	
Medium Healthy	Medium Healthy	
Medium Healthy	Unhealthy	

Unhealthy	Healthy	
Unhealthy	Medium Healthy	
Unhealthy	Unhealthy	

As appeared in Table 12, assessing the anticipated qualities and watched estimations of toxin considerations together, the last alert hues are named.

3.5.6.3 Performance Metrics:

The following well-known methods, such as Precision, Confusion Matrix (CM), False Negative Rate (FNR), Accuracy metrics, F1-Score, Specificity, False Positive Rate (FPR) and G-mean are used to estimate the alarm color matching performance of the novel prediction strategy.

Precision and Recall then showed as;

$$\text{Pr}ecision = \frac{TP}{(TP + FP)} \quad \textbf{21}$$

$$\text{Re}call = \frac{TP}{TP + FN} \quad \textbf{22}$$

$$Specificity = \frac{TN}{(FP + TN)} \quad \textbf{23}$$

$$FPR = \frac{FP}{FP + TN} \quad \textbf{24}$$

$$FNR = \frac{FN}{FN + TP} \quad \textbf{25}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \textbf{26}$$

$$Sensitivity = \frac{TP}{TP + FN} \quad \textbf{27}$$

$$G - mean = \sqrt{Sensitivity \times precision} \quad \textbf{28}$$

The precision and recall harmonic mean is said to be a combine measure between recall and precision, the balanced F-score or traditional F-measure:

$$F - Score = 2 \times \frac{\Pr ecision \times recall}{\Pr ecision + recall} \quad \textbf{29}$$

Where, True Negative is indicated by *TN*, True Positive is indicated by *TP*, False Negative is represented by *FN* and False Positive is denoted as *FP*.

3.5.6.4 Results Analysis:

Here the results of twofold air poison fixations and appraisal of air quality forecast models are presented. To assess that expectation execution from a few points of view, we utilized Precision, False Negative Rate, G-Mean, False Positive Rate, F1-Score, Recall, Accuracy measurements, and Specificity.

As far as to mistake criteria, the proposed technique gives an improvement, Exactness, Confusion Matrix (CM), F1-Score, and Recall esteems are acquired by the outcomes from the course of action and which is set up by utilizing limit esteems. The Table 10 gives the edge esteems, and Table 11 gives the choice qualities.

Next, Table 12 speaks to the presentation measurement estimations of the Precision, False Negative Rate, G-Mean, False Positive Rate, Recall, F1-Score, explicitness, and Accuracy measurements dependent on the current techniques HMM-SVM, ANN, and KNN.

Table 12 Evaluation of Performance Metrics

Methods	Precision	Recall	Specificity	FPR	FNR	F1-measure	G-mean	Accuracy
Proposed	0.94	0.95	0.92	0.57	0.15	0.941	0.943	0.978
HMM-SVM	0.92	0.9	0.9	0.5	0.23	0.92	0.91	0.96
ANN	0.84	0.81	0.86	0.45	0.65	0.55	0.82	0.85
KNN	0.74	0.79	0.79	0.49	0.61	0.81	0.76	0.81

Table 13 Pollution Level Ranking of Cities

CITIES	POLLUTION LEVEL RANKING
1	9
2	3
3	10
4	2
5	1
6	4
7	8

8	7
9	5
10	6

Table 13 gives the contamination level positioning of urban areas. Here, the positioning of the urban areas assessed by figuring the contamination rates between the poisons' ozone and nitrogen dioxide. As indicated by the air quality list, the wellbeing concern level determined, and the caution has anticipated the shading. The Smart City positioned by utilizing this alert. Because of the undesirable Smart City are positioned.

Give us a chance to think about ten urban areas, among that we discover the risk level of every city. As indicated by the risky degree of the city, the most reduced level is said to be rank 1. Further levels are referenced as rank 2, position 3…, and so forth. The graphical speaking to the positioning of urban areas has appeared in Table 13.

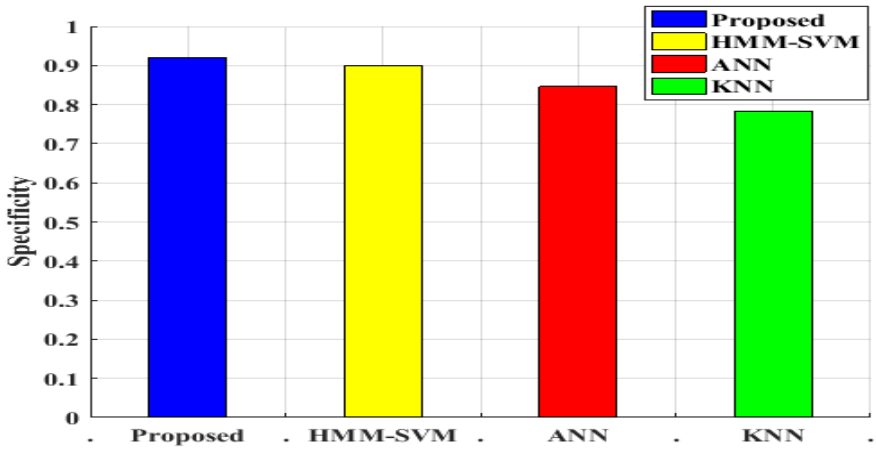


Figure 3-11 Performance of Specificity

The accompanying figures are the presentation examination of the extraordinary grids.

The presentation examination of graphical portrayal has appeared in Figure 3.11. It spoke to the proposed technique has great particularity execution contrasted with other existing strategies HMM-SVM, ANN, KNN.

Figure 3.12 and Figure 3.13 gives a graphical portrayal of the review and accuracy execution. It has high qualities instead of the HMM-SVM, ANN and KNN

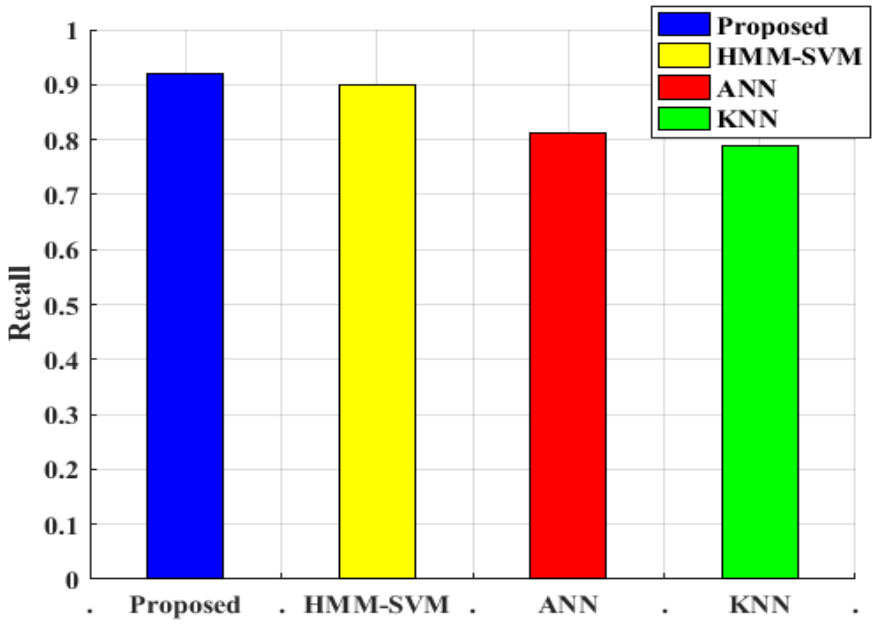


Figure 3-12 Recall Performance

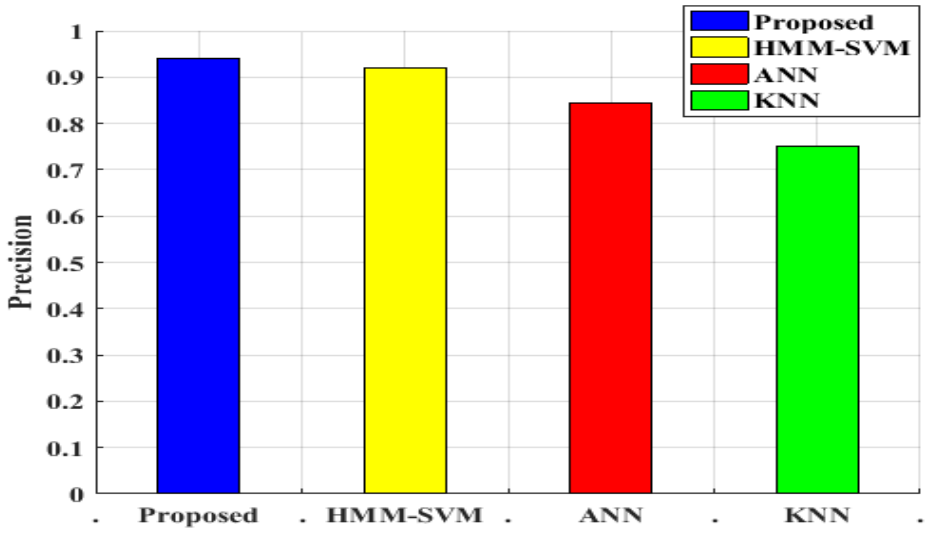


Figure 3-13 Precision Performance

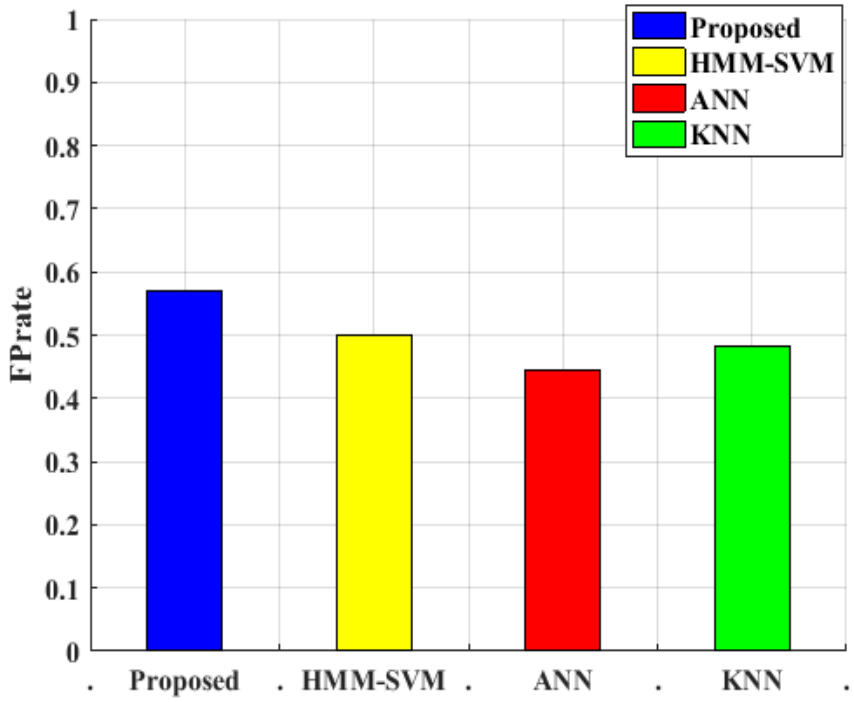


Figure 3-14 FPR Performance

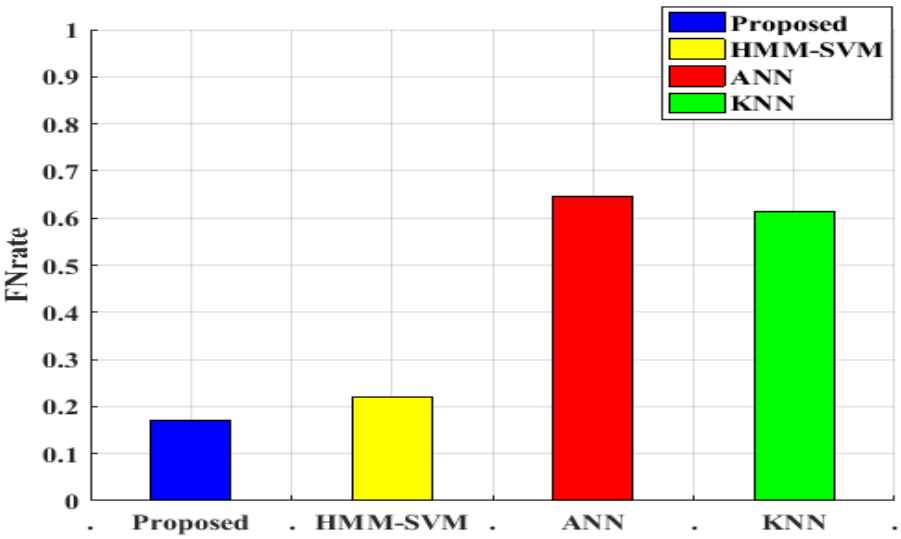


Figure 3-15 FNR Performance

Figure 3.14 and Figure 3.15 speaks to false positive and negative rate execution. When contrasted with the current methodologies, for example, HMM-SVM, ANN, and KNN, our proposed technique has a decent exhibition rate.

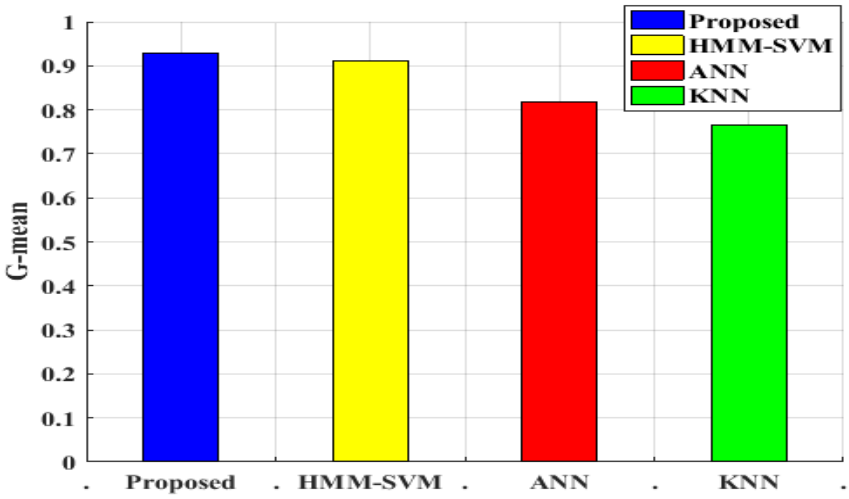


Figure 3-16 G-mean Performance

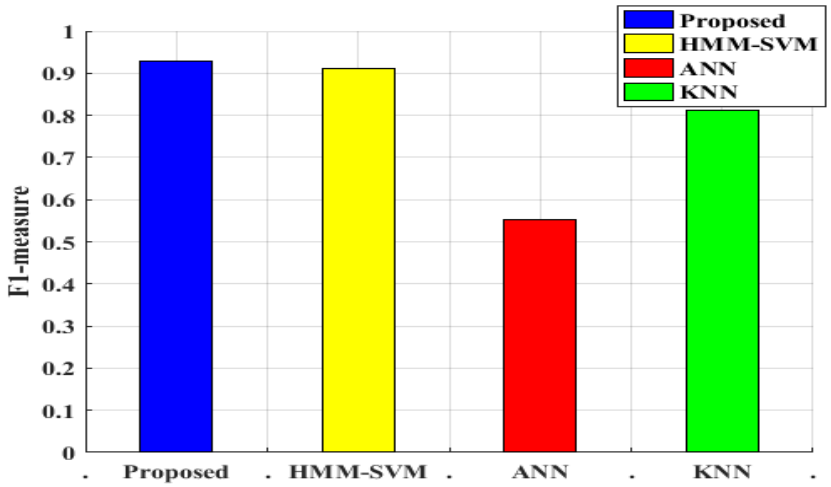


Figure 3-17 F1-measure Performance

Figure 3.16 and 3.17 shows the graphical representation of the performance of G-mean and F1-measure. By seeing this representation, our proposed method has a high-performance value.

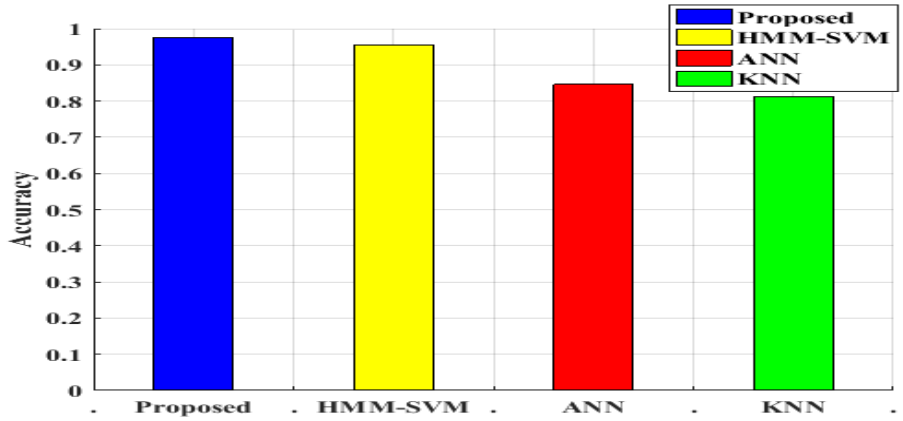


Figure 3-18 Performance of Accuracy

Figure 3.18 speaks to the presentation exactness graphical portrayal. The exactness of the proposed strategy is high contrasted with existing strategies, for example,

HMM-SVM, ANN, and KNN.

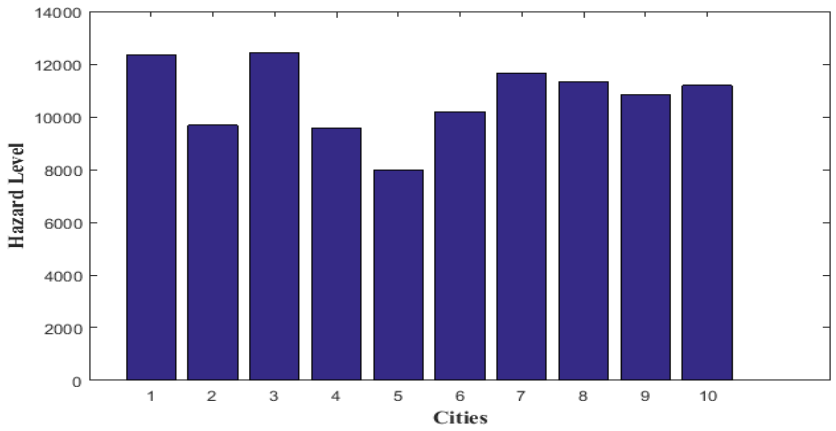


Figure 3-19 Ranking Level of Different Cities

Figure 3.19 gives a graphical portrayal of various urban areas. In which, the city having the most reduced perilous level set apart as rank 1. City 5 having rank one position that is risky level is low.

3.6 Summary

This chapter has given the detail of the proposed model for Indian smart city with data set according to the eighty indicators of eight dimensions. Each dimension has explained with its indicators of use. The general recommendation system model has given in detail with results. The adaptive whale optimization algorithm has given air quality hazard monitoring to save the environment.

The next chapter is about the ranking algorithm to rank the Indian smart city based on the proposed Indian smart city model data set given in this chapter. The ranking of cities and dimension rank of the individual city is explaining with the help of graphs.

Chapter Four: Ranking of Smart Cities

This chapter defines the "Indian Smart City Ranking Model," using the DBA and TDBA. The chapter is explained the "distance-based algorithm (DBA)" and Taxicab distance-based approach (TDBA) to rank the smart city using the indicators and dimensions of Indian Smart city Proposed Model. The simulation results explained with the result table and charts.

4.1 Introduction

The ranking is a relative comparison view, which gives knowledge of the standard relative position of growth. The smart city ranking provides a relative view of growth position among the cities, the race of improvement measured by ranking. The rank of each dimension of a city gives a clear direction for improvement. The algorithm DBA and TDBA defined for computing the optimal ranking solution for the smart city.

4.2 Indian Smart City Ranking Using DBA

4.2.1 Introduction

The Distance-based algorithm used for finding the optimal ranking solution for the ranking problem. In this, we are using all indicators value as an input to an algorithm to process and find the optimal ranking solution.

4.2.2 Distance-Based Algorithm (DBA)

The DBA [160]working methodology aims to specify the optimal objective point defines through the optimal model, i.e., Cities and the finest value in every criterion of the

ranking having in the computation activity. The all ranking criteria set of finest values in vector Cities, (c_1, c_2, …. c_n) space optimal points in n-dimensional.

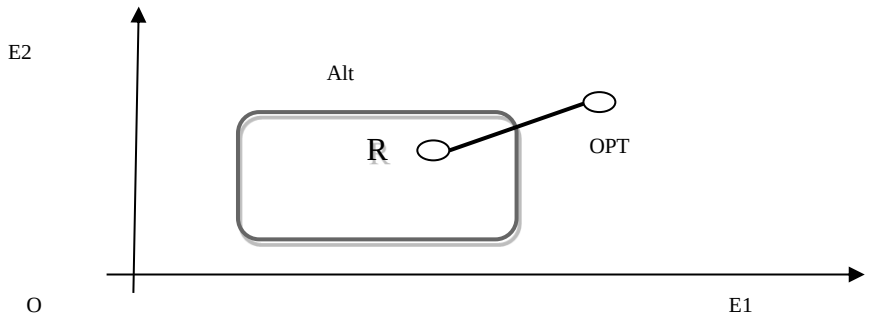


Figure 4-1 DBA Distance Approach

In this strategy for research, the ideal/best worth processed from all criteria taken the best estimation of the criteria. Thus, OPT implies different Cities having the best incentive for each foundation estimated. The probability of the presence of any elective

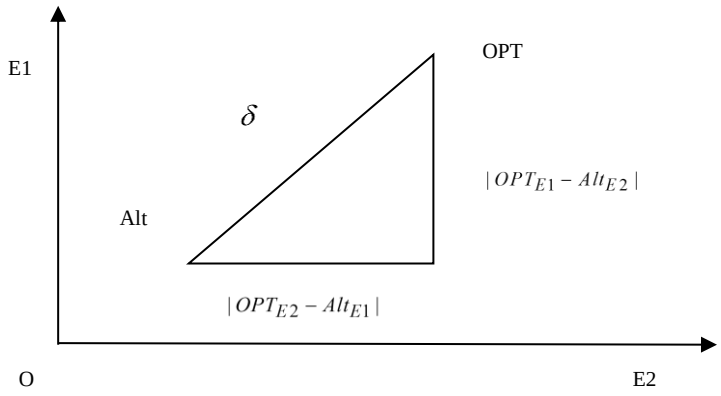


Figure 4-2 DBA Distance of Real Vector

Cities is less than outcomes regarding numerous choices to recommend the urban areas. Or this situation, OPT works as a suggestion point to compare every choice.

The optimal position, i.e., OPT fitness value calculated by the quantitative comparison of alternate cities. Therefore, here, we need to find the solution closest to OPT to solve this decision problem. The objective function can be defined to find a solution as:

$$Min \ \ \delta\left\{Alt(y), OPT\right\} \qquad \textbf{1}$$

Subject to y ς Y

Here { $Alt(y)$ }, and δ demonstrate a Cities in the n-dimensional space, and the good ways from OPT, separately. The arrangement is reliant on the two targets, OPT. The arrangement portrayal in the 2D region given in Figure 4.1. Here R is a conceivable region, and OPT is the optimal point.

The DBA locates the point in the possible region 'R' closest to the OPT, and this shows in Figure 4.2. The figure shows that the axis E1 and E2 have a parallel line $(Alt\text{-}OPT)_{E1}$ and $(Alt-OPT)_{E2}$ correspondingly. So $(Alt-OPT)_{E1}=|OPT_{E1}-Alt_{E1}|$, and $(Alt-OPT)_{E2}=|OPT_{E2}-Alt_{E2}|$. In 2-dimensional space, δ is given by

The DBA finds the point in the conceivable region 'R' nearest to the OPT, and this shows in Figure 4.2. The figure demonstrates that the axis E1 and E2 have a parallel line $(Alt\text{-}OPT)_{E1}$ and $(Alt-OPT)_{E2}$ likewise. Along these lines, $(Alt-OPT)_{E1}=|OPT_{E1}-Alt_{E1}|$, and $(Alt-OPT)_{E2}=|OPT_{E2}-Alt_{E2}|$ In 2-dimensional space, δ is given by

$$\delta=\left[(OPT_{E1}-Alt_{E1})^2+(OPT_{E2}-Alt_{E2})^2\right]^{1/2} \qquad \textbf{2}$$

In exact terms, the "distance δ" entrapped as

$$\delta = \left[\sum (OPT_{xy} - Alt_{xy})^2 \right]^{1/2} \quad \textbf{3}$$

Here x = 1, 2... n = alternate Cities and y=1, 2... m = ranking criteria.

To apply DBA needs to replicate 'n' Cities and 'm' positioning criteria related to every City, for example $Alt_1(c_{11}, c_{12}, ..., c_{1m})$ $Alt_2(c_{21}, c_{22}, ..., c_{2m})$, $Alt_n(c_{n1}, c_{n2}, ..., c_{nm})$ and the Cities $(c_{b1}, c_{b2}, ..., c_{bm})$ where c_{bm} = the ranking criteria best worth 'm.'

Along these lines, the total arrangement of cities can be appeared by the estimations of the ranking criteria in the accompanying matrix:

$$[c] = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1m} \\ c_{21} & c_{22} & \cdots & c_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nm} \\ c_{b1} & c_{b2} & \cdots & c_{bm} \end{bmatrix} \quad \textbf{4}$$

Presently expelled the estimation units' impact by grid utilizing the equation:

$$Z_{xy} = \frac{c_{xy} - \overline{c_y}}{S_y} \quad \textbf{5}$$

Here, $\bar{c}_y = \frac{1}{n} \sum_{x=1}^{n} c_{xy}$, and **6**

$$S_y = \left[\frac{1}{n} \sum_{x=1}^{n} (c_{xy} - \bar{c}_y)^2 \right]^{1/2} \quad \textbf{7}$$

Here x = 1, 2, ... , n, and y = 1, 2… , m.

For each ranking criteria of all alternative Cities calculate $\bar{c}_y$, and s_y provides the average of value and the standard deviation.

$$[z_s] = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1m} \\ Z_{21} & Z_{22} & \cdots & Z_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ Z_{n1} & Z_{n2} & \cdots & Z_{nm} \\ Z_{OPT1} & Z_{OPT2} & \cdots & Z_{OPTm} \end{bmatrix} \quad \textbf{8}$$

Here $Z_{11} = \frac{c_{11} - \bar{c}_1}{S_1}$, $Z_{12} = \frac{c_{12} - \bar{c}_2}{S_2}$, $Z_{1m} = \frac{c_{1m} - \bar{c}_m}{S_m}$.

After this, calculate the distance matrix [zd]by subtracting the value of the optimal set by a subsequent value in the alternatives set. This matrix gives the cities differences to the OPT**.**

$$[z_d] = \begin{bmatrix} Z_{\mathrm{OPT}1} - Z_{11} & Z_{OPT2} - Z_{12} & \cdots & Z_{OPTm} - Z_{1m} \\ Z_{\mathrm{OPT}1} - Z_{21} & Z_{OPT2} - Z_{22} & \cdots & Z_{OPTm} - Z_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ Z_{\mathrm{OPT}1} - Z_{n1} & Z_{OPT2} - Z_{n2} & \cdots & Z_{OPTm} - Z_{nm} \end{bmatrix} \quad \textbf{9}$$

Presently, at last, the distance, CD determined in one another Cities to the OPT utilizing:

$$CD_{OPT-Alt} = \left[\sum_{y=1}^{m} (Z_{OPTy} - Z_{xy})^2 \right]^{1/2} \quad \textbf{10}$$

This Euclidean distance is mathematical expiration to calculate many distances on each rank criteria for which cities are evaluated and finally ranked. The flow of the algorithm shows in Figure 4.3.

4.2.3 Evaluating "Indian Smart cities" Ranking

Information on data set portrayed in Table-6,7,8,9 has gathered from a report distributed by govt. of India. Measure 8 criteria are "Demographics Profile," "Economic Profile," "Infrastructure Profile," e-Governance and Computerization, "Finance," "Environmental," "Progress Track," "Security," and 80 Indicators/sub-criteria. All Indicators/sub-criteria have equivalent wattage 59 number of criteria's have high qualities as ideal/best criteria, and 21 number of criteria's having low an incentive as ideal/best. There are 20 Smart City in table-3 that is to positioned dependent on these 80 criteria. Ligated qualities have been changed over to numerical qualities by doling out 0 and 1 for criteria E-Governance and Computerization. In Table 3, the cities spoke to C1 to C20, and all criteria are symbolled as an AS A1-A13, B AS B1-B21, C AS C1-C15, D AS D1-D12, E AS E1-E11, F AS F1-F2, G AS G1-G4 AND H AS H1-H2. All the emblematic portrayal clarified in Table-3,4,5. The info grid has 20 lines spoken to competitor cities for ranking, and 80 sections of the framework spoke to the criteria for the count of the rank of these 20 cities.

Case -1: When a lower value of a criteria shows the best match with the real data, then adjusted criteria value = maximum criteria value- criteria value in the data set so value $(c = c_{max} - c_i)$.

Case-2: When a higher value of a criteria shows the best match with real data, then adjusted criteria value= criteria value – minimum criteria value in data set so value $(c = c_i - c_{min})$.

The condition 5 and 9 are utilized to ascertain the institutionalized framework and separation lattice generally, and the condition 10 gives the CD "Euclidean composite separation" between each elective city to the ideal position, OPT. The CD and positioning of "Indian Smart City" in light of all impacting criteria appeared in Table 14. The DBA discover positioning dependent on the CD, which gauging every one of the 80 affecting criteria of every one of the other cities. The rank one assigned to the smallest CD in respect and by the next smallest given position two and proceeds.

The outcome demonstrates that the Chennai has accomplished position number one after investigation 80 criteria and next succession is Pune, Jaipur, Ahmadabad, Surat, Vishakhapatnam, Indore, Coimbatore, Bhopal, Davanagere, Delhi, Udaipur, Ludhiana, Bhubaneswar, Jabalpur, Kochi, Guwahati, Belgaum, Solapur, Kakinada respectively.

The individual positioning of all criteria appears in Table no. 15. This outcome knows the development of positioning in each component of the city. The individual position thinks about shows which city needs more development in which measurements. City C1 (Kakinada) having rank 5 in b (“Economic Profile”) measurement and 20 in measurement e(“Finance”), so this demonstrates the city has financial profile is dynamic, yet fund the executives is expected to improve. The city C9 (Kochi) got rank 1 in measurement f (“Environmental”) and 20 in a(“Demographics Profile”), so it needs to focus on measurement “Demographics Profile,” etc. The accompanying diagrams demonstrate the positioning situation of each component of each city. Figure 4.4 to 4.25 gives a graphical view of ranking.

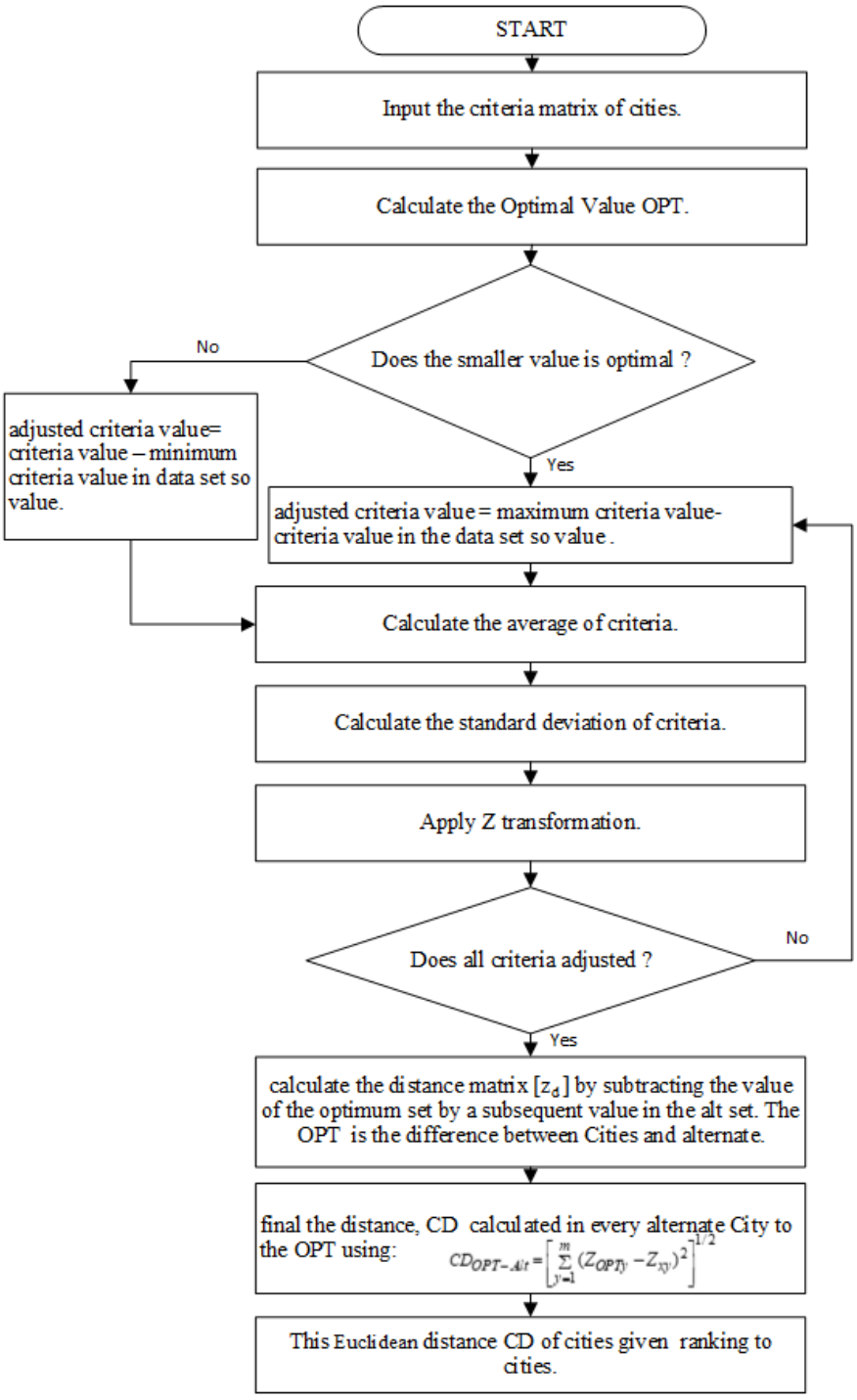


Figure 4-3 Flow Chart of Smart City Ranking Using DBA

4.2.4 Result and Charts

Table 14 Indian Smart City Ranking Result Using DBA

S. No	City	Sum	CD	Rank
1	Kakinada	595.6275	24.4055	20
2	Vishakhapatnam	441.6107	21.0145	6
3	Guwahati	530.7805	23.0387	17
4	Delhi	471.2837	21.7091	11
5	Ahmadabad	404.0672	20.1014	4
6	Surat	437.8815	20.9256	5
7	Belgaum	544.017	23.3242	18
8	Davanagere	464.8986	21.5615	10
9	Kochi	528.3442	22.9857	16
10	Bhopal	461.8994	21.4918	9
11	Indore	450.1009	21.2156	7
12	Jabalpur	516.5996	22.7288	15
13	Solapur	550.2184	23.4567	19
14	Pune	358.2065	18.9263	2
15	Bhubaneswar	514.9869	22.6933	14
16	Ludhiana	493.736	22.2202	13
17	Jaipur	394.2057	19.8546	3
18	Udaipur	489.0682	22.1149	12
19	Chennai	343.7742	18.5411	1
20	Coimbatore	458.2866	21.4076	8

Table 15 Indian Smart City Each Dimension Ranking Result Using DBA

	A	B	C	D	E	F	G	H	OVERALL
C1	19	5	11	19	20	16	17	12	20
C2	10	3	8	1	14	10	4	5	6
C3	9	7	12	20	12	7	8	15	17
C4	12	18	16	2	1	14	18	19	11
C5	2	12	17	3	3	8	5	9	4
C6	3	20	15	4	4	11	2	1	5
C7	17	4	6	18	16	17	19	8	18
C8	15	13	5	10	9	18	12	2	10
C9	20	2	20	11	8	1	16	13	16
C10	8	14	7	12	13	13	10	11	9
C11	5	19	10	5	6	4	7	18	7
C12	7	17	19	6	18	5	11	20	15
C13	18	11	14	17	19	19	14	6	19
C14	4	10	3	13	2	15	1	7	2
C15	6	15	18	14	17	12	13	16	14
C16	13	16	13	16	7	2	15	17	13
C17	11	1	1	7	10	3	9	14	3
C18	14	9	2	15	15	20	20	10	12
C19	1	8	4	8	5	9	3	3	1
C20	16	6	9	9	11	6	6	4	8

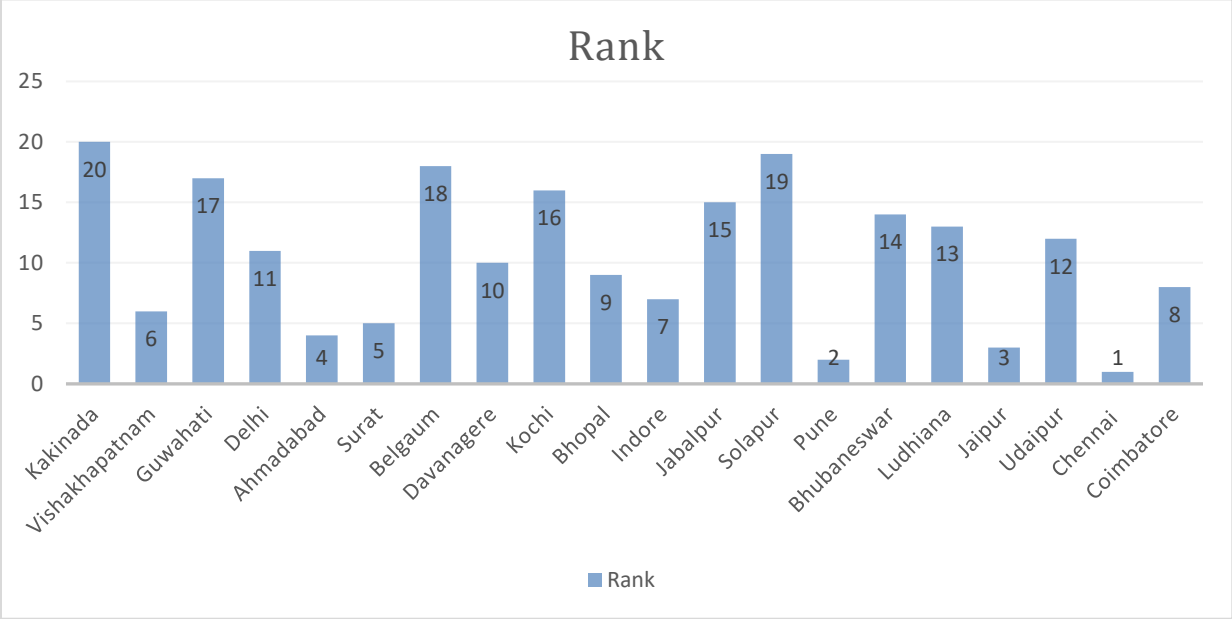


Figure 4-4 Indian Smart City Ranking Result Using DBA

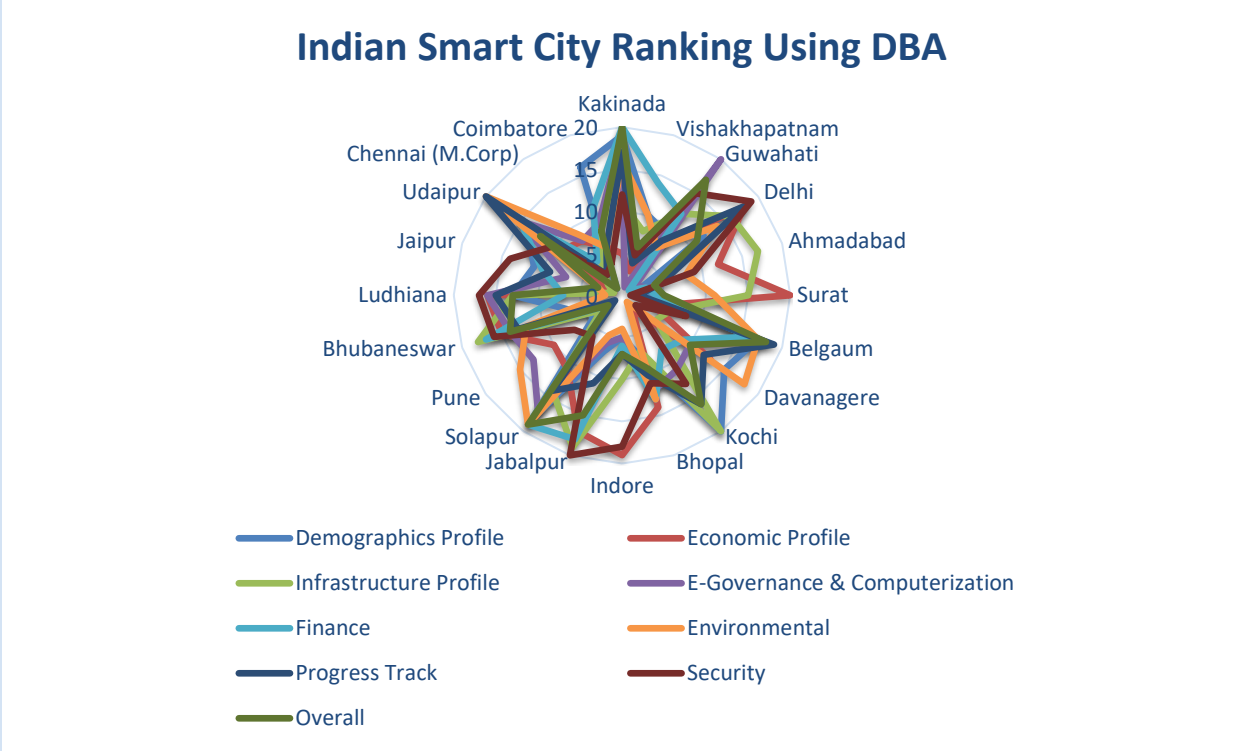


Figure 4-5 Indian Smart City Each Dimension Ranking Result Using DBA

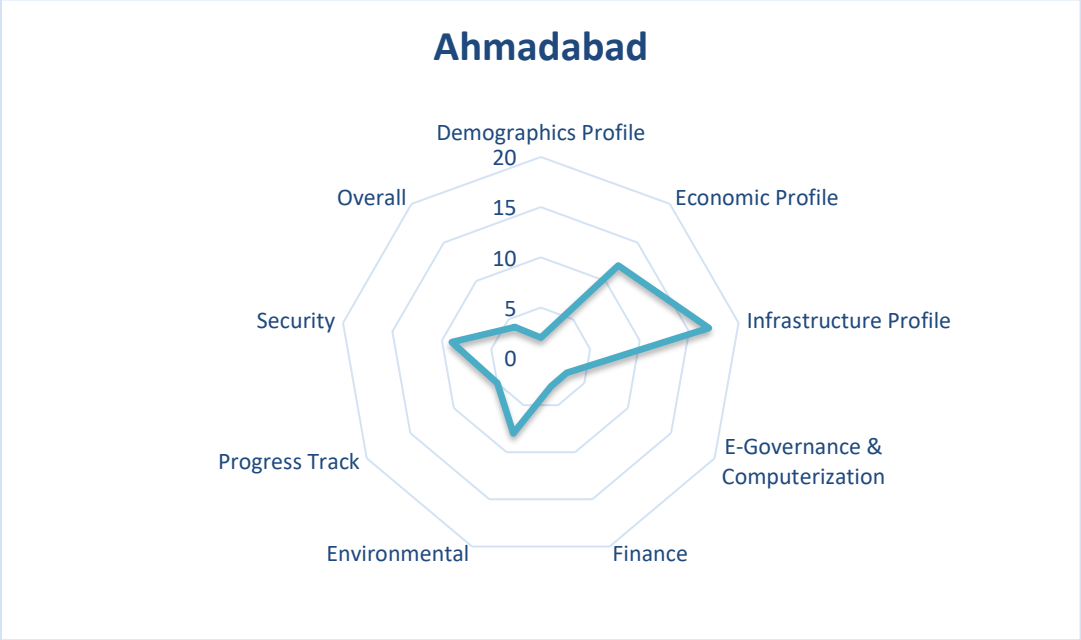


Figure 4-6 Ahmadabad Dimensions Ranking using DBA

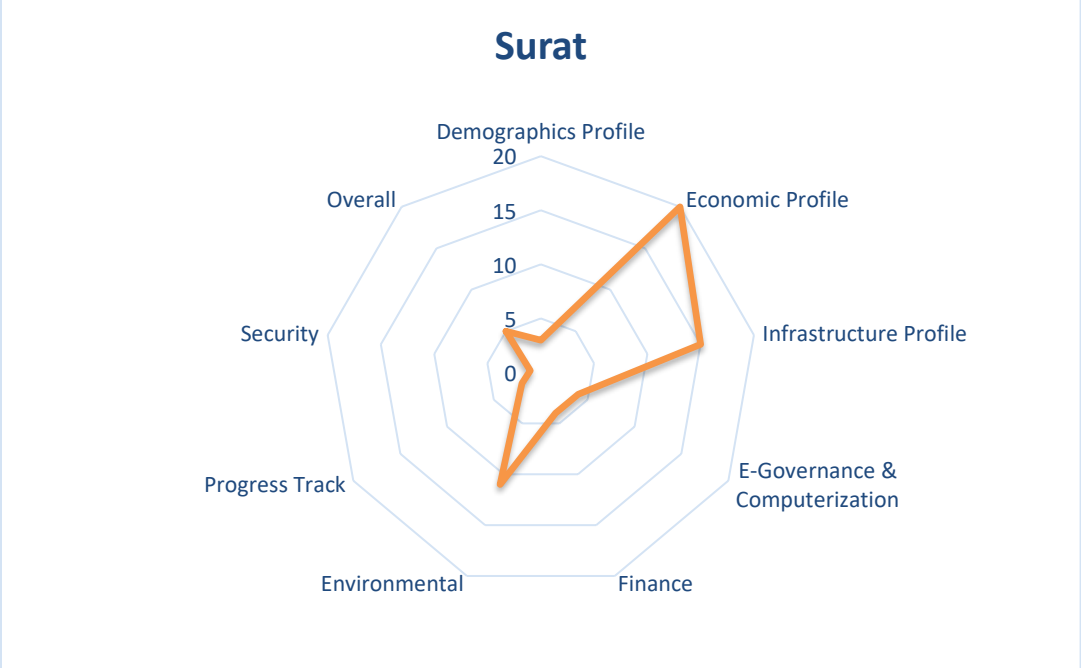


Figure 4-7 Surat Dimensions Ranking using DBA

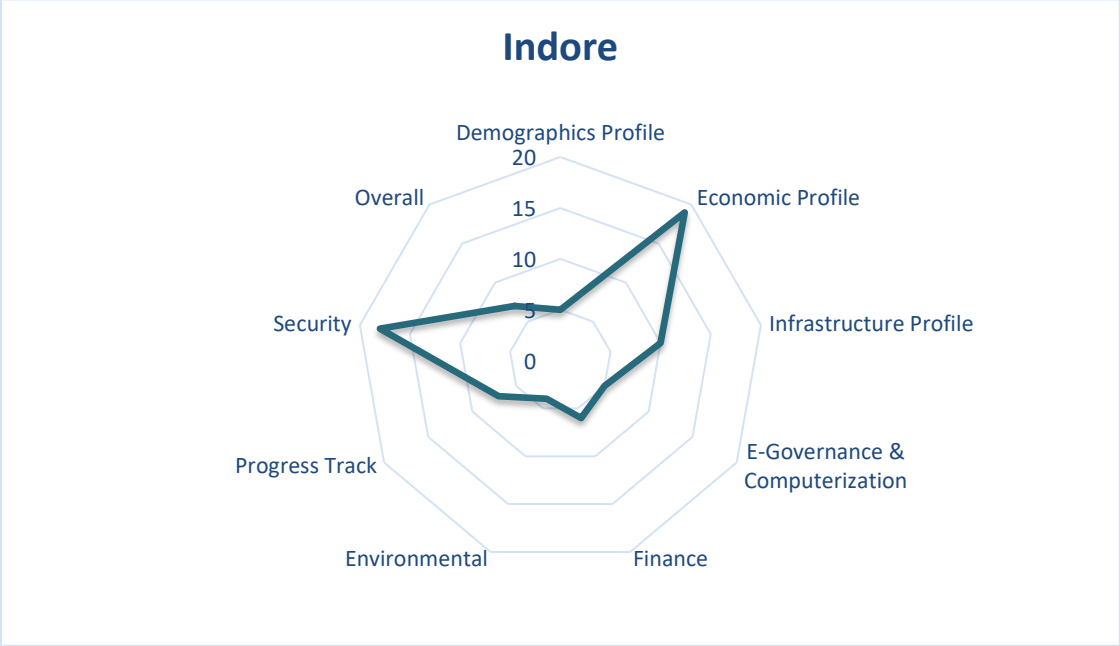


Figure 4-8 Indore Dimensions Ranking using DBA

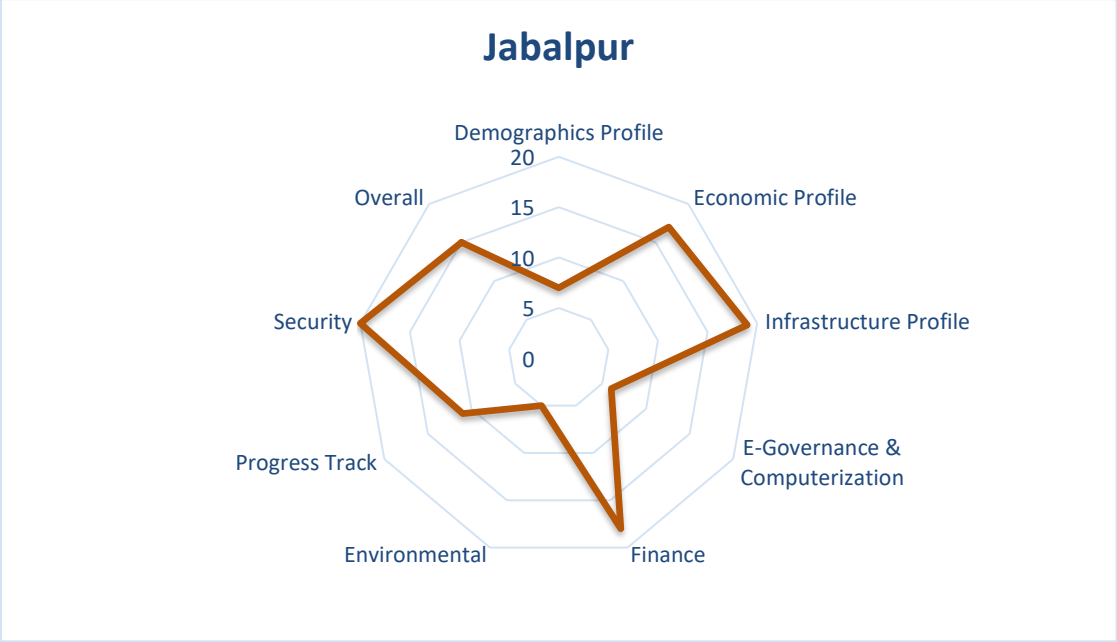


Figure 4-9 Jabalpur Dimensions Ranking using DBA

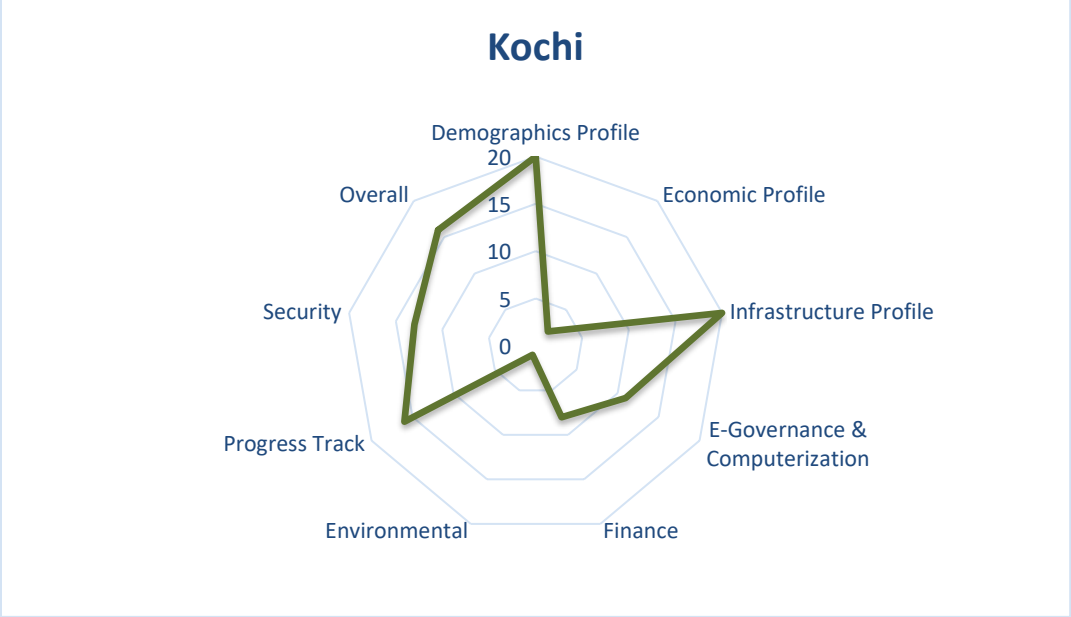


Figure 4-10 Kochi Dimensions Ranking using DBA

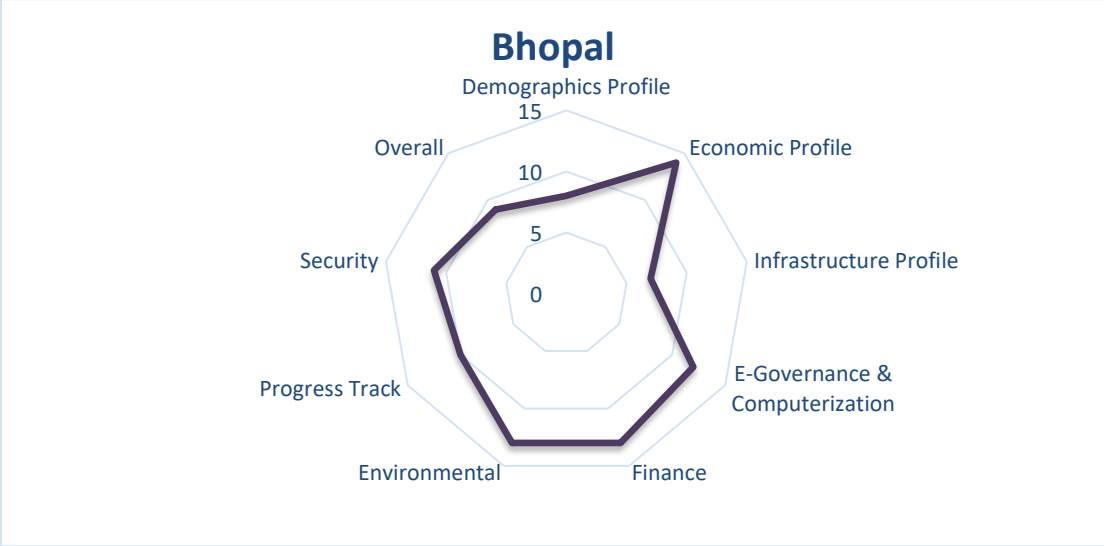


Figure 4-11 Bhopal Dimensions Ranking using DBA

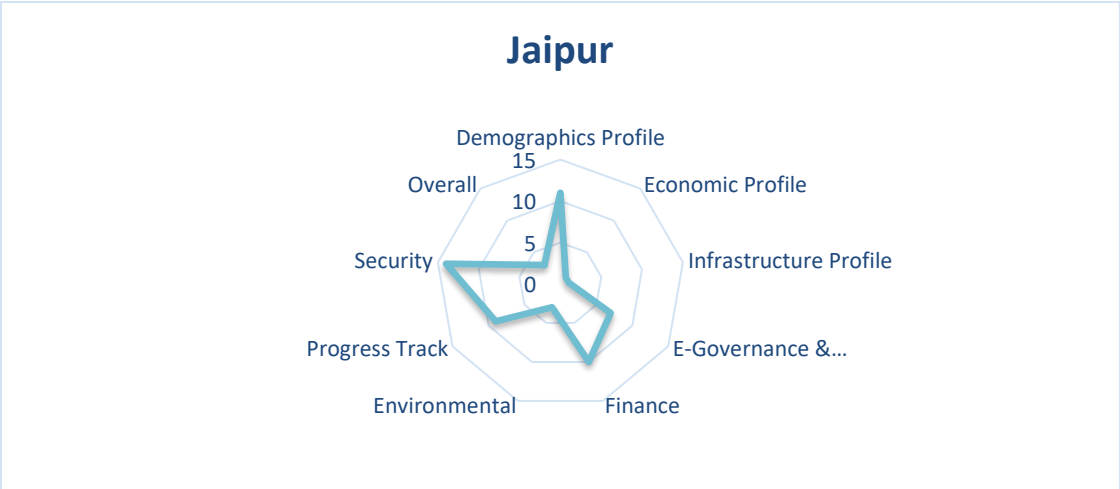


Figure 4-12 Jaipur Dimensions Ranking using DBA

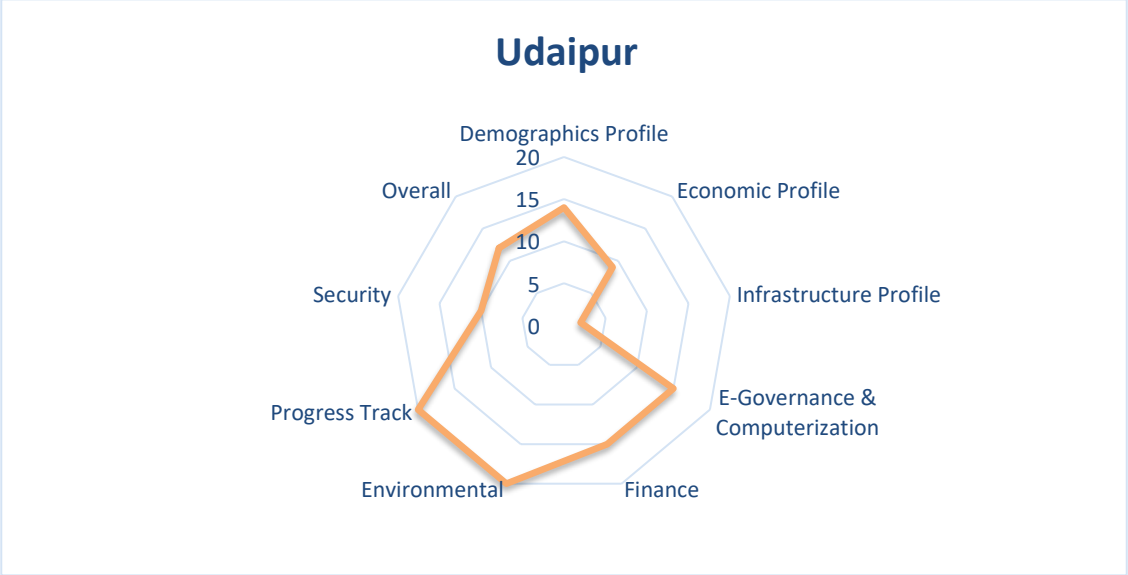


Figure 4-13 Udaipur Dimensions Ranking using DBA

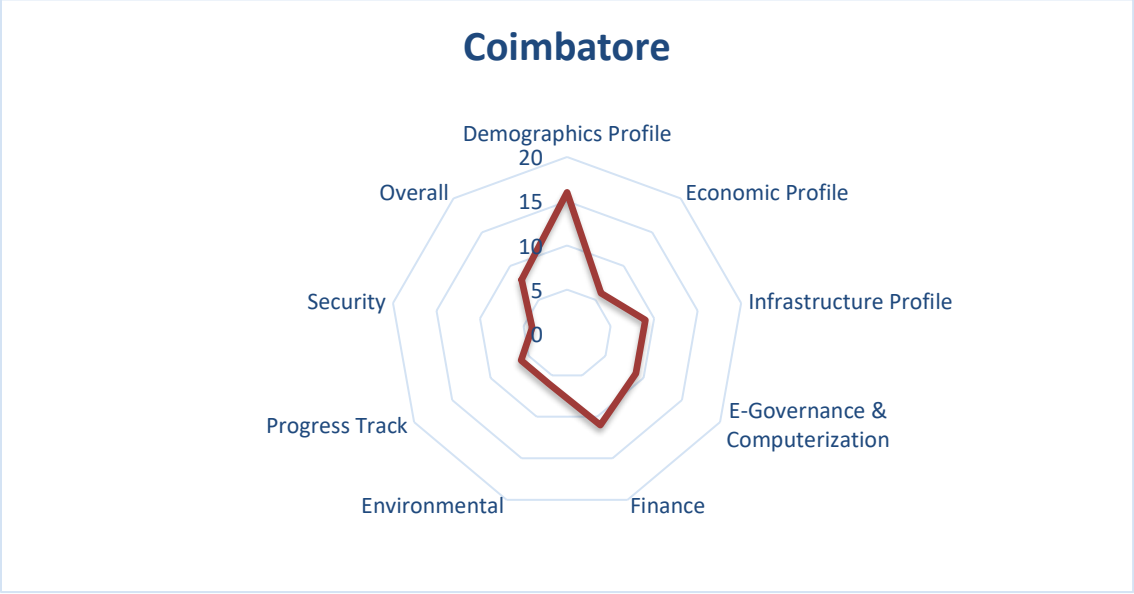


Figure 4-14 Coimbatore Dimensions Ranking using DBA

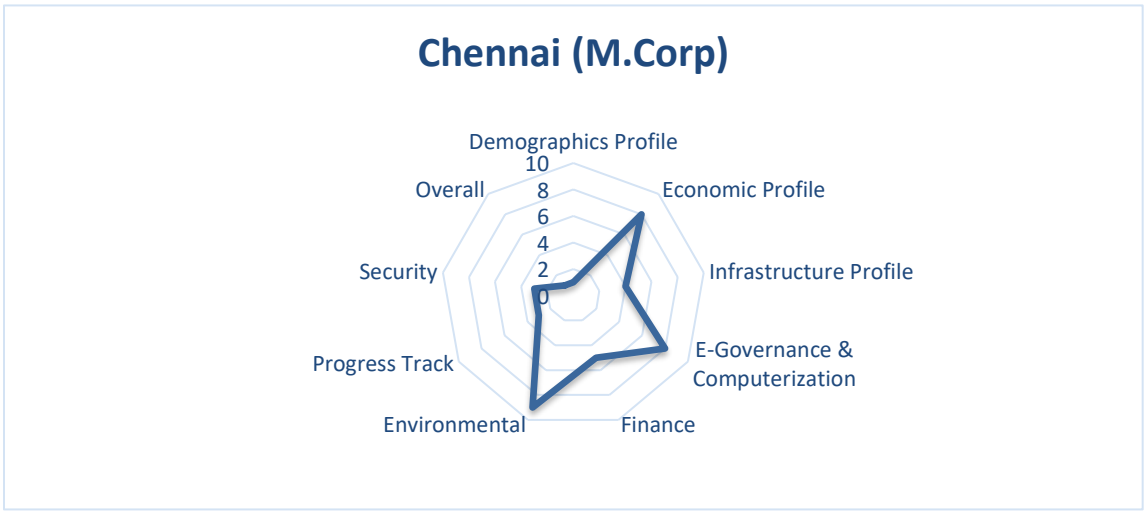


Figure 4-15 Chennai Dimensions Ranking using DBA

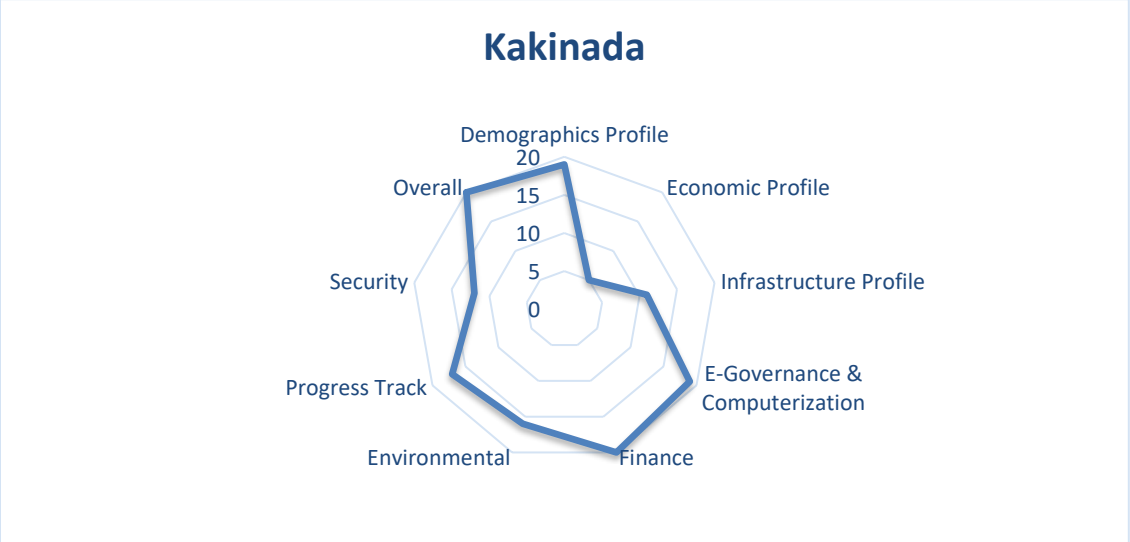


Figure 4-16 Kakinada Dimensions Ranking using DBA

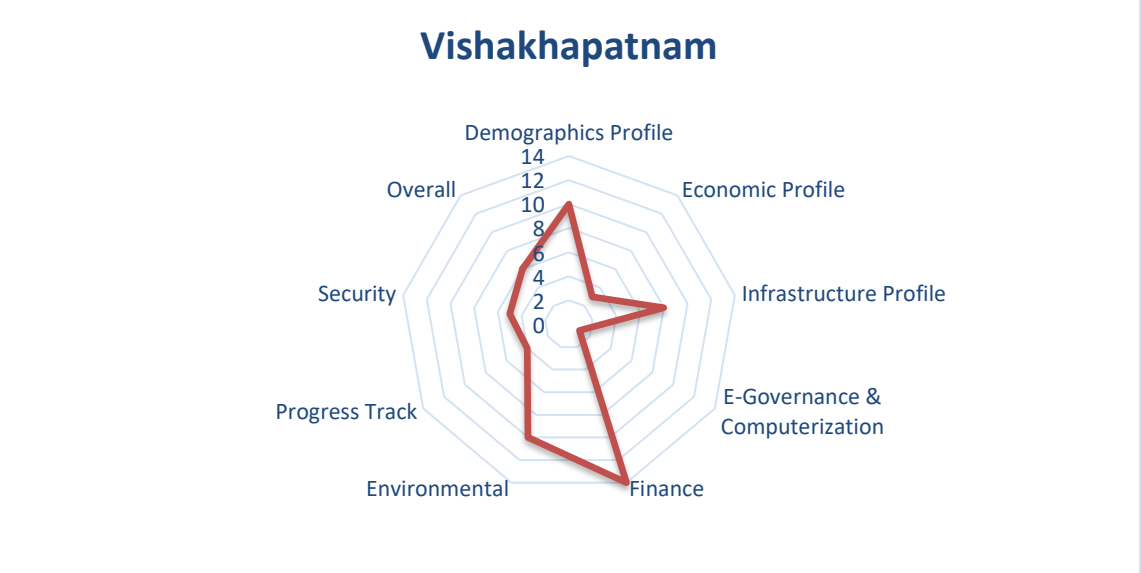


Figure 4-17 Vishakhapatnam Dimensions Ranking using DBA

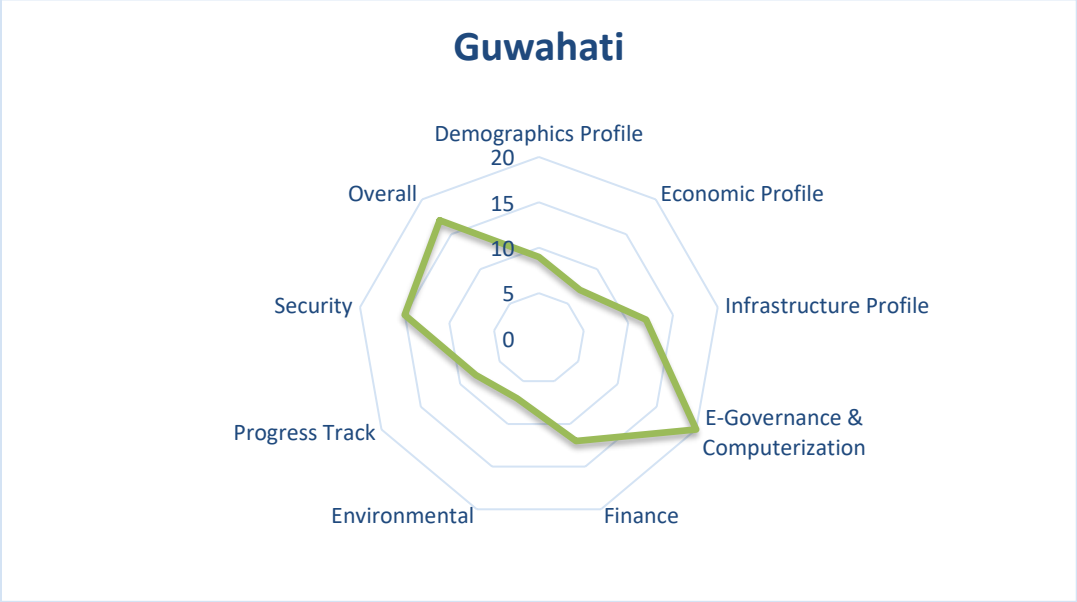


Figure 4-18 Guwahati Dimensions Ranking using DBA

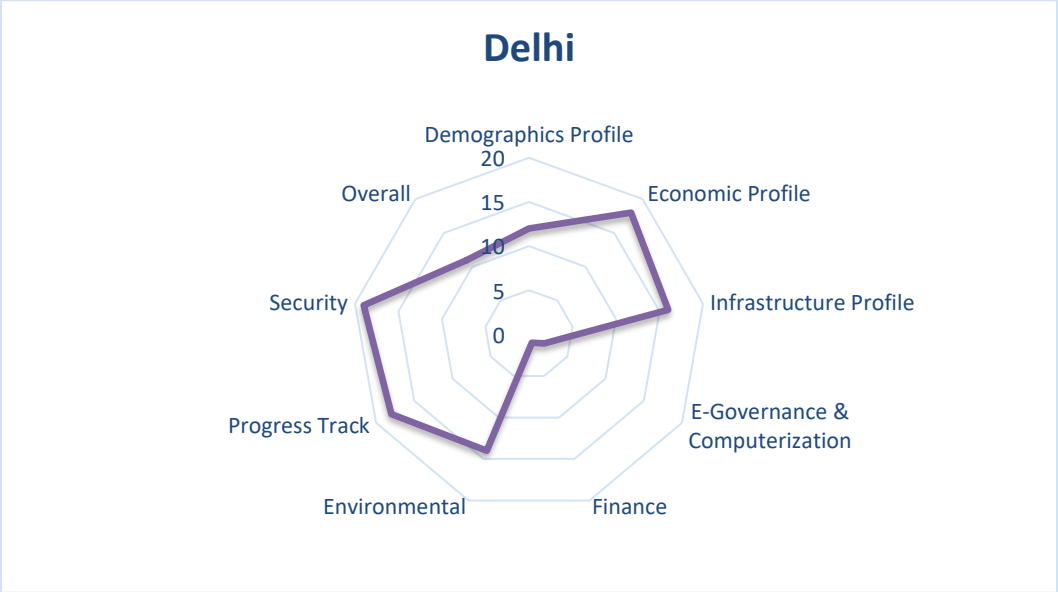


Figure 4-19 Delhi Dimensions Ranking using DBA

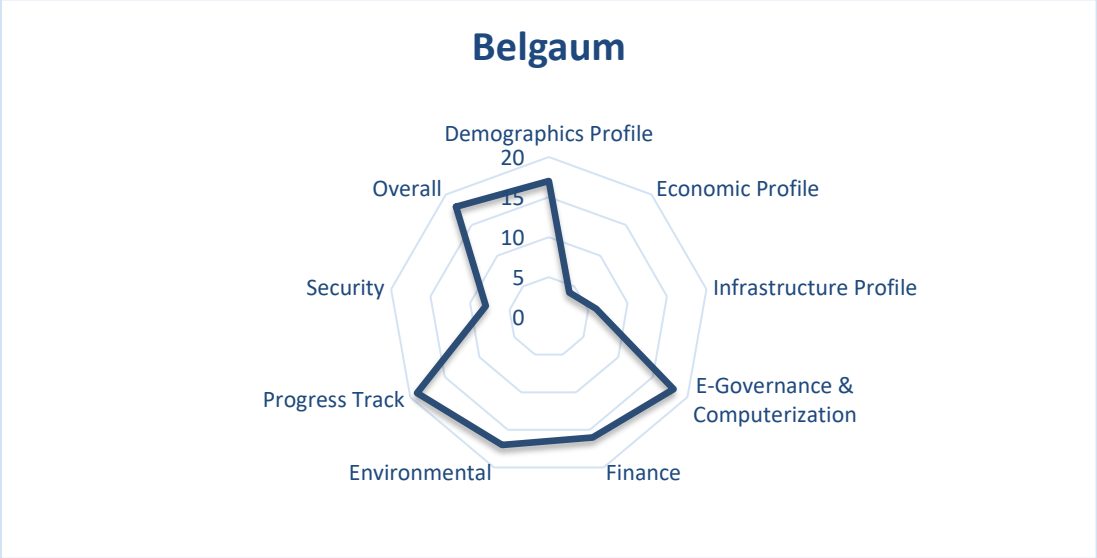


Figure 4-20 Belgaum Dimensions Ranking using DBA

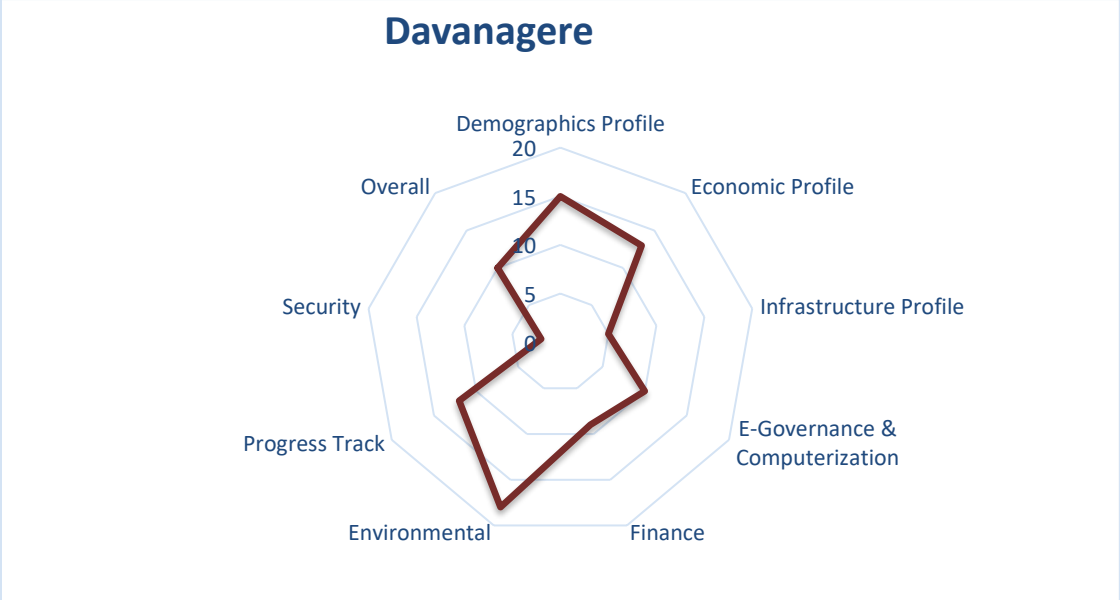


Figure 4-21 Davanagere Dimensions using DBA

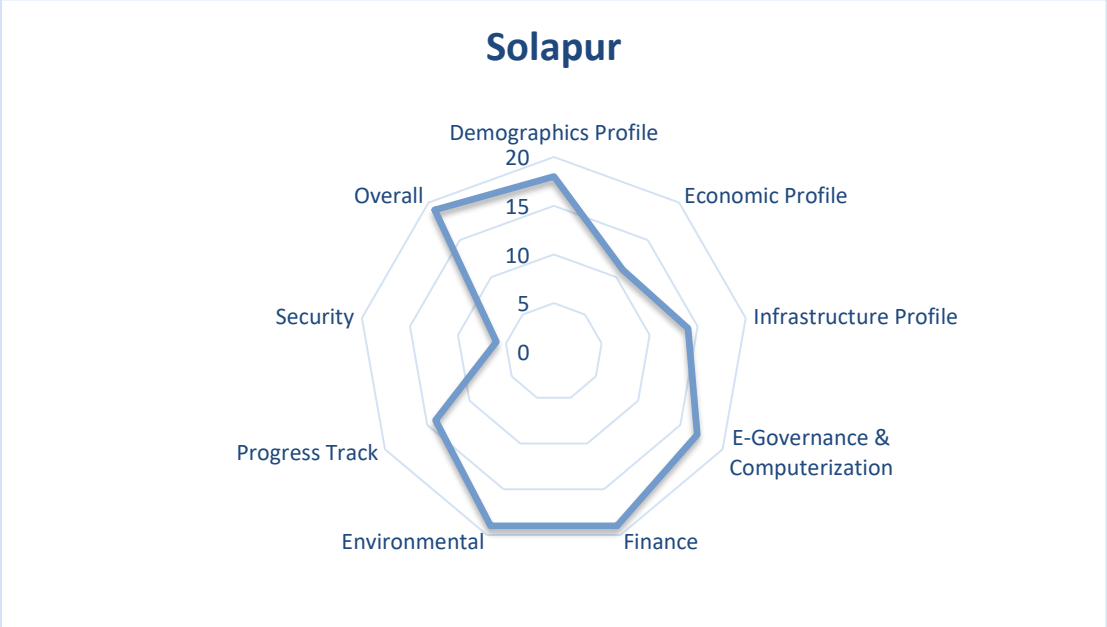


Figure 4-22 Solapur Dimensions Ranking using DBA

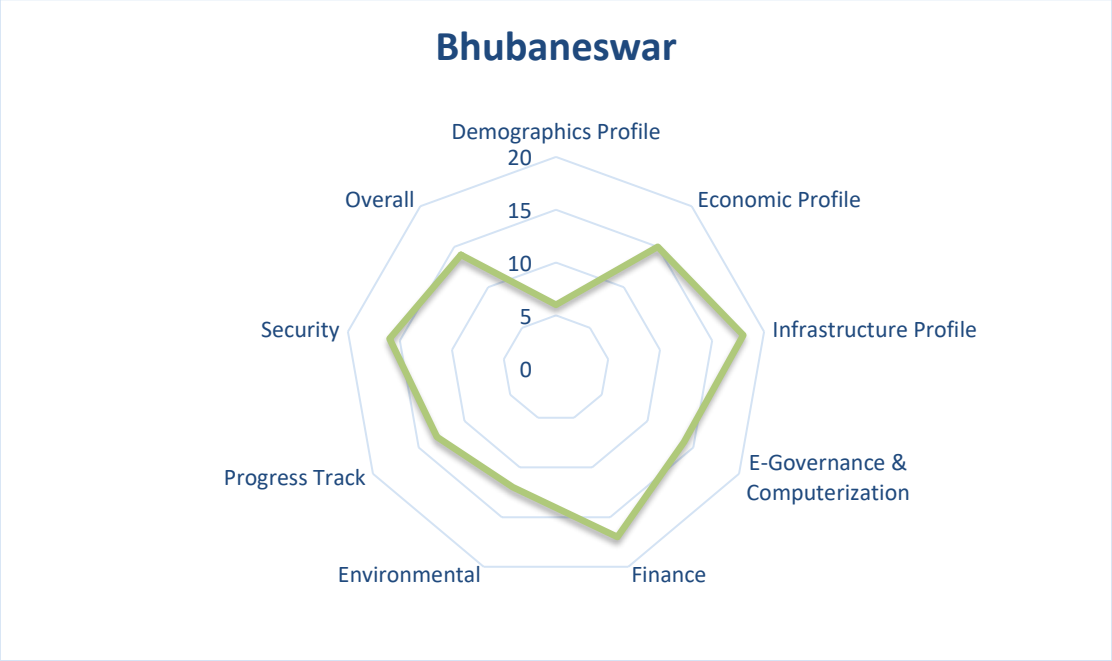


Figure 4-23 Bhubaneshwar Dimensions Ranking using DBA

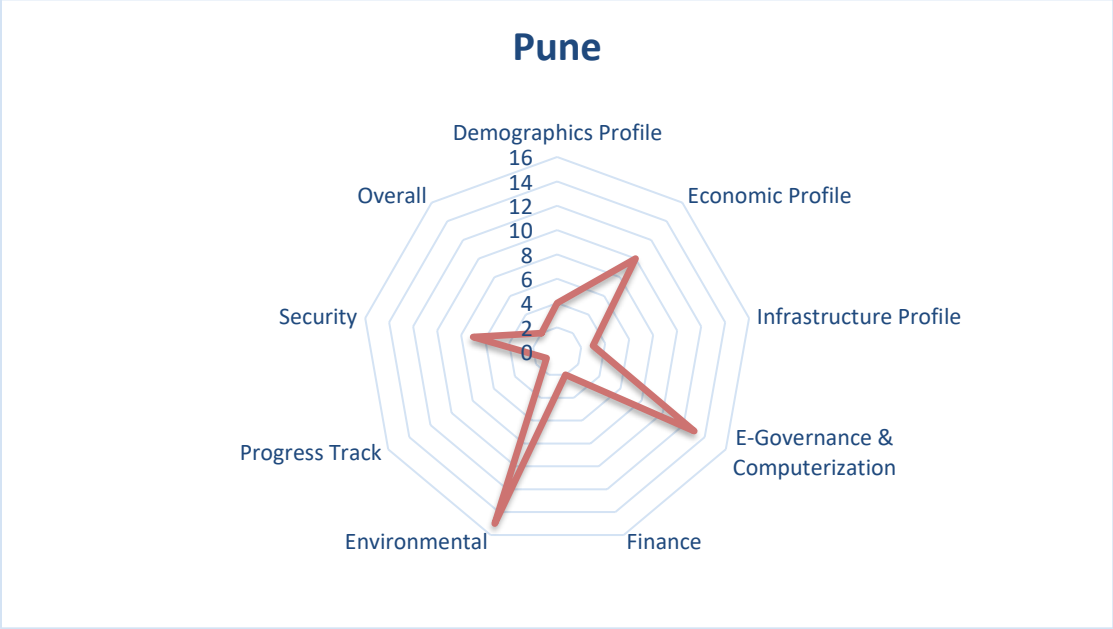


Figure 4-24 Pune Dimensions Ranking using DBA

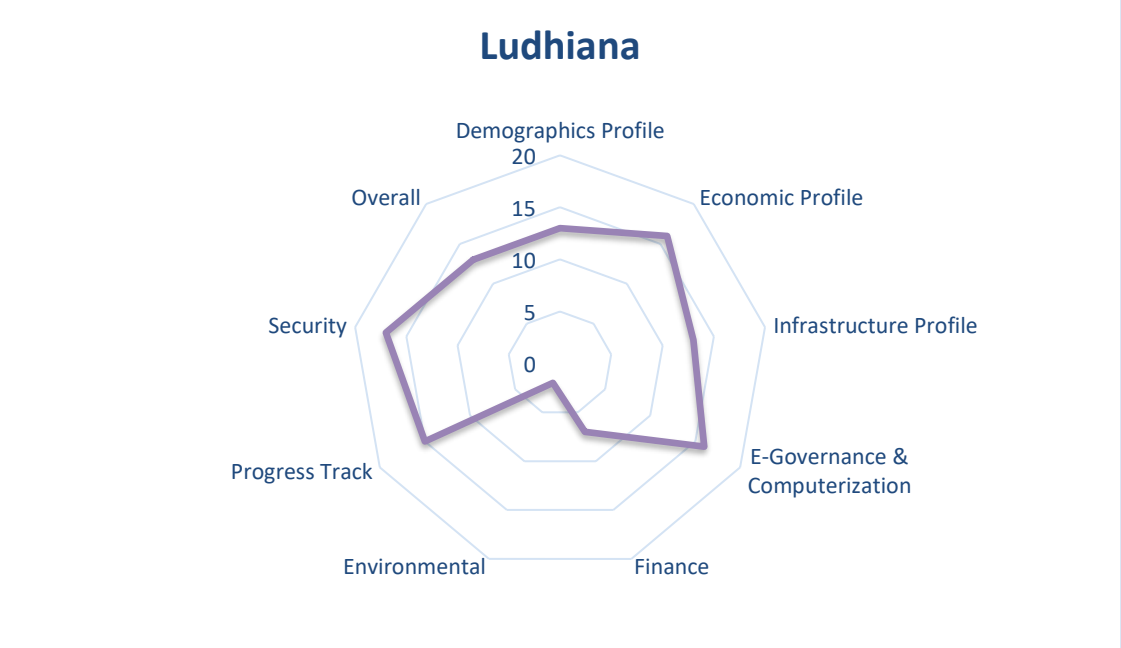


Figure 4-25 Ludhiana Dimensions Ranking using DBA

4.3 Indian Smart City Ranking Using TDBA

4.3.1 Introduction

The ranking is a multiple criterion selection Problem which is solved by this proposed taxicab distance algorithm. In the algorithm, the optimal selection of criteria based on the optimal distance solution from each alternative criterion. The ranking of smart cities is needed to compare multiple criteria of various dimensions, so this approach of the algorithm provided an optimal ranking solution.

4.3.2 Taxicab Distance Based Smart City Ranking (TDBA)

The TDBA operational standard is to specify the optimum objective spot identifies within the optimum model as criteria of cities and the noblest value in each criterion of the

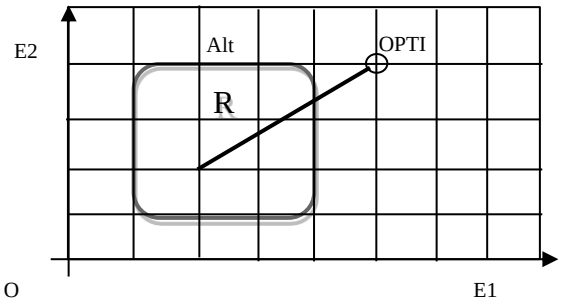


Figure 4-26 Taxicab Distance Approach

cities having in the computation action. The cities ranking criteria indicators are quality value set, which represented a vector (C_1, C_2, …. C_N) in n-dimensional area optimum spot. In the approach, the optimum quality value evaluated through all the criteria indicators noblest value. So, OPTI represents to the alternate Cities taking the noblest value for each criteria indicator computed. The chance of the presence of several alternate Cities is very fewer that results in respect of several alternates to advise the cities. In this case, OPTI

works as a suggestion spot to compare every alternate. The optimum spot, i.e., OPTI fitness value evaluated by the likely comparison of alternate's Cities. Therefore, here we need to find an answer nearby to OPTI to resolve this decision-based problem. The objective function formulated to obtain a solution as:

$$Min \ \ \delta\left\{Alt(y), OPTI\right\} \qquad \textbf{11}$$

Subject to $_{y}$ ς Y

Here, { $_{Alt(y)}$ }, and δ indicate cities in an n-dimensional area, and the distance from OPTI, correspondingly. The result is conditional on the two objectives, OPTI and δ. The result demonstration in the 2D area provided in Figure 4.26. Here possible region is R, and OPTI is the optimum spot.

The TDBA finds the spot in the possible region 'R' nearby to the OPTI, and this appears in Figure 4.27. In the figure, axis E1 and E2 have a parallel line $_{(Alt\text{-}OPTI)_{E1}}$ and $(Alt-OPTI)_{E2}$ respectively. So $(Alt-OPTI)_{E1}=|OPTI_{E1}-Alt_{E1}|$, and $(Alt-OPTI)_{E2}=|OPTI_{E2}-Alt_{E2}|$. In 2-D, δ is given by

$$\delta=|OPTI_{E1}-Alt_{E1}|+|OPTI_{E2}-Alt_{E2}| \qquad \textbf{12}$$

In specific terms, the "distance δ" taxicab distance use to frame as

$$\delta=\sum|OPTI_{xy}-Alt_{xy}| \qquad \textbf{13}$$

Here x = 1 to n = alternate Cities and y=1 to m ranking criteria.

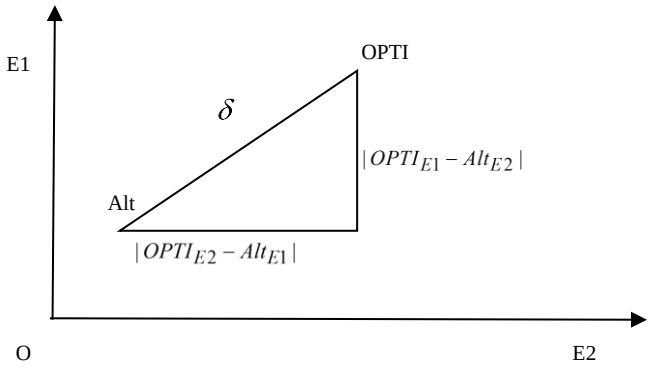


Figure 4-27 Taxicab Distance

To implement TDBA requirement to signify 'n' Cities and 'm' criteria for ranking associated with every City, e.g. $Alt_1(C_{11},C_{12},...,C_{1m})$, $Alt_2(C_{21},C_{22},...,C_{2m})$, $Alt_n(C_{n1},C_{n2},...,C_{nm})$, and the Cities $(C_{b1},C_{b2},...,C_{bm})$ where C_{bm} = the ranking criteria quality value 'm.'

So, the whole set of cities can be appeared by the values of every given ranking criteria indicators in this matrix:

$$[C]=\begin{bmatrix} C_{11} & C_{12} & \dots & C_{1m} \\ C_{21} & C_{22} & \dots & C_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nm} \\ C_{b1} & C_{b2} & \dots & C_{bm} \end{bmatrix}. \qquad \textbf{14}$$

Now eliminated the quantity units influence by standard matrix using the formula:

$$Z_{xy}=\frac{C_{xy}-\overline{C_y}}{S_y} \qquad \textbf{15}$$

$$\text{Here, } \overline{C}_y=\frac{1}{n}\sum_{x=1}^{n}C_{xy} \text{ , and} \qquad \textbf{16}$$

$$S_y=\left[\frac{1}{n}\sum_{x=1}^{n}(C_{xy}-\overline{C}_y)^2\right]^{1/2} \qquad \text{Here x = 1 to n, and y = 1 to m.} \qquad \textbf{17}$$

For every ranking criterion, indicators of all alternate Cities evaluate $\overline{C}_y$ donates the average value and s_y donates the standard deviation. The matrix for standardized framed as:

$$[z_s] = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1m} \\ Z_{21} & Z_{22} & \cdots & Z_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ Z_{n1} & Z_{n2} & \cdots & Z_{nm} \\ Z_{OPTI1} & Z_{OPTI2} & \cdots & Z_{OPTIm} \end{bmatrix} \quad \textbf{18}$$

Here $Z_{11} = \frac{C_{11} - \overline{C}_1}{S_1}$, $Z_{12} = \frac{C_{12} - \overline{C}_2}{S_2}$, $Z_{1m} = \frac{C_{1m} - \overline{C}_m}{S_m}$.

After this evaluation, the distance matrix [z_d] by deducting the value of the optimum set by the following value in the alt set. The OPTI is the difference between Cities and alternate.

$$[z_d] = \begin{bmatrix} Z_{\text{OPTI1}} - Z_{11} & Z_{OPTI2} - Z_{12} & \cdots & Z_{OPTIm} - Z_{1m} \\ Z_{\text{OPTI1}} - Z_{21} & Z_{OPTI2} - Z_{22} & \cdots & Z_{OPTIm} - Z_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ Z_{\text{OPTI1}} - Z_{n1} & Z_{OPTI2} - Z_{n2} & \cdots & Z_{OPTIm} - Z_{nm} \end{bmatrix} \quad \textbf{19}$$

Now at finishing the distance, TCD evaluated in every alternate City to the OPTI using:

$$TCD_{OPTI-Alt} = \sum_{y=1}^{m} | Z_{OPTIy} - Z_{xy} | \quad \textbf{20}$$

This Taxicab distance is measured in conclusion to evaluate many distances on every rank criteria indicators of cities are valued and, lastly, given a ranking. The flow of the algorithm shows in Figure 4.28.

4.3.3 Evaluating "Indian Smart Cities" Ranking

The data set the Table 6 to 9 has depicted from the published report by the "Government of India." The eight criteria have 80 indicators/sub-criteria. All eight criteria are having the same wattage of their 80 indicators. Among the 80 indicators, 59 indicators having high value as optimal/noblest and 21 indicators having low value as optimal/noblest. So, the 59 indicators values calculated in case 1 and 21 values calculated in case 2.

The data set of 20 cities' names mentioned in Table 3 has ranked based on 80 indicators of 8 criteria. The Ligated values in the data set have converted into 0 & 1 for criteria "E-Governance & computerization." The symbolic orientations of cities as is given in Table 3 and the criteria and sub-criteria represented, as shown in Table 4,5. The information about indicators, as explained in Table 5. As data set, the input matrix having 20 rows shows the candidates cite for ranking, and 80 columns of indicators are using for evaluating the ranking of 20 cities. The input matrix is now calculating the optimal value for each criteria's indicators, the 59 indicators having high value as optimal/noblest, and 21 indicators having low value as optimal/noblest. So, the following cases are:

Case -1: The first case the lower value is noblest to real value then $(c = c_{max} - c_i)$.

Case-2: The second case the higher value is noblest to real value then $(c = c_i - c_{min})$.

c=adjusted criteria value, c_{max}= maximum criteria value, c_{min}= minimum criteria value, c_i= criteria value in the data set so value.

Now calculate the standard matrix by using equation 15. the distance matrix calculated with the help of equation number 9. The equation 20 is used to calculate the TCD “Taxicab Composite Distance” between every alternate city to the optimum spot, OPTI. This TCD gives the optimal result for ranking the cities. The results show TCD and rank in Table 16. The result calculated based on 80 indicators of the criteria. The individual ranking of all criteria shows in Table no. 17. This result helps to know the growth of ranking in each dimension of the city. The individual rank compares shows which city needs more growth in which dimensions. The influence of all optimal criteria in the rank. The rank 1 allotted to the smallest TCD value and next to allotted rank two and so on.

The result indications that the Pune has attained the rank number one after evaluation 80 indicators of criteria and next order in Chennai, Jaipur, Ahmadabad, Surat, Delhi, Vishakhapatnam, Indore, Davanagere, Bhopal, Coimbatore, Udaipur, Kochi, Ludhiana, Jabalpur, Bhubaneswar, Guwahati, Solapur, Belgaum, Kakinada correspondingly. Figure 4.29 to 4.50 gives a graphical view of ranking.

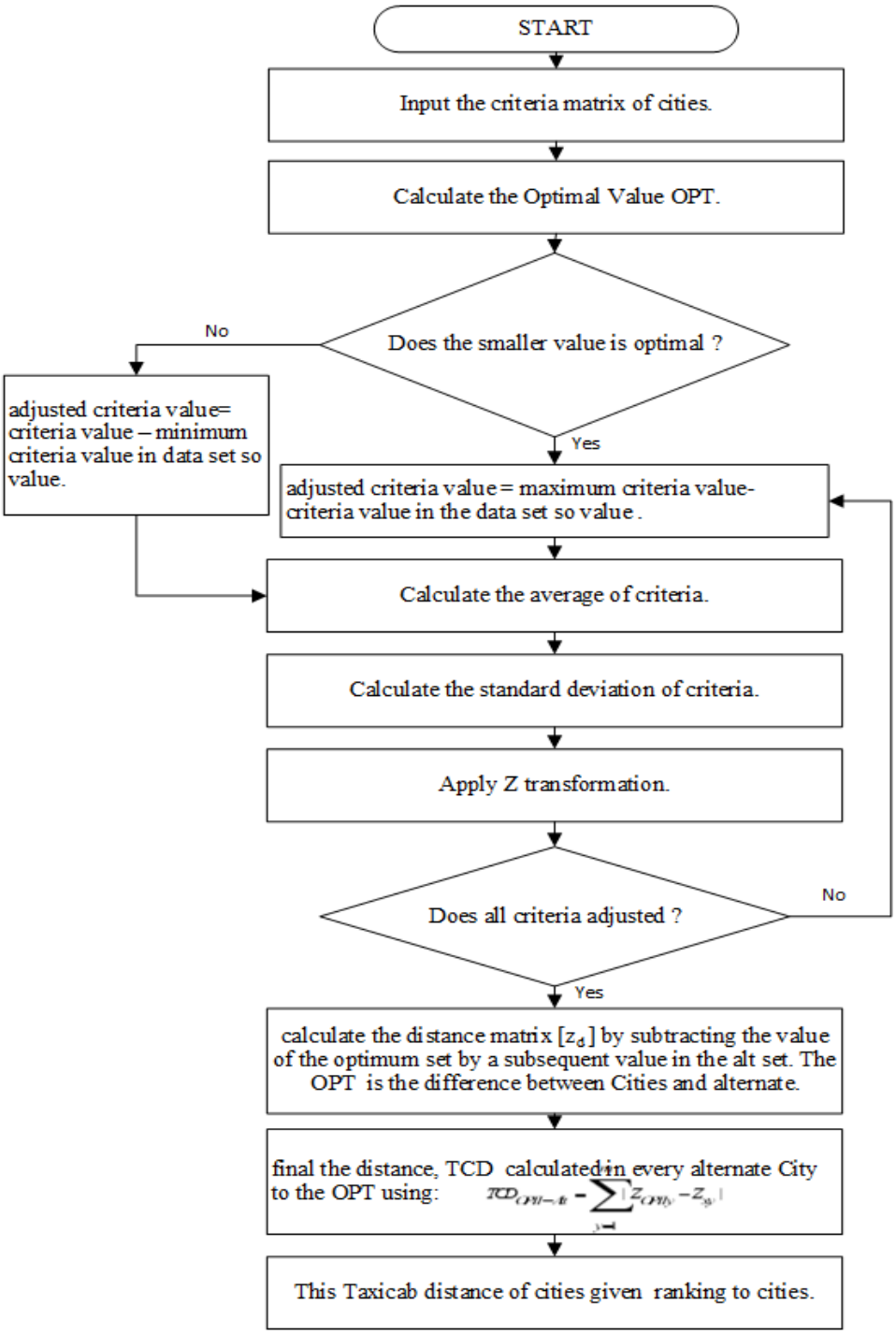


Figure 4-28 Flow Chart of Smart City Ranking Using TDBA

4.3.4 Result and Charts

Table 16 Indian Smart City Ranking Result Using TDBA

S. No	City	TCD	Rank
1	**Kakinada**	199.8288	20
2	**Vishakhapatnam**	153.3956	7
3	**Guwahati**	176.2668	17
4	**Delhi**	149.8682	6
5	**Ahmadabad**	144.4897	4
6	**Surat**	145.5258	5
7	**Belgaum**	183.074	19
8	**Davanagere**	158.2838	9
9	**Kochi**	166.1073	13
10	**Bhopal**	162.7193	10
11	**Indore**	154.2875	8
12	**Jabalpur**	169.6948	15
13	**Solapur**	178.7641	18
14	**Pune**	131.4859	1
15	**Bhubaneswar**	172.8226	16
16	**Ludhiana**	168.1034	14
17	**Jaipur**	142.7679	3
18	**Udaipur**	165.5154	12
19	**Chennai**	131.6461	2
20	**Coimbatore**	163.8349	11

Table 17 Indian Smart City Each Dimension Ranking Result Using TDBA

	A	B	C	D	E	F	G	H	OVERALL
C1	19	9	10	19	20	16	17	1	20
C2	4	5	12	1	13	9	5	2	7
C3	12	4	11	20	9	8	6	3	17
C4	5	18	7	2	2	14	18	4	6
C5	3	15	17	3	3	7	7	18	4
C6	2	19	13	4	4	10	3	5	5
C7	18	1	6	18	16	17	19	6	19
C8	14	14	5	10	10	18	4	7	9
C9	20	2	18	11	7	1	16	20	13
C10	10	13	8	12	14	13	12	8	10
C11	8	20	15	5	6	4	8	9	8
C12	9	17	19	6	18	5	10	10	15
C13	16	6	14	17	19	19	13	11	18
C14	6	8	3	13	1	15	1	12	1
C15	7	10	20	14	17	12	14	13	16
C16	15	16	16	16	8	2	15	14	14
C17	11	3	2	7	11	3	9	15	3
C18	13	11	1	15	15	20	20	19	12
C19	1	7	4	8	5	11	2	16	2
C20	17	12	9	9	12	6	11	17	11

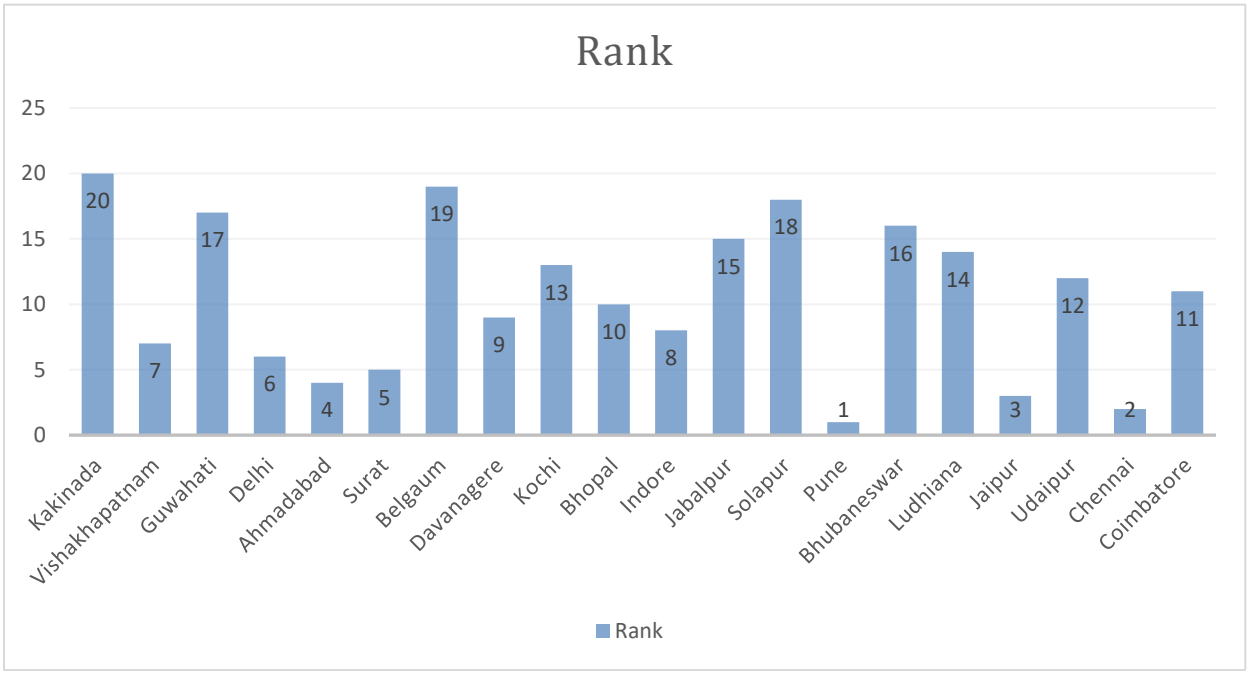


Figure 4-29 Indian Smart City Ranking Result Using TDBA

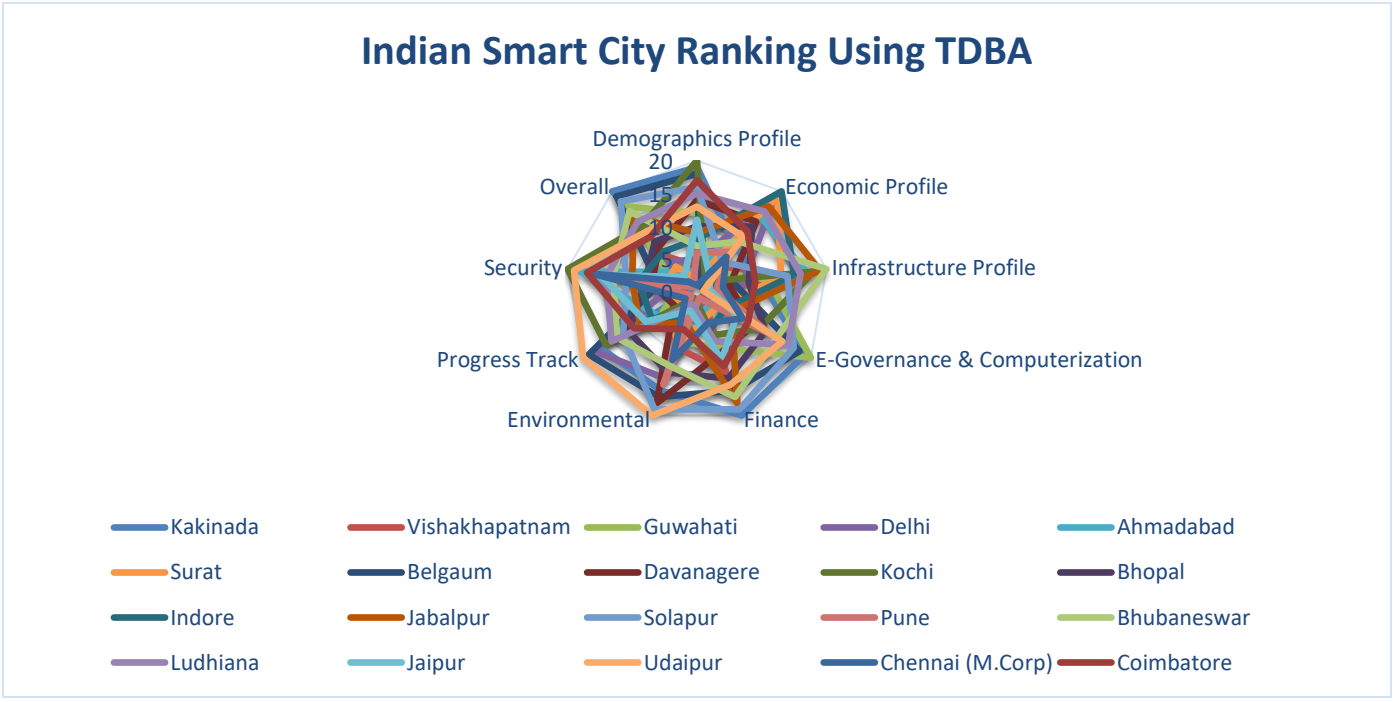


Figure 4-30 Indian Smart City Each Dimension Ranking Result Using TDBA

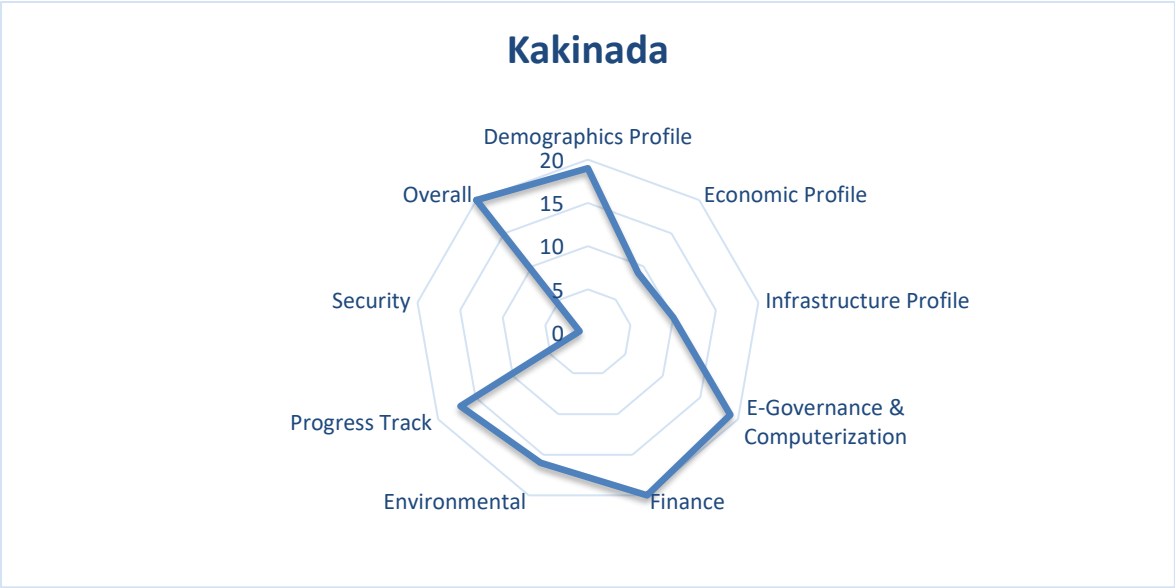


Figure 4-31 Kakinada Dimensions Ranking using TDBA

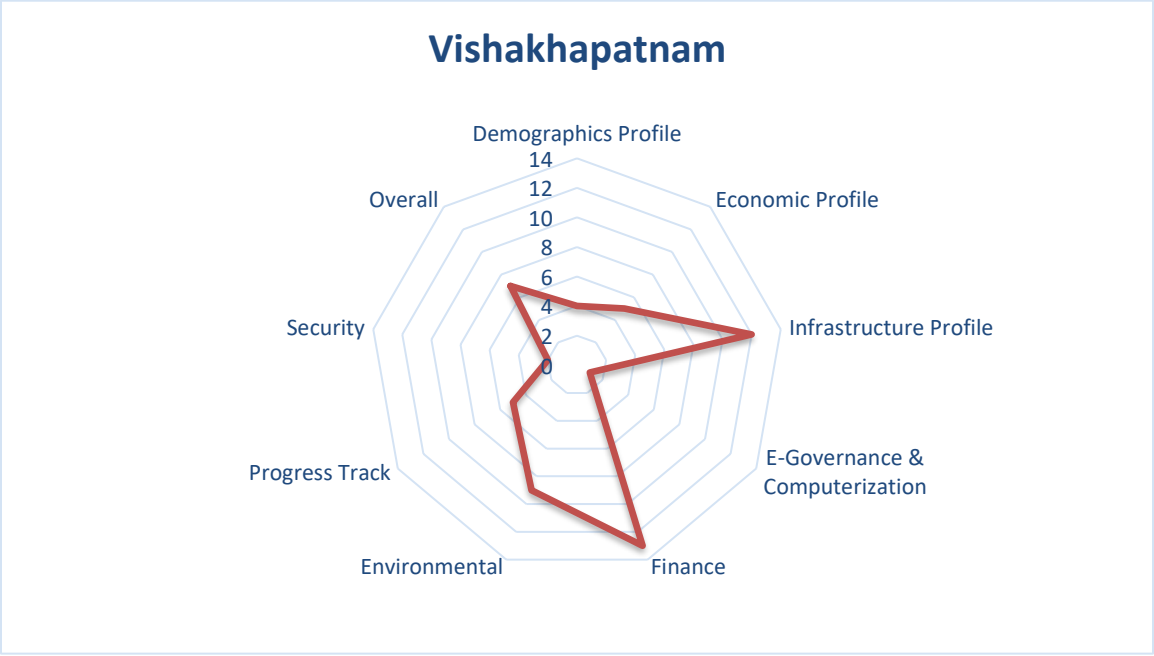


Figure 4-32 Vishakhapatnam Dimensions Ranking using TDBA

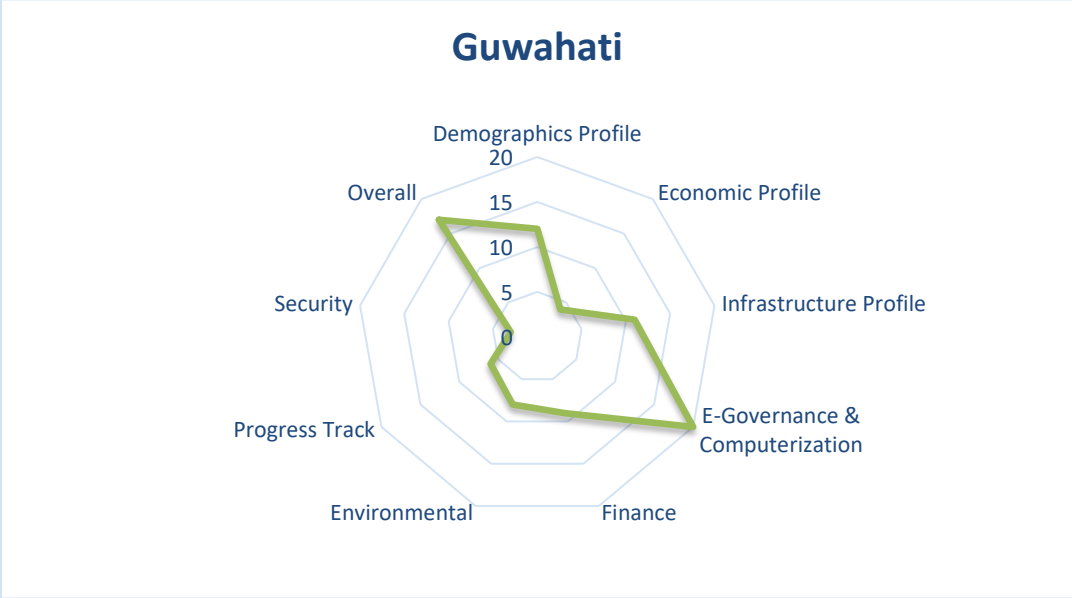


Figure 4-33 Guwahati Dimensions Ranking using TDBA

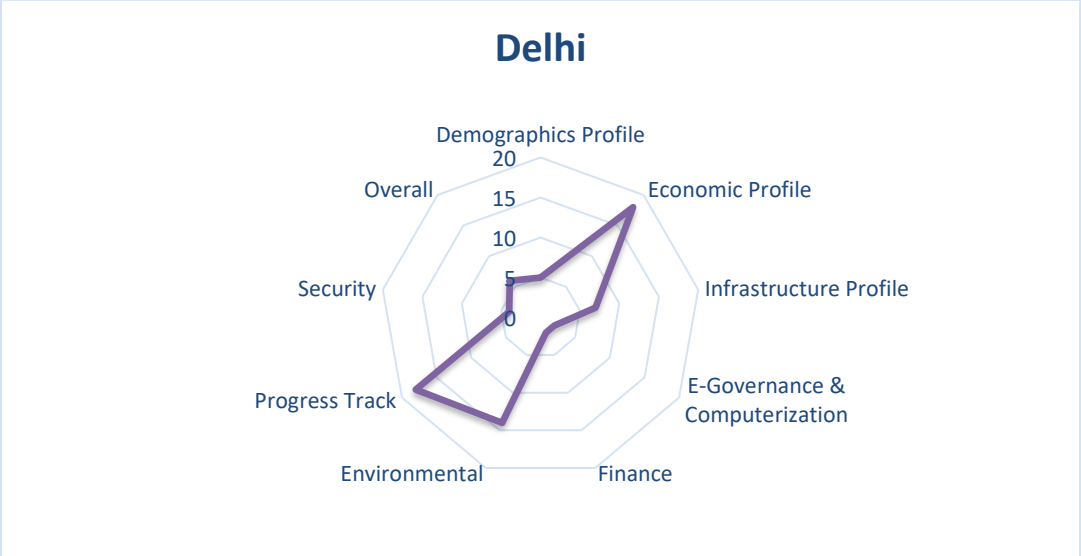


Figure 4-34 Delhi Dimensions Ranking using TDBA

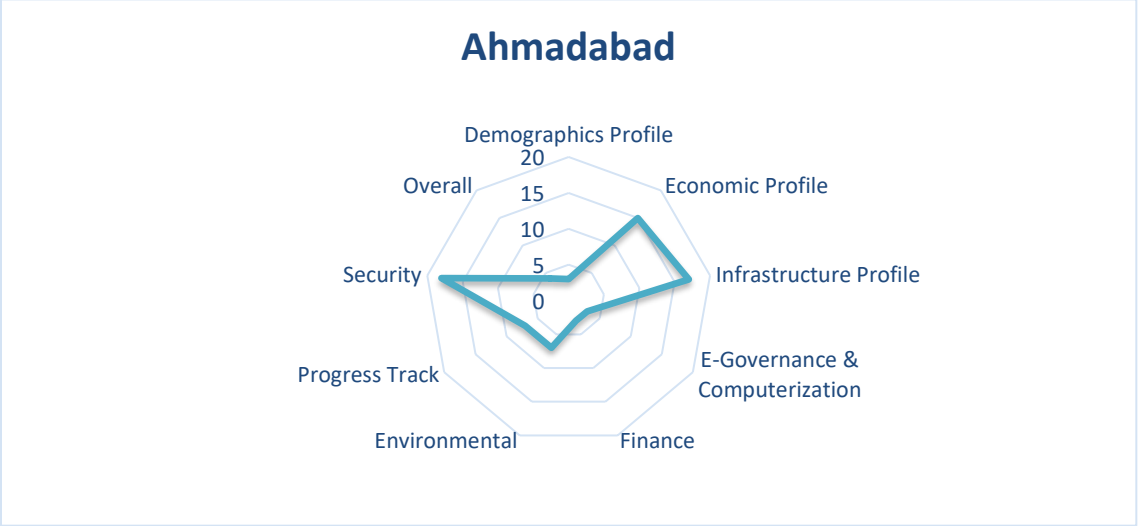


Figure 4-35 Ahmadabad Dimensions Ranking using TDBA

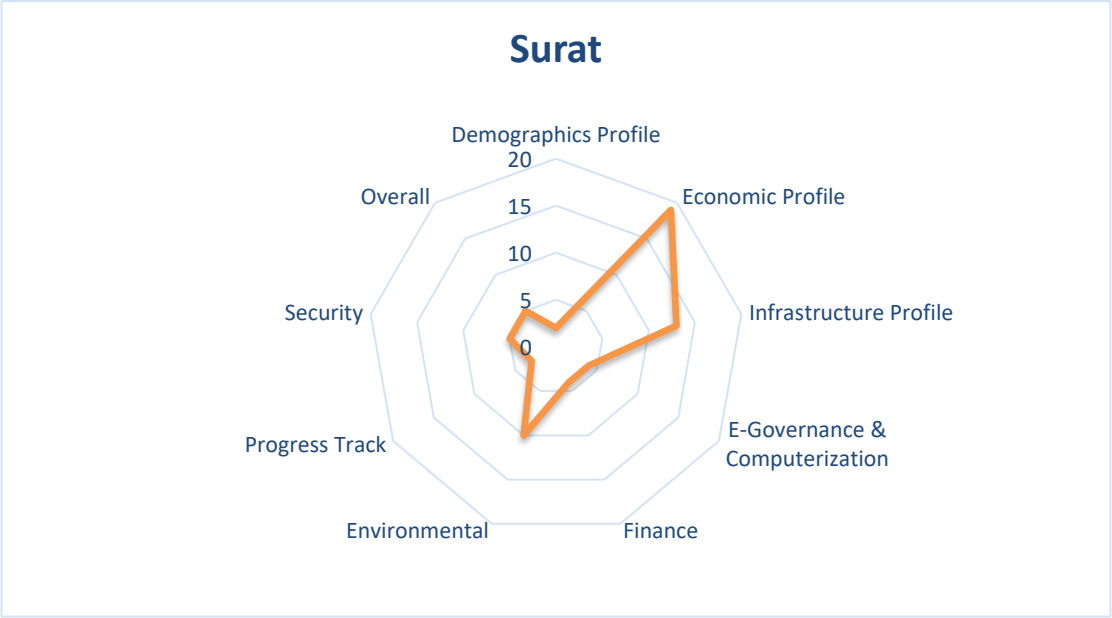


Figure 4-36 Surat Dimensions Ranking using TDBA

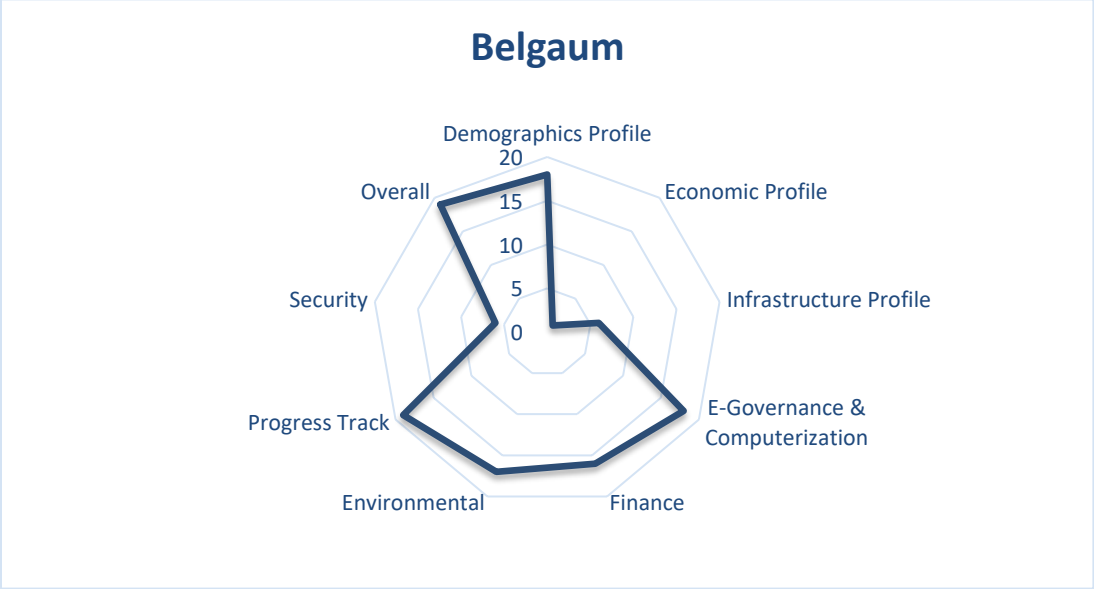


Figure 4-37 Belgaum Dimensions Ranking using TDBA

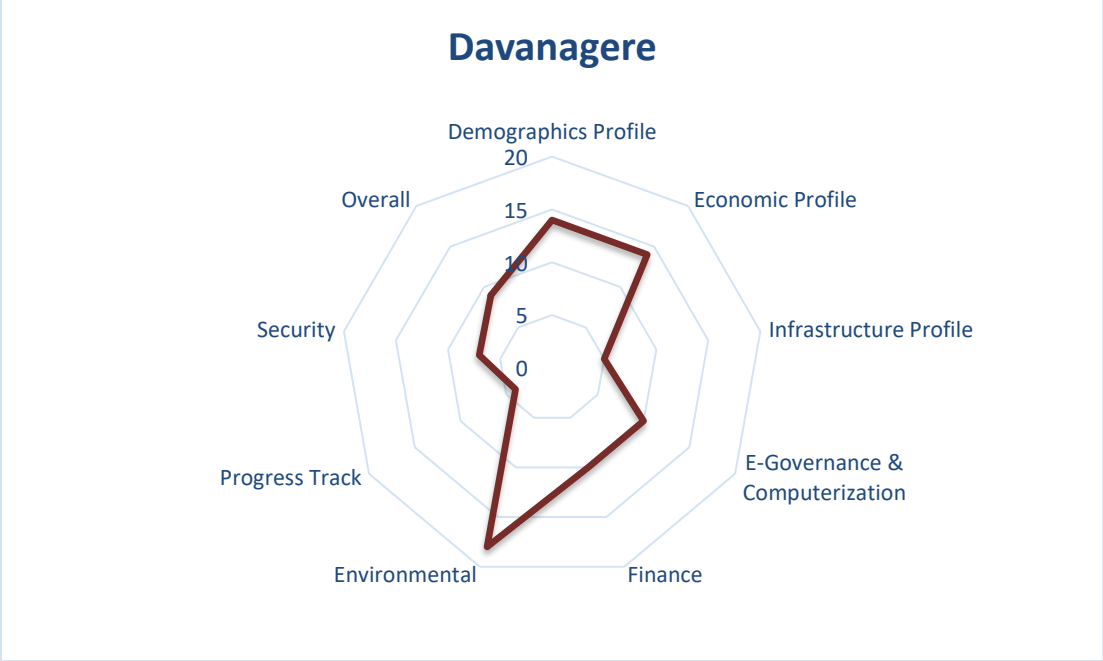


Figure 4-38 Davanagere Dimensions Ranking using TDBA

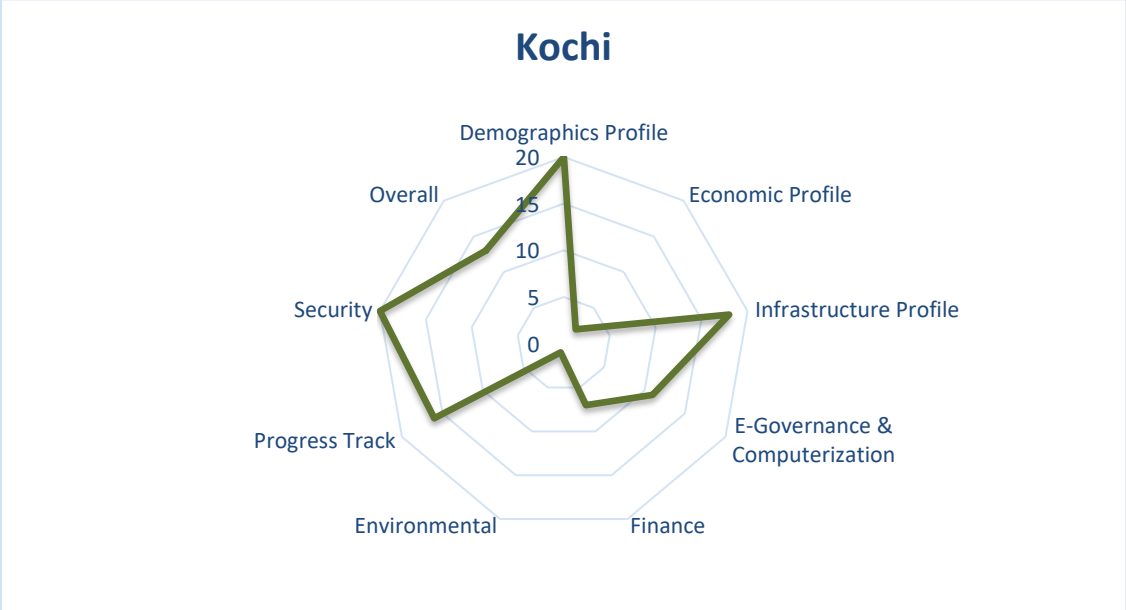


Figure 4-39 Kochi Dimensions Ranking using TDBA

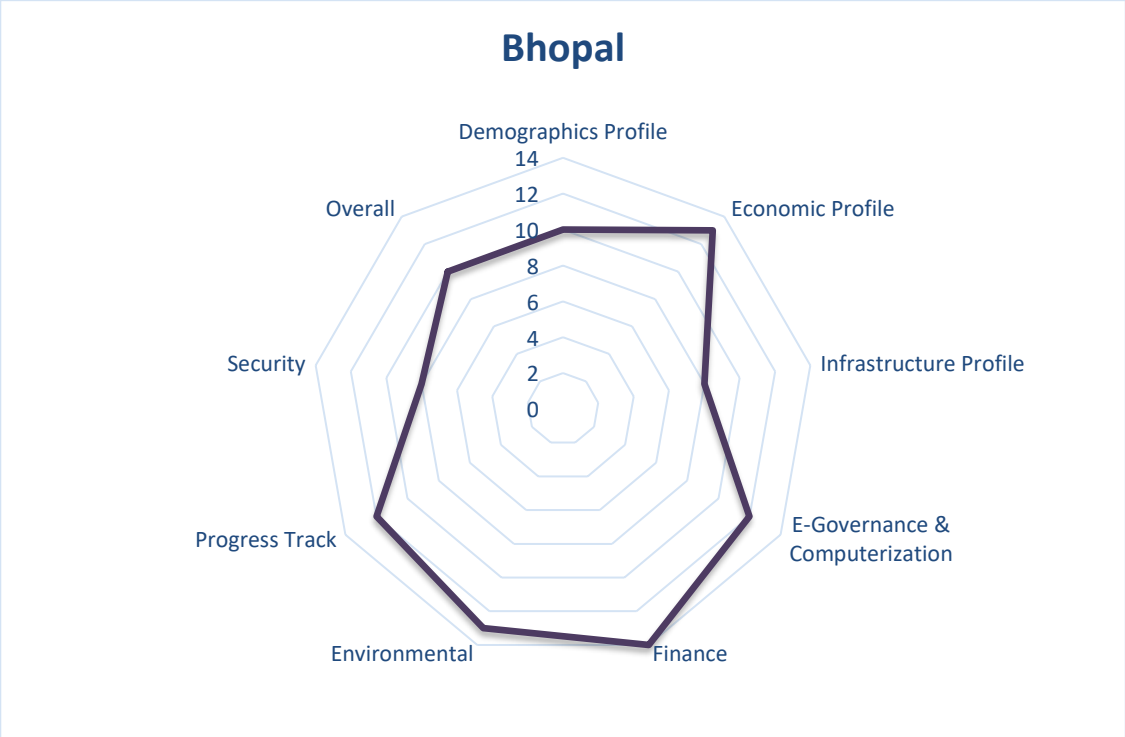


Figure 4-40 Bhopal Dimensions Ranking using TDBA

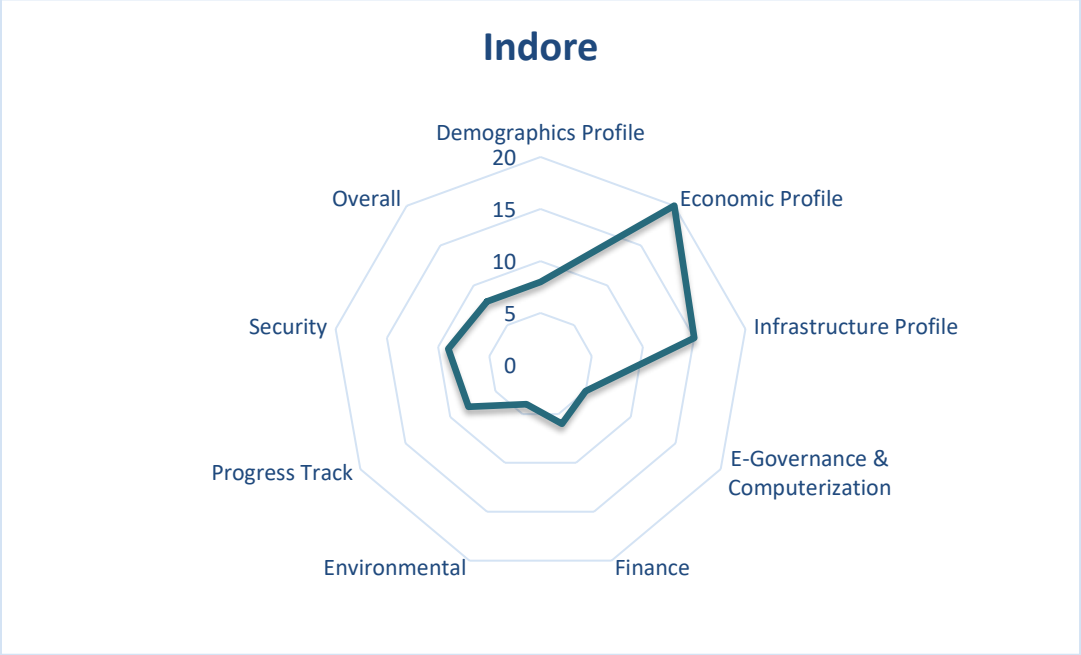


Figure 4-41 Indore Dimensions Ranking using TDBA

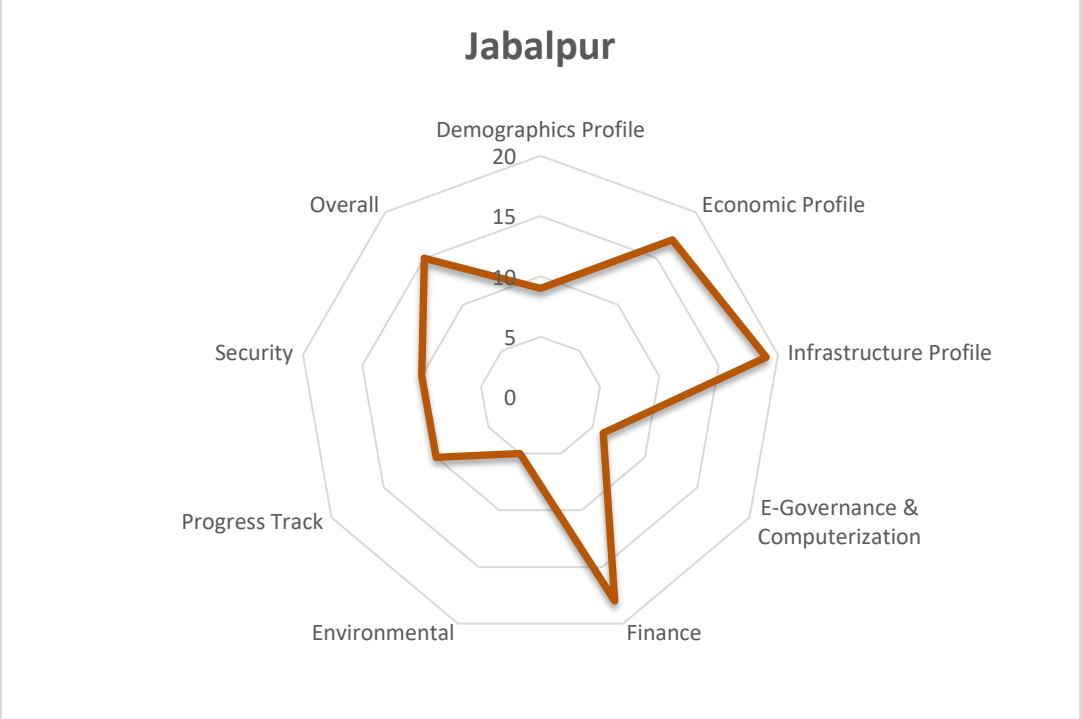


Figure 4-42 Jabalpur Dimensions Ranking using TDBA

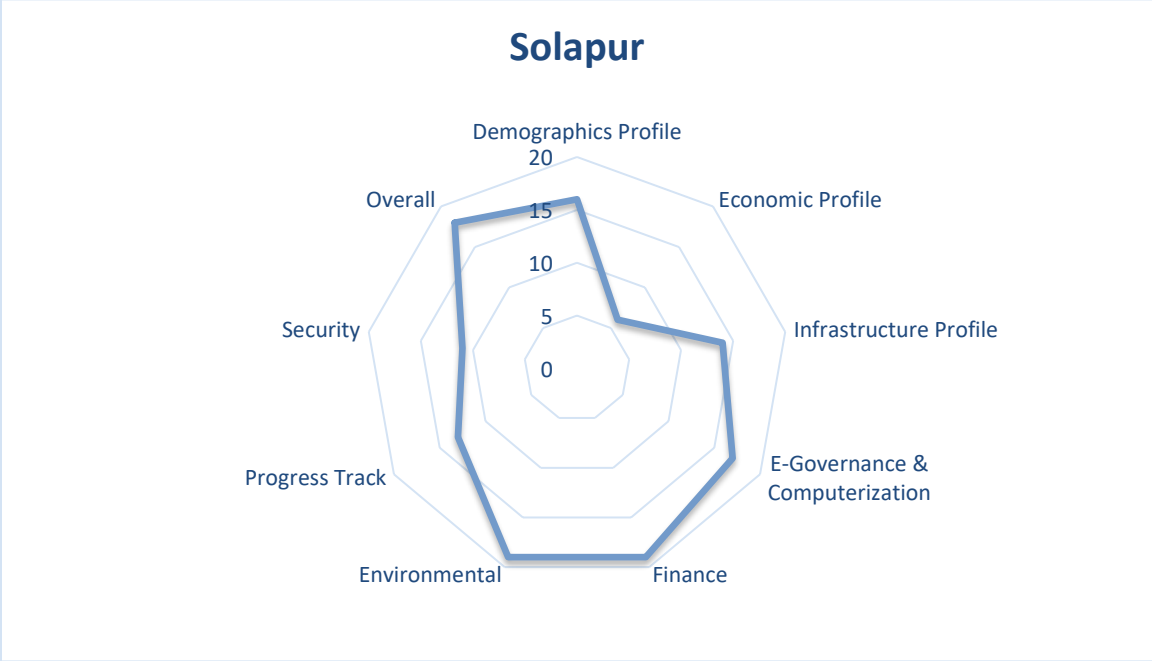


Figure 4-43 Solapur Dimensions Ranking using TDBA

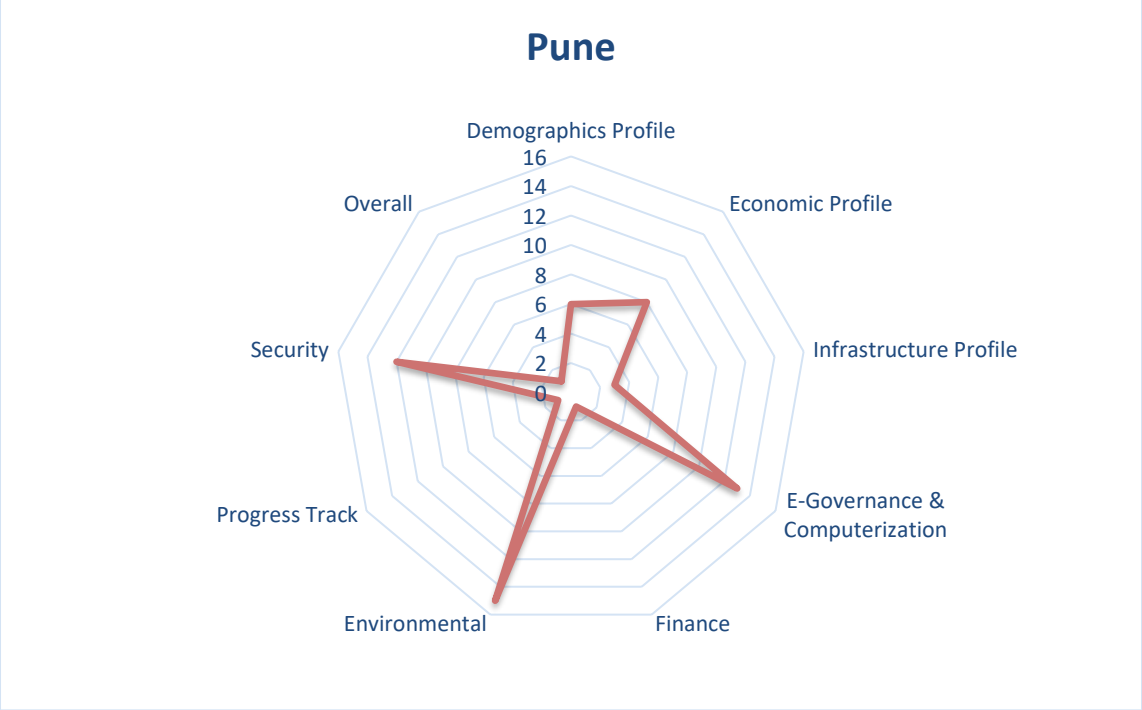


Figure 4-44 Pune Dimensions Ranking using TDBA

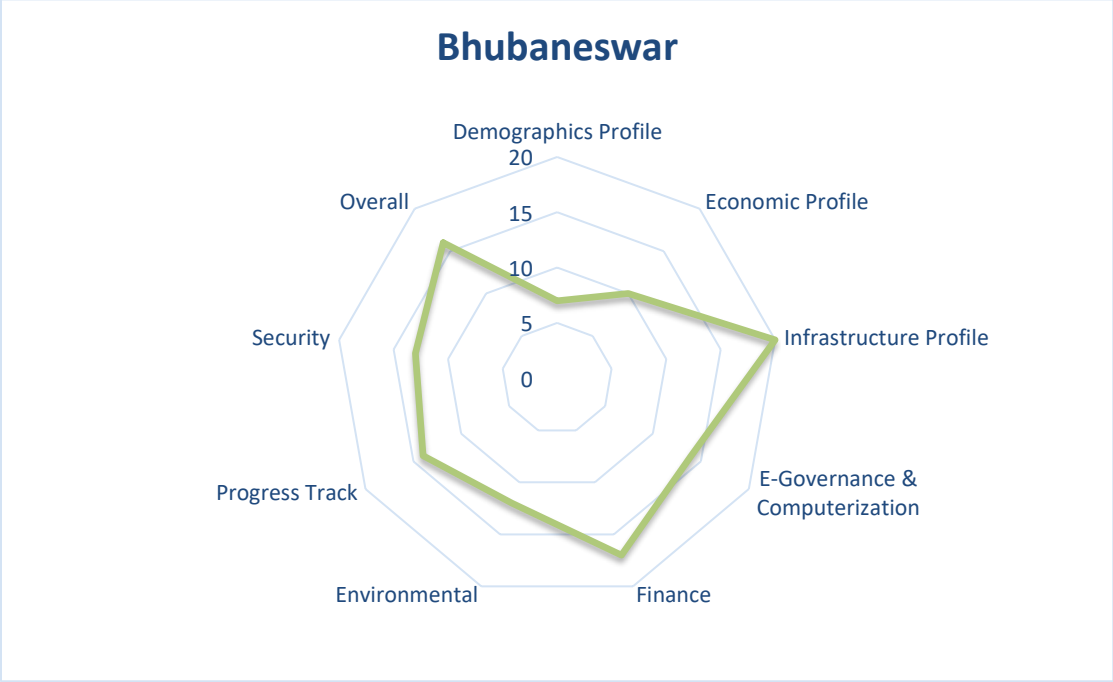


Figure 4-45 Bhubaneshwar Dimensions Ranking using TDBA

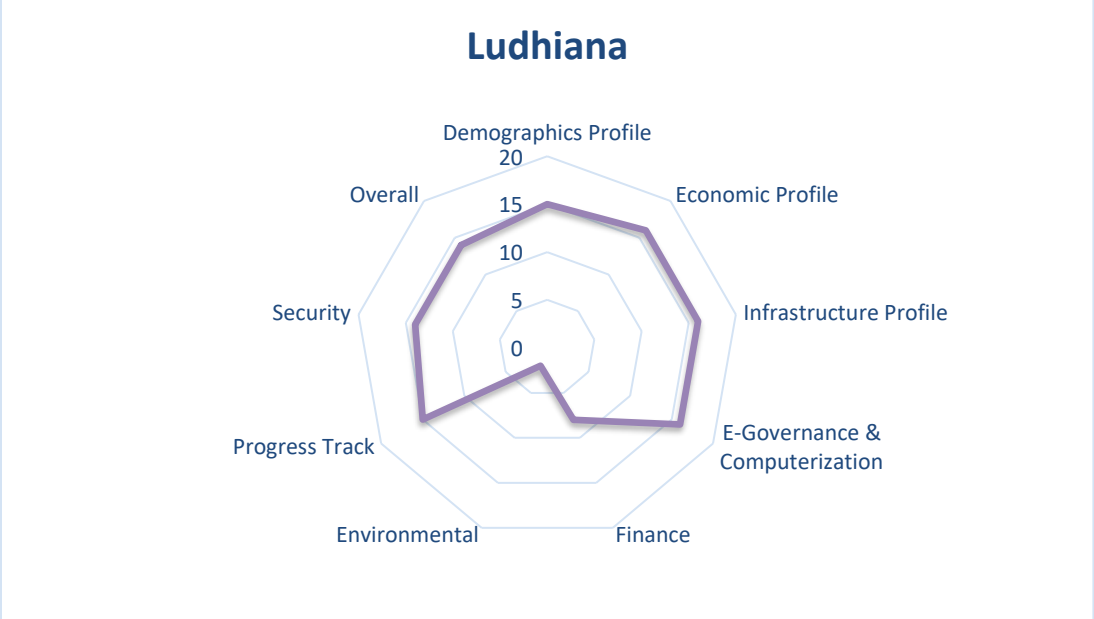


Figure 4-46 Ludhiana Dimensions Ranking using TDBA

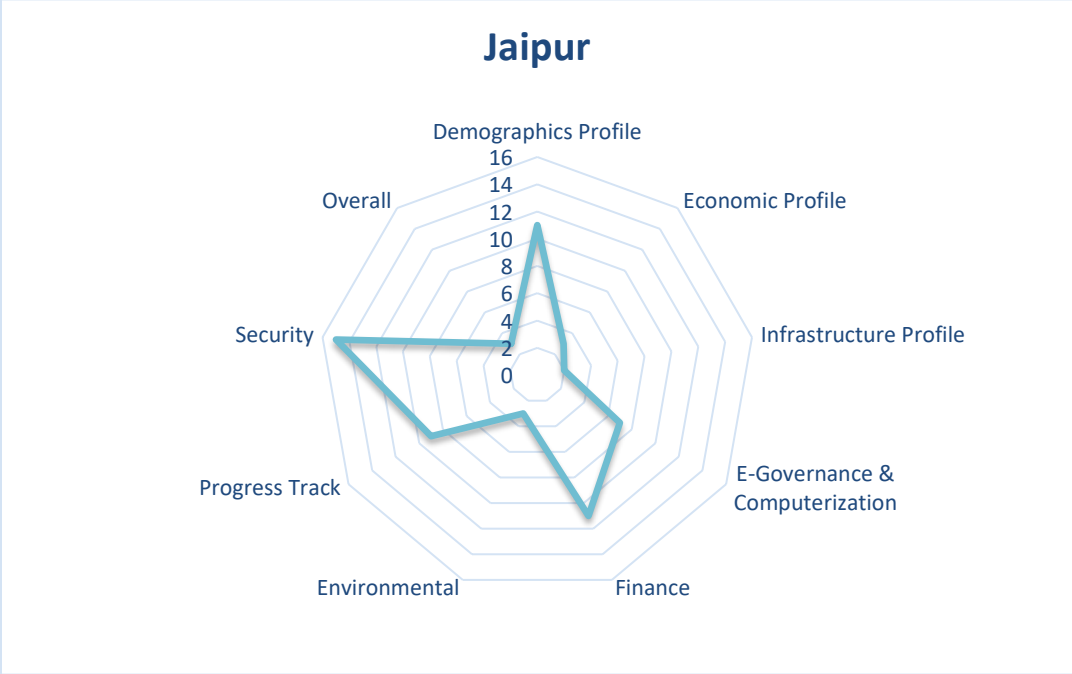


Figure 4-47 Jaipur Dimensions Ranking using TDBA

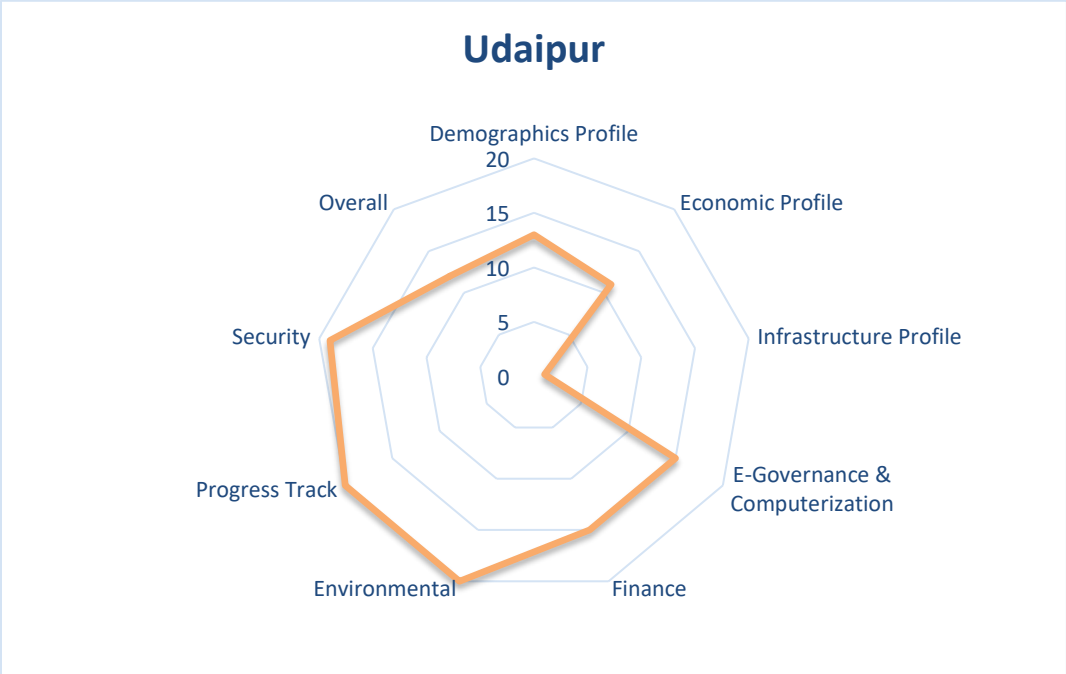


Figure 4-48 Udaipur Dimensions Ranking using TDBA

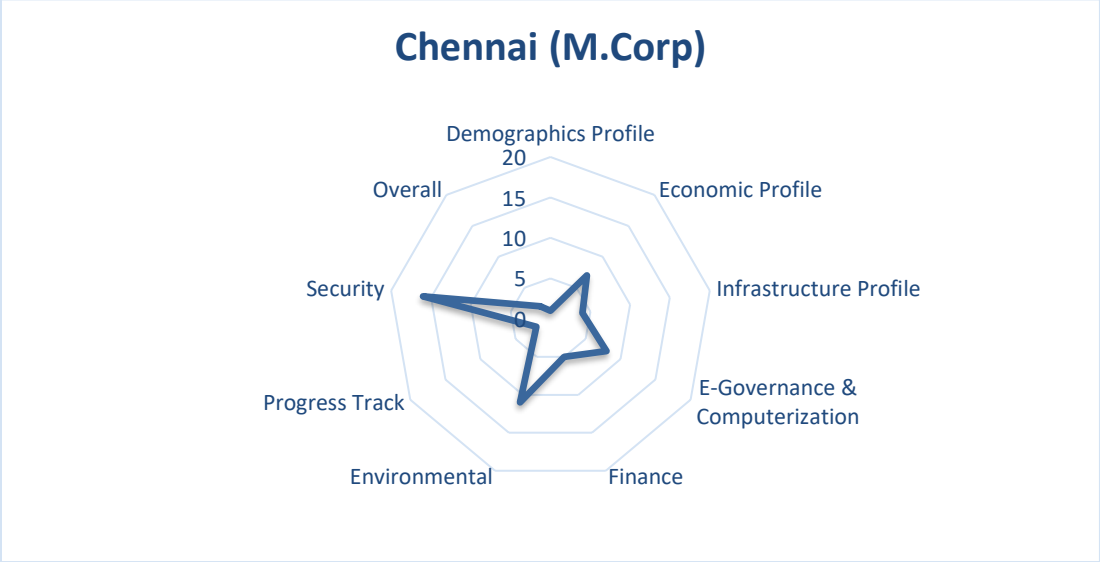


Figure 4-49 Chennai Dimensions Ranking using TDBA

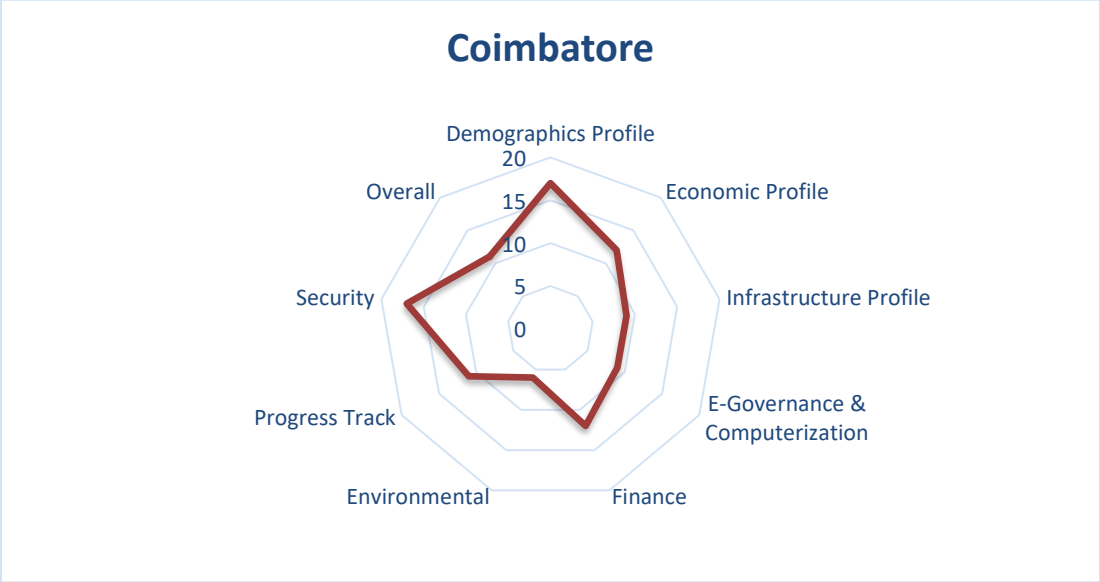


Figure 4-50 Coimbatore Dimensions Ranking using TDBA

4.4 The Spearman's rank correlation coefficient

The “Spearman's rank correlation coefficient” [133] is computed between the DBA ranking result in TDBA ranking Result.

The formula of “Spearman's rank correlation coefficient” is

$$r_s = 1 - \frac{\sum d_i^2}{n(n^2 - 1)}$$

The value of r_s computed as 0.96. It indicates a strong positive association between the ranks of DBA and TDBA.

4.5 Summary

This chapter has explained two algorithms DBA and TDBA, for ranking Smart City on the base of 80 indicators of twenty Indian smart cities. The graphical presentation of rank given a perfect comparative understanding of the growth of smart cities. Spearman's rank correlation coefficient of both ranking algorithms shows a similarity association of both algorithms in the same direction ranking.

The next chapter concludes the research work with the future scope of the research.

Chapter Five: Conclusion and Scope for Future Research

This chapter emphasizes the detail conclusion and also looks at the opportunities for future research based on the given algorithm and methodology.

5.1 Introduction

Smart cities have the most indispensable part in altering distinctive regions of human life, touching segments. Big Data can do in the field of “Make in India,” Smart Cities. It will lead to the planning of smart cities. The smart cities developed with the help of big data will have all the resources needed for the growing population demand. Big data is used to collect data from various sources of information.

The information is used to rank the cities according to the “Indian Smart City Model.” The cities can change and update their planning according to city rank. The data gathering from different departments and surveys of the cities. This information used for evaluation of city rank using the DBA and TDBA approach. DBA and TDBA is a mathematical tool that has been used to aggregate and convert data into a standard form that is used to rank “Indian smart cities.” The result shows the growth of the city in the individual dimension of the city through the ranking comparison of each dimension of the smart city. The policymaker can make decisions based on the result for achieving the successful, manageable target of the smart city.

5.2 Major Results

The specific major findings of the research work are:

- The smart city model developed for the Indian smart cities with eight dimensions, “Economic Profile,” “Demographics Profile,” e-Governance& Computerization, “Infrastructure Profile,” “Environmental,” “Finance,” “Progress Track,” and “Security.” The model has eighty indicators, as in Table 5.
- The Indian Smart cites data collected according to the model to find the rank of smart cities, and the individual dimensions rank also finds to show the growth level of the individual cites in each dimension.
- The distance-based algorithm (DBA) applied for ranking of twenty different Indian smart cities based on a set of eighty selection criteria/indicators. The outcome demonstrates that the Chennai has accomplished position number one after investigation 80 criteria and next succession is Pune, Jaipur, Ahmadabad, Surat, Vishakhapatnam, Indore, Coimbatore, Bhopal, Davanagere, Delhi, Udaipur, Ludhiana, Bhubaneswar, Jabalpur, Kochi, Guwahati, Belgaum, Solapur, Kakinada respectively.
- The Taxicab Distance-Based Approach has been proposed to rank the “Indian smart city.” The TDBA finds the optimal solution on the bases of “Indian smart city” indicators multiple values. The result indications that the Pune has attained the rank number one after evaluation 80 indicators of criteria and next order in Chennai, Jaipur, Ahmadabad, Surat, Delhi, Vishakhapatnam, Indore, Davanagere, Bhopal, Coimbatore, Udaipur, Kochi, Ludhiana, Jabalpur, Bhubaneswar, Guwahati, Solapur, Belgaum, Kakinada respectively.

5.3 Scope for Future Research

This technique needs improvement due to the variety of changes in data. The future is going toward the information age, which provides more future work for cities. The Indian cities are in the developing phase, so the scope for improvement is always occurring.

The smart city data collection process can be more advanced with the development of new technology. The model can adopt new advanced techniques. The IoT can embed in the future for collecting the information bases on sensors. The smart city models each dimension can be exploring and evolve with new technological terms.

References

[1] "A tool for Smart Cities Index Evaluating Cities," *Indian School of Business*, New Delhi - 110067, Feb-2017.

[2] K. Sharma and S. Tayal, "Indian Smart City Ranking Model," *Int. J. Recent Technol. Eng. ISSN2277-3878*, vol. 8, no. 2, pp. 4820–4832, 2019.

[3] A. V. Michael Lindfield, Florian Steinberg, Steffen Lehmann, Thomas Hurst, Debra Lam, Malcolm Ball, Lloyd Wright, Alan Baird, Theresa Audrey O. Esteban, Reynar R. Rollan, Vergel Latay, Vince Michael Docta, *Urban Development Series: Green cities*, no. 1. 2012.

[4] O. F. Cities, "Trends in Smart City Development National League of Cities," *Natl. Leag. Cities-Center City Solut. Appl. Res.*, 2016.

[5] A. Datta, "A 100 smart cities, a 100 utopias," *Dialogues Hum. Geogr.*, 2015.

[6] E. Estevez, N. V. Lopes, and T. Janowski, "Smart Sustainable Cities. Reconnaissance Study," p. 330, 2016.

[7] S. Bhattacharya and S. Rathi, "Reconceptualising Smart Cities: A Reference Framework for India," *Cent. Study Sci. Technol. Policy*, vol. September, p. 80, 2015.

[8] "Smart Cities Mission Statement & Guidelines," *Smart Cities Mission Statement Guidel.*, vol. 3, no. 3, pp. 42–49, 2015.

[9] S. Bhattacharya, S. Rathi, S. A. Patro, and N. Tepa, "Reconceptualising Smart Cities:," no. September, p. 80, 2015.

[10] D. Pérez-González and R. Díaz-Díaz, "Public services provided with ICT in the smart city environment: The case of Spanish cities," *J. Univers. Comput. Sci.*, vol. 21, no. 2, pp. 248–267, 2015.

[11] B. Benamrou, B. Mohamed, A. S. Bernoussi, and O. Mustapha, "Ranking models of smart cities," *Colloq. Inf. Sci. Technol. Cist*, pp. 872–879, 2017.

[12] P. P. Berrone *et al.*, "IESE Cities in Motion Index," 2017.

[13] J. García-Arurmí, V. Martinez, L. Sararols, and B. Corcostegui, "Vitreoretinal surgery for cystoid macular edema associated with retinitis pigmentosa,"

Ophthalmology, vol. 110, no. 6, pp. 1164–1169, 2003.

[14] S. Tayal, N. Nagwal, and K. Sharma, "Role of big data in make in India," in *Advances in Intelligent Systems and Computing*, 2018, vol. 564, pp. 431–437.

[15] M. M. Rathore, A. Ahmad, A. Paul, and S. Rho, "Urban planning and building smart cities based on the Internet of Things using Big Data analytics," *Comput. Networks*, 2016.

[16] M. Fazio, A. Celesti, A. Puliafito, and M. Villari, "Big data storage in the cloud for smart environment monitoring," *Procedia Comput. Sci.*, vol. 52, no. 1, pp. 500–506, 2015.

[17] S. Consoli *et al.*, "A Smart City Data Model based on Semantics Best Practice and Principles," *Proc. 24th Int. Conf. World Wide Web - WWW '15 Companion*, pp. 1395–1400, 2015.

[18] E. Bertino, "Big data - Opportunities and challenges: Panel position paper," *Proc. - Int. Comput. Softw. Appl. Conf.*, pp. 479–480, 2013.

[19] M. Krämer, "Controlling the processing of smart city data in the cloud with domain-specific languages," *Proc. - 2014 IEEE/ACM 7th Int. Conf. Util. Cloud Comput. UCC 2014*, pp. 824–829, 2014.

[20] A. M. S. S. Osman, "A novel big data analytics framework for smart cities," *Futur. Gener. Comput. Syst.*, vol. 91, pp. 620–633, Feb. 2019.

[21] Smart Cities Council, "Smart Cities Open Data Guide," 2014.

[22] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: promise and potential," *Heal. Inf. Sci. Syst.*, vol. 2, no. 1, p. 3, 2014.

[23] D. Paliwal, P. Chauhan, S. Tayal, Bt. Mait, and A. Professor MAIT, "Categorizing Regions using R programming Regional Analysis using R programming," in *Futuristic Trends in Computational Analysis and Knowledge Management*, 2015, p. 2014.

[24] A. Kumar, "Distributed and Big Data Storage Management in Grid Computing," *Int. J. Grid Comput. Appl.*, vol. 3, no. 2, pp. 19–28, 2012.

[25] Leonidas G. Anthopoulos, *Understanding Government or an Tool for Smart Smart*

Cities: A Industrial Trick? Springer, 2017.

[26] R. Achaerandio, R. Bigliani, J. Curto, and G. Gallotti, “Smart Cities Analysis in Spain 2012 — The Smart Journey,” 2012.

[27] “Smart Cities Readiness Guide,” *Smart Cities Counc.*, 2015.

[28] R. K. Jain, J. M. F. Moura, and C. E. Kontokosta, “Big data + big cities: Graph signals of urban air pollution [Exploratory SP],” *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 130–136, 2014.

[29] G. Piro, I. Cianci, L. A. A. Grieco, G. Boggia, and P. Camarda, “Information centric services in Smart Cities,” *J. Syst. Softw.*, vol. 88, no. 1, pp. 169–188, Feb. 2014.

[30] S. Joshi, S. Saxena, T. Godbole, and Shreya, “Developing Smart Cities: An Integrated Framework,” *Procedia Comput. Sci.*, 2016.

[31] A. J. Jara, D. Genoud, and Y. Bocchi, “Big data for smart cities with KNIME a real experience in the SmartSantander testbed,” *Softw. - Pract. Exp.*, vol. 45, no. 8, pp. 1145–1160, 2015.

[32] M. Technologies, “Investigation into Big Data Impact on Digital Marketing Ketty Grishikashvili, Open University, UK S. Dibb, Open University, UK M. Meadows, Open University, UK,” no. October, pp. 26–37, 2014.

[33] Y. Pan, Y. Tian, X. Liu, D. Gu, and G. Hua, “Urban Big Data and the Development of City Intelligence,” *Engineering*, 2016.

[34] J. M. Shapiro, “Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital,” *Rev. Econ. Stat.*, vol. 88, no. 2, pp. 324–335, May 2006.

[35] “Cities in Motion Strategies Research Projects IESE Cities in Motion Strategies.”

[36] “IESE Cities in Motion Index,” 2015.

[37] chombosan, “IESE Cities in Motion Index IESE Cities in Motion Index Introduction: The Need for a Global Vision 10,” 2018.

[38] “Data Analyst Cities - 2THINKNOW DATA.” [Online]. Available: https://2thinknow.com/data-analyst-cities/. [Accessed: 12-Oct-2019].

[39] 2thinknow, “Innovation Cities™ Index 2018 - Global City Ranking by

2thinknow," *innovation-cities*, 2018. [Online]. Available: https://www.innovation-cities.com/innovation-cities-index-2018-global/13935/. [Accessed: 19-Jun-2019].

[40] A. Caragliu, C. Del Bo, and P. Nijkamp, "Smart cities in Europe," *J. Urban Technol.*, vol. 18, no. 2, pp. 65–82, Apr. 2011.

[41] C. Europea, "European smart cities," *Conf. Stamp.*, 2007.

[42] L. Anthopoulos, M. Janssen, and V. Weerakkody, "A Unified Smart City Model (USCM) for Smart City Conceptualization and Benchmarking," *Int. J. Electron. Gov. Res.*, vol. 12, no. 2, pp. 77–93, Apr. 2016.

[43] L. E. Edvinsson, "Aspects on the city as a knowledge tool," *J. Knowl. Manag.*, vol. 10, no. 5, pp. 6–13, Sep. 2006.

[44] N. S. L.- Ericsson, "Networked Society Index," 2016.

[45] T. Castro and M. Quartin, "A summary of the Green City Index research series," *Mon. Not. R. Astron. Soc. Lett.*, vol. 443, no. 1, p. 25, 2014.

[46] P. B. Tno, S. J. Tno, and H. N. Ait, "Recommendations for a smart city index," no. January, 2017.

[47] "Network Society City Index 2014 Appendix 1-Methodology," Ericsson Stockholm, Sweden, 2014.

[48] B. Mattoni, F. Gugliermetti, and F. Bisegna, "A multilevel method to assess and design the renovation and integration of Smart Cities," *Sustain. Cities Soc.*, 2015.

[49] M. Wang, J. Wang, and F. Tian, "'City intelligent energy and transportation network policy' 'based on the big data analysis,'" *Procedia Comput. Sci.*, vol. 32, pp. 85–92, 2014.

[50] M. Batty *et al.*, "Smart cities of the future," *Eur. Phys. J. Spec. Top.*, vol. 214, no. 1, pp. 481–518, Nov. 2012.

[51] J. Minx *et al.*, "CLIMATECON Climatecon Working Paper Series DEVELOPING A PRAGMATIC APPROACH TO ASSESS URBAN METABOLISM IN EUROPE Developing a Pragmatic Approach to Assess Urban Metabolism in Europe," Department of Climate Change Economics Technische Universität Berlin, 2011.

[52] S. P. Mohanty, U. Choppali, and E. Kougianos, “Everything you wanted to know about smart cities,” *IEEE Consum. Electron. Mag.*, 2016.

[53] UN-Habitat, “Issue Paper on Smart Cities,” *HABITAT III Issue Pap.*, vol. 21, no. May, pp. 0–10, 2015.

[54] ITU-T SG20, “Key performance indicators for smart sustainable cities to assess the achievement of sustainable development goals,” *Recommendation ITU-T Y.4903/L.1603*. pp. 1–50, 2017.

[55] Z. Sang, “standard\ITU-T\ITU-T S,Sept 2014\Bangkok training\Ziqin smart city KPIs and monitoring V2,” China, 2014.

[56] J. M. Shapiro, “Growth Effects of Human Capital,” vol. 88, no. May, pp. 324–335, 2006.

[57] L. M. A. Bettencourt, “The Uses of Big Data in Cities,” *Big Data*, vol. 2, no. 1, pp. 12–22, 2014.

[58] K. Kourtit, P. Nijkamp, and D. Arribas, “Smart cities in perspective - a comparative European study by means of self-organizing maps,” *Innovation*, vol. 25, no. 2, pp. 229–246, 2012.

[59] R. Giffinger, “Smart cities ranking: an effective instrument for the positioning of the cities?,” *ACE Archit. City Environ.*, no. June, 2017.

[60] I. A. T. Hashem *et al.*, “The role of big data in smart city,” *Int. J. Inf. Manage.*, vol. 36, no. 5, pp. 748–758, Oct. 2016.

[61] George Firican, “The 10 Vs of Big Data | Transforming Data with Intelligence,” *February 8, 2017*, 2017. [Online]. Available: https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx. [Accessed: 06-Oct-2019].

[62] W. Topics and L. Objectives, “Big Data & Ontology Organized by Learning Objectives How to Apply Electrical Engineering Department IIT Delhi , New Delhi-110016,” 2015.

[63] M. Sookhak, H. Tang, Y. He, and F. R. Yu, “Security and Privacy of Smart Cities: A Survey, Research Issues and Challenges,” *IEEE Commun. Surv. Tutorials*, vol.

PP, no. c, p. 1, 2018.

[64] E. K. Lee, P. Chu, and R. Gadh, “Fine-grained access to smart building energy resources,” *IEEE Internet Comput.*, vol. 17, no. 6, pp. 48–56, 2013.

[65] S. Zygiaris and Z. Sotiris, “Smart City Reference Model: Assisting Planners to Conceptualize the Building of Smart City Innovation Ecosystems,” *J. Knowl. Econ.*, vol. 4, no. June, pp. 217–231, 2012.

[66] L. Rodríguez-Mazahua, C. A. Rodríguez-Enríquez, J. L. Sánchez-Cervantes, J. Cervantes, J. L. García-Alcaraz, and G. Alor-Hernández, “A general perspective of Big Data: applications, tools, challenges and trends,” *J. Supercomput.*, vol. 72, no. 8, pp. 3073–3113, Aug. 2016.

[67] S. P. Pawar, “Smart City with Internet of Things (Sensor networks) and Big Data,” *Academia.edu*, vol. 1, no. 9860027825, p. 10, 2015.

[68] U. Aguilera, O. Peña, O. Belmonte, and D. López-de-Ipiña, “Citizen-centric data services for smarter cities,” *Futur. Gener. Comput. Syst.*, vol. 76, pp. 234–247, 2017.

[69] M. Roscia, M. Longo, and G. C. Lazaroiu, “Smart City by multi-agent systems,” *Proc. 2013 Int. Conf. Renew. Energy Res. Appl. ICRERA 2013*, no. October, pp. 371–376, 2013.

[70] R. K. R. Kummitha and N. Crutzen, “How do we understand smart cities? An evolutionary perspective,” *Cities*, vol. 67, pp. 43–52, 2017.

[71] G. P. Hancke, B. de C. de Silva, G. P. Hancke, B. Silva, and G. Hancke, Jr., *The role of advanced sensing in smart cities*, vol. 13, no. 1. 2013.

[72] P. D. Mullins and S. T. Shwayri, “Green Cities and ‘IT839’: A New Paradigm for Economic Growth in South Korea,” *J. Urban Technol.*, vol. 23, no. 2, pp. 47–64, Apr. 2016.

[73] H. Kumar, M. K. Singh, M. P. Gupta, and J. Madaan, “Moving towards smart cities: Solutions that lead to the Smart City Transformation Framework,” *Technol. Forecast. Soc. Change*, 2018.

[74] British Standars Institute, “Mapping Smart City Standards,” p. 47, 2014.

[75] C. In, S. Up, and S. Cities, “MODERN,” no. January, 2015.

[76] M. N. Kamel Boulos and N. M. Al-Shorbaji, “On the Internet of Things, smart cities and the WHO Healthy Cities,” *Int. J. Health Geogr.*, vol. 13, 2014.

[77] D. Li, J. Cao, and Y. Yao, “Big data in smart cities,” *Sci. China Inf. Sci.*, vol. 58, no. 10, pp. 1–12, 2015.

[78] S. T. Shwayri, “A Model Korean Ubiquitous Eco-City? The Politics of Making Songdo,” *J. Urban Technol.*, vol. 20, no. 1, pp. 39–55, Jan. 2013.

[79] P. Hinssen, “Open data power smart cities,” 2012.

[80] J. H. Lee, R. Phaal, and S.-H. H. Lee, “An integrated service-device-technology roadmap for smart city development,” *Technol. Forecast. Soc. Change*, vol. 80, no. 2, pp. 286–306, Feb. 2013.

[81] D. Alahakoon and X. Yu, “Advanced analytics for harnessing the power of smart meter big data,” *Proc. - 2013 IEEE Int. Work. Intell. Energy Syst. IWIES 2013*, pp. 40–45, 2013.

[82] W. Yuan, P. Deng, T. Taleb, J. Wan, and C. Bi, “An Unlicensed Taxi Identification Model Based on Big Data Analysis,” *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1703–1713, Jun. 2016.

[83] J. Bai, J. Fan, and R. Tsay, “Special Issue on Big Data,” *J. Bus. Econ. Stat.*, vol. 34, no. 4, pp. 487–488, 2016.

[84] M. V. Moreno *et al.*, “Applicability of Big Data Techniques to Smart Cities Deployments,” *IEEE Trans. Ind. Informatics*, 2017.

[85] B. S. Tice, *Big data*. 2014.

[86] K. E. Skouby and P. Lynggaard, “Smart home and smart city solutions enabled by 5G, IoT, AAI and CoT services,” *Proc. 2014 Int. Conf. Contemp. Comput. Informatics, IC3I 2014*, pp. 874–878, 2014.

[87] Y. Al-Amrani, M. Lazaar, and K. E. Elkadiri, “Sentiment Analysis using supervised classification algorithms,” in *Proceedings of the 2nd international Conference on Big Data, Cloud and Applications - BDCA’17*, 2017, pp. 1–8.

[88] M. Chen, S. Mao, and Y. Liu, “Big data: A survey,” in *Mobile Networks and*

Applications, 2014, vol. 19, no. 2, pp. 171–209.

[89] D. Boyd and K. Crawford, “Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon,” *Inf. Commun. Soc.*, vol. 15, no. 5, pp. 662–679, 2012.

[90] S. Azzi, A. Bouzouane, S. Giroux, C. Dallaire, and B. Bouchard, “Human activity recognition in big data smart home context,” *Proc. - 2014 IEEE Int. Conf. Big Data, IEEE Big Data 2014*, pp. 1–8, 2015.

[91] CiM; IESE Business School, “IESE Cities in Motion Index 2016 Center for Globalization and Strategy Introduction: The Need for a Global Vision,” *IESE Bus. Sch.*, pp. 1–74, 2016.

[92] J. E. R. Pascual Berrone;(Schneider Electric Sustainability and Strategy Chair) and Carl;(Carl Schroeder Chair of Strategic Management), “IESE Cities in Motion Index 2017,” *Univ. Navarra*, no. IESE Business School-IESE Cities in Motion Index (ST-442-E), 2017.

[93] N. Austin, “Name Address of Shortlisted Firms for Smart Cities Proposal Name Address of Shortlisted Firms for Smart Cities Proposal,” no. 1, 2016.

[94] C. F. Calvillo, A. Sánchez-Miralles, and J. Villar, “Energy management and planning in smart cities Energy management and planning in smart cities,” *Renew. Sustain. Energy Rev.*, vol. 55, no. March, pp. 273–287, Mar. 2016.

[95] L. G. Anthopoulos, *Understanding Smart Cities: A Tool for Smart Government or an Industrial Trick?*, vol. 22, no. November. 2017.

[96] V. Nunes, J. P. Gouveia, A. M. Rodrigues, and T. Simão, “INSMART - towards the new distribution systems operators potential roles in low carbon future and integrated frameworks for smart cities,” in *CIRED - Open Access Proceedings Journal*, 2017, vol. 2017, no. 1, pp. 2797–2799.

[97] L. Vidiasova, P. Kachurina, and F. Cronemberger, “Smart Cities Prospects from the Results of the World Practice Expert Benchmarking,” *Procedia Comput. Sci.*, 2017.

[98] L. G. Anthopoulos, “Understanding the Smart City Domain: A Literature Review,”

in *Public Administration and Information Technology*, vol. 8, Cham: Springer International Publishing, 2015, pp. 9–21.

[99] M.-L. L. Marsal-Llacuna, J. Colomer-Llinàs, and J. Meléndez-Frigola, “Lessons in urban monitoring taken from sustainable and livable cities to better address the Smart Cities initiative,” *Technol. Forecast. Soc. Change*, vol. 90, no. PB, pp. 611–622, Jan. 2015.

[100] N. Tsolakis and L. Anthopoulos, “Eco-cities: An integrated system dynamics framework and a concise research taxonomy,” *Sustain. Cities Soc.*, vol. 17, pp. 1–14, Sep. 2015.

[101] P. Bellini, M. Benigni, R. Billero, P. Nesi, and N. Rauch, “Km4City ontology building vs data harvesting and cleaning for smart-city services,” *J. Vis. Lang. Comput.*, vol. 25, no. 6, pp. 827–839, Dec. 2014.

[102] N. F. da Cruz and R. C. Marques, “Scorecards for sustainable local governments,” *Cities*, vol. 39, pp. 165–170, Aug. 2014.

[103] F. Duarte *et al.*, “A Conceptual Framework for Assessing Digital Cities and the Brazilian Index of Digital Cities: Analysis of Curitiba, the First-Ranked City,” *J. Urban Technol.*, vol. 21, no. 3, pp. 37–48, Jul. 2014.

[104] L. G. Anthopoulos, “The rise of the smart city,” in *Public Administration and Information Technology*, vol. 22, 2017, pp. 5–45.

[105] A. Staffans and L. Horelli, “Expanded Urban Planning as a Vehicle for Understanding and Shaping Smart, Liveable Cities,” *J. Community Informatics*, vol. 10, no. 3, p. online, 2014.

[106] ISO, “Sustainable development of communities: Indicators for city services and quality of life,” no. July. p. 112, 2014.

[107] L. G. Anthopoulos, M. Janssen, and V. Weerakkody, “Comparing Smart Cities with different modeling approaches,” *Proc. 24th Int. Conf. World Wide Web - WWW ’15 Companion*, vol. 1997, pp. 525–528, 2015.

[108] S. Moreno Pires, T. Fidélis, and T. B. Ramos, “Measuring and comparing local sustainable development through common indicators: Constraints and

achievements in practice," *Cities*, vol. 39, pp. 1–9, Aug. 2014.

[109] P. Berrone and J. E. Ricart, "IESE Cities in Motion, Methodology and Modeling Index," *Navarra IESE Bus. Sch. Univ. Navarra*, 2014.

[110] K. C. Desouza and T. H. Flanery, "Designing, planning, and managing resilient cities: A conceptual framework," *Cities*, vol. 35, pp. 89–99, Dec. 2013.

[111] S. Singhal, S. McGreal, and J. Berry, "Application of a hierarchical model for city competitiveness in cities of India," *Cities*, vol. 31, pp. 114–122, Apr. 2013.

[112] S. Batty, "Sustainable urban planning: models and institutions," *Implement. Underst. Sustain. Dev.*, vol. II, pp. 71–86, 2003.

[113] G. C. Lazaroiu and M. Roscia, "Definition methodology for the smart cities model," *Energy*, vol. 47, no. 1, pp. 326–332, Nov. 2012.

[114] K. Mori and A. Christodoulou, "Review of sustainability indices and indicators: Towards a new City Sustainability Index (CSI)," *Environmental Impact Assessment Review*, vol. 32, no. 1. Elsevier, pp. 94–106, 01-Jan-2012.

[115] U. C. L. Centre, F. O. R. Advanced, and S. Analysis, "UCL Centre For Advanced Spatial Analysis Papers Smart Cities of the Future," vol. 44, no. 0, p. 40, 2012.

[116] L. Leydesdorff and M. Deakin, "The triple-helix model of smart cities: A neo-evolutionary perspective," *J. Urban Technol.*, vol. 18, no. 2, pp. 53–63, Apr. 2011.

[117] M. Thite, "Smart cities: Implications of urban planning for human resource development," *Hum. Resour. Dev. Int.*, vol. 14, no. 5, pp. 623–631, Nov. 2011.

[118] J. Malek, M. Laroussi, H. B. E. N. Ghezala, and H. Ben Ghezala, "A Design Framework For Smart City Learning Scenarios," *2013 9th Int. Conf. Intell. Environ.*, pp. 9–15, Jul. 2013.

[119] L. Chan *et al.*, "User's manual on the Singapore Index on Cities' Biodiversity," 2014.

[120] G. S. Yovanof and G. N. Hazapis, "An architectural framework and enabling wireless technologies for digital cities & Intelligent urban environments," in *Wireless Personal Communications*, 2009, vol. 49, no. 3, pp. 445–463.

[121] R. G. Hollands, "Will the real smart city please stand up?," *City*, vol. 12, no. 3, pp.

303–320, Dec. 2008.

[122] R. Giffinger, H. Kramar, N. Pichler-Milanović, and F. Strohmayer, “SMART CITY PROFILES Deliverable 2.1 Part 1 Introduction The Smart Cities’ Approach,” no. May, pp. 1–7, 2014.

[123] L. Chan *et al.*, “User’s manual on the Singapore Index on Cities’ Biodiversity,” 2014.

[124] Zurinah Tahir, Jalaluddin Abdul Malek, Z. Tahir, and J. A. Malek, “Main criteria in the development of smart cities: determined using analytical method,” *Plan. Malaysia*, vol. XIV, no. 5, pp. 1–14, Nov. 2016.

[125] V. Kučerová, P. Hanzel, and A. Hajtman, “Peritonzilárny absces u 15-mesačného dietata,” *Otorinolaryngol. a Foniatr.*, vol. 64, no. 3, pp. 177–180, 2015.

[126] MoUD-GoI, “Draft Concept Note on Smart City Scheme,” *Smart Cities*, vol. 2014, pp. 1–46, 2014.

[127] I. J. 1, “Smart cities - Preliminary Report 2014,” *Inf. Technol.*, pp. 1–71, 2015.

[128] I. J. 1, “trolle.2003 - Mammal survey in Rio Negro Basin.pdf,” *Inf. Technol.*, pp. 1–71, 2015.

[129] BSI, “Smart cities-Vocabulary BSI Standards Publication Bio-based products-PAS 600:2013 Part 0: Subtitle BSI Standards Publication,” 2014.

[130] S. E. Bibri and J. Krogstie, “Smart sustainable cities of the future: An extensive interdisciplinary literature review,” *Sustain. Cities Soc.*, 2017.

[131] S. Madakam and R. Ramaswamy, “100 New Smart Cities (India’s Smart Vision),” *2015 5th Natl. Symp. Inf. Technol. Towar. New Smart World, NSITNSW 2015*, no. February, pp. 1–6, Feb. 2015.

[132] S. Tayal and K. Sharma, “The Recommender Systems Model for Smart Cities,” *Int. J. Recent Technol. Eng.*, vol. 8, no. 2S7, pp. 451–456, 2019.

[133] “Spearman Rank Correlation Coefficient BT - The Concise Encyclopedia of Statistics,” in *The Concise Encyclopedia of Statistics*, New York, NY: Springer New York, 2008, pp. 502–505.

[134] J. V. Winters, “Why Are Smart Cities Growing? Who Moves And Who Stays,” *J.*

Reg. Sci., vol. 51, no. 2, pp. 253–270, May 2011.

[135] A. S. Elmaghraby and M. M. Losavio, "Cyber security challenges in smart cities: Safety, security and privacy," *J. Adv. Res.*, vol. 5, no. 4, pp. 491–497, Jul. 2014.

[136] S. F. Naya and K. E. Lee, "The characteristics of Asia-Pacific growth triangles," *J. Asia Pacific Econ.*, vol. 1, no. 1, pp. 123–129, 1996.

[137] P. Chamoso, A. González-Briones, S. Rodríguez, and J. M. Corchado, "Tendencies of Technologies and Platforms in Smart Cities: A State-of-the-Art Review," *Wirel. Commun. Mob. Comput.*, vol. 2018, pp. 1–17, Aug. 2018.

[138] C. Pérez-Galván and I. D. L. Bogle, "Global optimisation for dynamic systems using interval analysis," *Comput. Chem. Eng.*, vol. 107, no. 1, pp. 343–356, Aug. 2017.

[139] IBM, "Smarter cities for smarter growth. How cities can optimize their systems for the talent-based economy," 2010.

[140] R. Carli, M. Dotoli, R. Pellegrino, and L. Ranieri, "Measuring and managing the smartness of cities: A framework for classifying performance indicators," *Proc. - 2013 IEEE Int. Conf. Syst. Man, Cybern. SMC 2013*, pp. 1288–1293, 2013.

[141] S. Tayal, S. K. Goel, and K. Sharma, "A comparative study of various text mining techniques," in *2015 International Conference on Computing for Sustainable Global Development, INDIACom 2015*, 2015, vol. 11, no. 5, pp. 76–81.

[142] W. M. Wey and J. Hsu, "New Urbanism and Smart Growth: Toward achieving a smart National Taipei University District," *Habitat Int.*, vol. 42, pp. 164–174, Apr. 2014.

[143] J. Yang, Y. Liu, X. Zhu, Z. Liu, and X. Zhang, "A new feature selection based on comprehensive measurement both in inter-category and intra-category for text categorization," *Inf. Process. Manag.*, vol. 48, no. 4, pp. 741–754, Jul. 2012.

[144] L. M. Abualigah, A. T. Khader, M. A. Al-Betar, and O. A. Alomari, "Text feature selection with a robust weight scheme and dynamic dimension reduction to text document clustering," *Expert Syst. Appl.*, vol. 84, no. C, pp. 24–36, Oct. 2017.

[145] R. Garg, R. K. Sharma, and K. Sharma, "Ranking and selection of commercial off-

the-shelf using fuzzy distance based approach," *Decis. Sci. Lett.*, vol. 5, no. 2, pp. 201–210, 2016.

[146] S. Tayal, P. P. Singh, N. Nagwal, and K. Sharma, "Development Indicators Analysis using R," 2017.

[147] M. Mafarja and S. Mirjalili, "Whale optimization approaches for wrapper feature selection," *Appl. Soft Comput. J.*, vol. 62, pp. 441–453, Jan. 2018.

[148] S. Okuya, "M2M and big data to realize the smart city," *NEC Tech. J.*, vol. 7, no. 2, pp. 67–71, 2012.

[149] A. Sheth, "Transforming Big Data into Smart Data: Deriving value via harnessing Volume, Variety, and Velocity using semantic techniques and technologies," *2014 IEEE 30th Int. Conf. Data Eng.*, pp. 2–2, 2014.

[150] H. M. Zawbaa, E. Emary, A. E. Hassanien, and B. Parv, "A wrapper approach for feature selection based on swarm optimization algorithm inspired from the behavior of social-spiders," in *Proceedings of the 2015 7th International Conference of Soft Computing and Pattern Recognition, SoCPaR 2015*, 2016, pp. 25–30.

[151] M. A. Tawhid and A. M. Ibrahim, "Feature selection based on rough set approach, wrapper approach, and binary whale optimization algorithm," *Int. J. Mach. Learn. Cybern.*, pp. 1–30, Aug. 2019.

[152] P. Bellini, P. Nesi, and N. Rauch, "Knowledge base construction process for smart-city services," *Proc. IEEE Int. Conf. Eng. Complex Comput. Syst. ICECCS*, pp. 186–189, 2014.

[153] Y. Li, "An efficient mathematical model to process data of smart city," *Proc. - 7th Int. Conf. Intell. Comput. Technol. Autom. ICICTA 2014*, pp. 64–67, 2015.

[154] M. Mafarja and S. Mirjalili, "Whale optimization approaches for wrapper feature selection," *Appl. Soft Comput. J.*, vol. 62, no. November, pp. 441–453, 2018.

[155] W. Pedrycz, "Applied soft computing: Editorial," *Applied Soft Computing Journal*, vol. 27. Elsevier B.V, p. 489, 2015.

[156] Y. Al Amrani, M. Lazaar, and K. E. El Kadirp, "Random forest and support vector

machine based hybrid approach to sentiment analysis," *Procedia Comput. Sci.*, vol. 127, pp. 511–520, Jan. 2018.

[157] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, Aug-2018.

[158] M. Liu, X. Lu, and J. Song, "A New Feature Selection Method for Text Categorization of Customer Reviews," in *Communications in Statistics: Simulation and Computation*, 2016, vol. 45, no. 4, pp. 1397–1409.

[159] P. B. and the C. Consortium, "Citypulse_Annual_Report," 2014.

[160] K. Sharma, R. Garg, C. K. Nagpal, and R. K. Garg, "Selection of optimal software reliability growth models using a distance based approach," *IEEE Trans. Reliab.*, vol. 59, no. 2, pp. 266–276, 2010.

Publications

Papers Accepted/Published in International Journals

1. **Kapil Sharma, Sandeep Tayal**," "Indian Smart City Ranking Model" using Taxicab Distance-Based Approach," International Journal of Energy Systems ISSN: 1868-3975 2019, https://doi.org/10.1007/s12667-019-00365-9, **Springer, (Scopus)**
2. **Kapil Sharma, Sandeep Tayal**," "Indian Smart City Ranking Model," International Journal of Recent Technology and Engineering ISSN: 2277-3878, Volume-8 Issue-2, July 2019, DOI: 10.35940/ijrte.B2472.078219 **(Scopus)**
3. **Kapil Sharma, Sandeep Tayal**," An Adaptive Whale Optimization Algorithm Guided Smart City Big Data Feature Identification for Fair Resource Utilization," International Journal of Innovative Technology and Exploring Engineering ISSN: 2278-3075, Volume-8, Issue-12, October 2019, DOI: 10.35940/ijitee.L3763.1081219 **(Scopus)**

Papers Accepted/Published in International Conferences

1. **Sandeep Tayal, Kapil Sharma**," The Recommender systems model for Smart Cities," International Journal of Recent Technology and Engineering ISSN: 2277-3878, Volume-8, Issue-2S7, July 2019, DOI: 10.35940/ijrte.B1083.0782S719 **(Scopus)**
2. **S. Tayal**, N. Nagwal, and **K. Sharma**, "Role of Big Data in Make in India," in *International Conference on Advanced Computing and Intelligent Engineering (ICACIE 2016) held during 21 - 23 Dec 2016 organized by C. V. Raman College of Engineering,* Bhubaneswar, India., 2016, https://doi.org/10.1007/978-981-10-6875-1_42, **Springer.**
3. **S. Tayal**, S. K. Goel and **K. Sharma**, "A comparative study of various text mining techniques," *Computing for Sustainable Global Development (INDIACom 2015) 2nd International Conference on 11th - 13th March,* New Delhi, 2015, pp. 1637-1642**,** IEEE. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7100525&isnumber=7100186