I J C T A, 10(8), 2017, pp. 705-710 ISSN: 0974-5572

Data mining Applications for Smart city: A Review

Supriya Nagarkar*

ABSTRACT

We are living in information era, where data and information are two most important aspects of day to day life. Most of the business and scientific organizations generates large amount of data. Effective analysis and utilization of gigantic data is a key factor for success in many business and service fields which includes the smart city domain. Many countries are adopting Data mining applications to support smart city components to achieve required goals.

The aim of this paper is to review the applications of data mining to support smart city mission.

Index terms: Applications of Smart City,Applications of Data mining, Data mining, Smart city.

1. INTRODUCTION

Smart city: Concept of Smart city was introduced in 1998, but its realization started in the year 2000. Urban population is constantly increasing which results in to increasing pressure on public utilities like transport, electricity, health care services and Education system. This emphasizes the necessity of Smart City to meet challenges of increasing population. Well defined components of the smart city are Smart mobility, Smart environment,Smart governance, Smart people plus its application areas Smart energy, Smart education and Smart healthcare.This generates huge amount of data. There is a need of good technique to interpret and extract useful information from this data.

Data mining (DM): DM sometimes called as a data or knowledge discovery is a computer assisted process of digging and analyzing enormous set of data and then extracting the meaning of data. It automatically searches large volume of data for models & patterns using computational techniques from statistics and machine learning and information theory. It is the ideal tool for extracting knowledge from hidden data.It can predict future trends & behavior which can be used for decision-making.

2. APPLICATIONS OF DATA MINING

Data mining is an interdisciplinary field; it covers areas like business, banking, science, research, risk analysis etc.

A. *Data mining for financial Data Analysis*–Finance and banking sectors produces large amount of information. In this sector data is always systematic, reliable and complete. In banking system loan payment prediction and credit analysis are important task. Classification and clustering these two techniques are used to identify customer behavior. DM can also be useful in detecting financial crimes.

B. *Data mining for Retails Industry*–Retail Industry is very important application area for data mining. In such application large amount of information such as sales, customers shopping history, goods, transportation history is recorded. Data mining in retail industry can help to identify customers buying behavior which will improve the quality services and customer satisfaction.

* Assistant Professor, TMV, *Email: supriyanagarkar@gmail.com*

C. *Data mining for Telecommunication Industry*–Telecommunication industry provides many services like Emails, Fax, Internet messengers. The information available with telecommunication is vast and multidimensional. So Data mining plays important role in understanding the business involved, identifying pattern, making best utilization of resources detecting the frauds, and to improve the quality services.

D. *Data mining For Medicine and health care data*–Hospital, health care organizations and insurance companies generates large volume of data. Data mining can permit healthcare organizations to forecast trends in the patient conditions and behaviors, which is accomplished by data analysis from different outlooks and learning connections and relations from superficially related information.

E. *Data mining in higher education*–Data mining concepts in Education field is Called EDM-Educational Data Mining. Educational communities have become very proactive now a day. They are constantly thinking about giving quality education to the students so that they can create a brand name. They are also trying to curtail the rate of failures and drops-outs.Predicting the failure of students in advance is a complex task as it needs to process large amount of data. So data Mining techniques can easily use in this context.

3. TECHNIQUES OF DATA MINING

A. *Classification*–This data mining Technique is used to unify items in a set of data into one of a predefined set of classes. Classification technique is based on mathematical techniques such as decision trees, linear programming and statistics. For example Classification can be used to categorize students which can predict the students' performance.

B. *Clustering*–Clustering is a useful techniqueof grouping data into different groups so that data in each group shares similar type of patterns. Clustering techniques defines the classes and puts objects in it. For example clustering technique can be used to keep similar types of books in one shelf so that reader can easily search the book.

C. *Association Rules*–Association is one of the best data mining techniques which are also known as Relation technique. It tries to discover frequent items set from voluminous data. For example customers usually buy butter with bread which intern saves time and increases sales.

D. *Regression& prediction*–As the name indicates this technique is used for prediction. Prediction model tries to discover relationship between independent variables and dependent variables. For example this technique can be used to predict the profit on sales where profit is a dependent variable and sale is independent variable. For this past data can be considered and future prediction can done.

4. SMART CITY–AN EVOLVING CONCEPT

Ever increasing urbanization is a global phenomenon. Urban population is constantly increasing which results in to increasing pressure on public utilities like transport, electricity, health- care services and Education system. This emphasizes the necessity of Smart City to meet challenges of increasing population. Smart city is a way to integrate multiple information and new technologies.

5. APPLICATIONS OF DATA MINING TO SMART CITY – LITERATURE SURVEY

This paper [1] discusses the energy consumption data with the parameters – public building, structure, construction, behavior pattern. This paper compares traditional method and data mining technology for energy consumption analysis. It focuses on use data mining techniques for efficient utilization to obtained knowledge and tries to fill that gap by utilizing data mining in the energy efficiency evaluation of buildings.

Table 1
Applications and services of Smart city

Components	*Areas Covered*
Smart Energy Management	- Smart street lights - Solar energy initiative - Energy efficient Green Building
Urban Mobility	- Automatic Traffic Management - Multi-model transportation services (Road, metro, rail, air, waterways …) - GIS Maps for city - Good connectivity to rural areas for transportation of goods / vegetables & other needs
E-Governance and citizen services	- Disaster Management solution - Online municipal services - Easy clearance of business services
Smart Environment	- Environment friendly Green building - Pollution control system - Early detection for natural disaster
Water & waste Management	- Smart meter system - Water Quality mapping meters - Waste gathering and clearance - Supply clean water for drinking - Water recycling
Smart Education	- Digital education - Smart Integrated system which will bring students, teachers and parents under one umbrella
Smart Health	- Up–to–date health care information system - Smart hospitals with enhanced patience interaction system

This paper[2]talk over various issues related with health care. Data mining in health care will help in organizing large amount of data, use of information technology for automation, maintaining the security, help in early diagnosis, help in predicting the trend and the analysis from various perspective. It further suggests the Classification and Regression, Association rule, Cluster analysis and text mining techniques for analysis of data.

This paper [3] compares the effectiveness and performances of several data mining techniques such as decision tree, regression in predicting irrigation water demand. The study compares SysFor with other classification techniques for the first time with accuracy and efficiency.

This paper [4] Emphasizes on challenges face by educational industry, applications in learning like performance prediction , predicting future failures and finding the solution, development of courseware, assessment and research. They have listed the use of different open source tool like MangoDB, Hadoop, Weka, and Orange.

This paper [5] presents a widespread study of big data driven smart energy management and also proposes theprocess model for its management. This paper also provides insight on Smart energy management in IT sector, data collection, integration processing, sharing and security.

This Paper[6] discusses six objectives of Mobility in context to Smart city. They are Pollution, Traffic blocking, people safety, Noise pollution, Transfer speed and transfer cost. Author has suggested 4 key factors such as Public companies, Private Companies, Local Government agencies and Integration of these threethat is Integrated Transport System (ITS). There are some innovative solutions like – Use of CNG Vehicles, Use ICT for smart phone based Integrated ticketing system, Car sharing , Car reservation etc.in

addition to this author also focuses on ICT based advance Intelligent transport system which includes video surveillance for security , Traffic control and traffic data collection system etc.

This paper [7] presents the concept of E-Governance with the parameters –Smart, open, agile government with participation and collaboration at all levels. It further explain some areas/opportunities/ challenges as Budget(finance), Modernization(Use of ICT – Information and communication Technology),Safety and security (Data Privacy &confidentiality),High speed connectivity(Internet), Participation and collaboration (Active involvement of every citizens and social media), open data (Accurate and reliable data) and open Government (Transparency and accountability). Finally Smart governance model discusses

- Norms – New standard smart budgetary systems for budgeting and controlling need to be developed.
- Policies – Newsustainable and adaptive policies have to be outlined.
- Practices – Interactionof Academicians, Technology and Government will play important role in this context.
- Information – Information must be open, timely, and sharable.
- Technologies – Useof Modern and advance technology is a backbone of this process
- Human Capital – Smart, skill people having the knowledge of Norms, process, and policies are must.

6. BENEFITS OF DATA MINING TO SMART CITY APPLICATION AREAS

Collaboration of Data mining Techniques and Smart city will bring many benefits to all stake-holders. Data mining technique will provide computer based solution in searching hidden patterns from large amount of data and fragment the data into clusters which can give better results.

6.1. SmartEnergy Management

To facilitate decision-making related distribution of electricity, actual demand of the citizens andother affecting conditions.

- It allowsnear-real time determination of demand using efficient analysis of the data collected using data mining techniques.
- It helps to achieve resource optimization using pricing plans consistent with demand, supply and generation methods of electricity.

6.2. Urban Mobility

Analysis of data will help to identify the traffic patterns. This will help for following

- To identify traffic-jams, accidents in the city roads' at various times and to reduce the same by adjusting suitable traffic controls. To send instructions / feedback to appropriate authorities for taking actions to improve or resolve a traffic problem.
- Enable sharing of traffic information (collected through sensors, smart traffic lights and on-vehicle devices to drivers via smartphones or other communication devices) to the commuters
- To support decisions like opening new roads, enhancing the infrastructure basedon congestion data, updation of car parking and alternative roads.
- Reduce supply chain delays by properly managing associated deliveries and shipping movements.

6.3. E-Governance and citizen services

E-governance is the collaboration of different government agencies and streamlining theirprocesses. This results in better & efficient operations, better handling of shared data, stronger enforcement of rules &regulations. The data mining techniques will help

- Improve business decisions regarding new policies or enhancing existing policies.
- Governmentagencies to improve the quality of the data, while citizens will show how they can use the data andtransfer it to new knowledge to enhance the quality of government services.
- Government to focus on issues / concerns related to health & social care, housing, education etc.

6.4. Smart Environment

Accurate and in-time weather information will help

- To improve the country's agriculture,
- To informpeople ofpossible hazardous conditions (like storms, heavy rains …) and better disaster management
- Alerts to health departments and associated agencies regarding contagious diseases and prevent epidemics

6.5. Water & Waste Management

- Better analysis of usage will help to predict demand (and
- To improve waste recycle there by reducing load on natural resource
- Alerts to health departments and associated agencies regarding contagious diseases and prevent epidemics

6.6. Smart Education

- Behavior and matchmaking will lead to new knowledge.
- Every student is unique. Analysis of student's information (assessment of graduates academics, attitudes, behavioral pattern etc.)can help the education institutes
  * To predict failure rate and measures to reduce the same
  * To check whether they are using their resources in the right places and producing the right results.
- Optimize academic research(for instance, astronomer can now analyze a huge astronomy dataset using powerfulcomputers instead of manual analyses. By analyzing and exploring high quality digital images taken from space,new discoveries may happen in the fields.) This is applicable to many fields such as medicalexperiments, manufacturing operations, environmental studies and economic and financial analysis.

6.7. Smart Healthcare

- Allow doctors to gather, analyze and utilize patient information (comprehensive and well maintained patient history records) for better diagnostics & treatment. This information can also beused by insurance companies and government agencies.
- To help predicting health issues and provide associated treatment
- Better management of contagious diseases and prevent epidemics

7. CONCLUSION

Based on the study following conclusions can be Drawn:

Data mining and smart-cities are two important concepts. Ever developing smart cities will be generating huge amount of data which can be characterized as 3V's – Velocity, Volume, Variety. Data mining techniques will help to extract useful information from this large data. The integration of data-mining techniques and smart cities will providevery useful information. This information can drive decisions for sustainable life style, smart governance, smart energy efficient buildings, smart traffic management and quality life.

REFERENCE

[1] Liang Zhao, Jili Zhang, ChongquanZhong,"The Application of Data Mining Technology in Building Energy Consumption Data Analysis," International Journal of Computer, Electrical, Automation, Control and Information Engineering, vol. 10, pp. 81-85, 2016.

[2] Boris Milovic, Milan Milovic, "Prediction and Decision Making in Health Care using Data Mining," International Journal of Public Health Science (IJPHS),vol. 1, pp. 69-78, 2012.

[3] Mahmood A. Khan, MdZahidul Islam, Mohsin Hafeez, "Evaluating the Performance of Several Data Mining Methods for Predicting Irrigation Water Requirement,"Proceedings of the Tenth Australasian Data Mining Conference (AusDM), Sydney, Australiavol. 134, pp. 199-207, 2012.

[4] Katrina Sin and LoganathanMuthu, "Application of Big Data in Education Data Mining and learning Analytics– A literature review," ICTAT Journal on Soft Computing, vol. 05, pp. 1035-1049, 2015.

[5] Kaile Zhou Chao Fu, Shanlin Yang, "Big data driven smart energy management: From big data to big insightsRenewable and Sustainable Energy Reviews," pp. 215-225, 2016.

[6] Clara Benevolo, Renata Paola Dameri and Beatrice D'Auria, "Smart Mobility in Smart city," Springer International Journal, pp. 13-26, 2016.

[7] Hans J. Scholl1 and Margit C. Scholl2,"Smart Governance: A Roadmap for Research and Practice," iConference, pp. 163–176, 2014.