Evolving Sentiments towards E-Governance using Opinion Mining

Pankaj Kumar
Assistant Professor
SRMGPC, Lucknow

Alisha Zaidi
Student - B. Tech
SRMGPC, Lucknow

Ankita Srivastava
Student - B. Tech
SRMGPC, Lucknow

Geet Jain
Student - B. Tech
SRMGPC, Lucknow

ABSTRACT
Today in current scenario, much of the research are focusing on opinion mining and sentiments analysis due to its high volume of opinion available in blogs, discussion forum, review sites, surveys, news, social network sites etc available in electronic format. Now researchers are anticipated to build up a system that can recognize and classify opinion as represented in digital form. We have a chance to make appropriate belief of a thousand views and discover problems as soon as they occur. Opinion mining can be made appropriately as a sub-discipline of computational linguistics that exactly concentrates on extracting people's opinion from the web, surveys and other which are available digitally. Opinion mining gives us direction in building a system by which we can bring together and inspect opinions about the product available digitally such as surveys, blogs and other online format.

Here in this paper we have considered E-Governance basically as the application of Information and Communications Technology to define the functioning of Government in order to bring about 'Simple, Moral, Accountable, Responsive and Transparent' (SMART) governance.

In this paper we are trying to evolve the sentiments towards E-Government projects and policies on the basis of opinion collected from the people. We try to analyze the effectiveness of E-governance using the concept of opinion mining. For this we took the opinions of people regarding smart city.

Keywords
E-governance, opinion mining, sentiment analysis.

1. INTRODUCTION
As citizens of India, we have to deal with government in our day-to-day lives. Citizens expect speedy service, proper treatment, and quick disposal of grievances or applications. This interaction, however, is not always pleasant as for different classes of people the perception and impact of government projects and policies are different. With the rising awareness among citizens about their rights, Government today, is expected to be transparent in its dealings, accountable for its activities and faster in its responses [1, 7]. The emergence of Information and Communications Technology (ICT) has provided means for faster and better communication, efficient storage, retrieval and processing of data and exchange and utilization of information to its users, be they individuals, groups, businesses, organizations or governments [2, 8]. Thus, the use of opinion mining and sentiment analysis tools can help in bridging this gap between the government and the citizens. According to the World Bank "E-Government refers to the use by government agencies of information technologies (such as Wide Area Networks, the Internet, and mobile computing) that have the ability to transform relations with citizens, businesses, and other arms of government. These technologies can serve a variety of different ends: better delivery of government services to citizens, improved interactions with business and industry, citizen empowerment through access to information, or more efficient government management. The resulting benefits can be less corruption, increased transparency, greater convenience, revenue growth, and/ or cost reductions."Thus, the concentration here is on use of information technologies in improving citizen-government connections, cost-cutting and creation of revenue and clearness [4,9].

Opinion mining is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. Sentiment analysis is a type of natural language processing for tracking the mood of the public about a particular product or topic [3, 4, 11]. It is about determining the subjectivity, polarity (positive or negative) and polarity strength (weakly positive, mildly positive, strongly positive, etc.) of a piece of text – Opinion mining differs from pure data and text mining in so far it deals with subjective statement. In this sense, it is a specific development of a discipline dealing with unstructured information extraction (IE) that was previously mainly working with objective data such as natural disasters or bibliographic information [5, 10].

In this paper, our main concern is to find out the sentiments of various classes of people regarding e-governance focusing mainly on the Smart City project. Concept of the smart city has been introduced as a strategic device to encompass modern urban production factors in a common framework and to highlight the growing importance of Information and Communication Technologies (ICTs), social and environmental capital in profiling the competitiveness of cities. The significance of these two assets - social and environmental capital - itself goes a long way to distinguish smart cities from their more technology-laden counterparts, drawing a clear line between them and what goes under the name of either digital or intelligent cities.

2. RELATED WORK
In the paper "e-Governance Project Lifecycle" published by National Institute for Smart Government, www.nisg.org on behalf of the Department of Electronics & Information Technology, Government of India, it is stated that E-Governance is the tool used by the government for discharging its functions for the welfare of the citizens. So it is has to be citizen-friendly. Delivery of services to citizens and promoting their rights are considered primary functions of the government. For doing so government has to change its outlook and its way of interacting with citizens. In a democratic nation of over one billion people like India, e-Governance should enable seamless access to information and

seamless flow of information across the state and central government in the federal set up.

The paper [8] "A Survey of Opinion Mining and Sentiment Analysis" Written by Bing Liu summarises as: With the explosive growth of social media (i.e., reviews, forum discussions, blogs and social networks) on the Web, individuals and organizations are increasingly using public opinions in these media for their decision making. However, finding and monitoring opinion sites on the Web and distilling the information contained in them remains a formidable task because of the proliferation of diverse sites.

"Sentiment Analysis and Opinion Mining: A Survey" by G.Vinodhini discusses the usage of different methods for analyzing the opinions and evolving opinions and sentiments out of extracted views of different people regarding various products, policies etc. The performance of different methods used for opinion mining is evaluated by calculating various metrics like precision, recall and F-measure. Precision is the fraction of retrieved instances that are relevant, while recall is the fraction of relevant instances that are retrieved. In the government context, opinion mining has long been in use as an intelligence tool, to detect hostile or negative communications [4].

3. CONCEPT OF SMART CITY

It is believed that on the back of better connectivity and better access to public information, we can manage cities more effectively, anticipate and solve problems in a more cost effective way, and raise the economic prospects and the quality of life of everyone. In so doing, the India can strengthen its position as a global hub of expertise at a time when cities throughout the world are seeking innovative solutions to the challenges of urbanization.

There are various key aspects to smarter approaches covering E-Governance, which are strongly information driven:

- E-Passports: The Passport Seva Project was launched by the Ministry of External Affairs with the intention of delivering Passport Services to the citizens in a comfortable environment with wider convenience and reliability.
- E-Visa & Immigration: In order to Modernize and upgrade the Immigration services, "Immigration, Visa and Foreigners Registration & Tracking (IVFRT)" has been identified and included as one of the projects to be undertaken by the Ministry of Home Affairs under the National e-Governance Plan.
- E-Reservation: Reservation of various trains' seats can be easily done through IRCTC.
- E-Health: Telehealth and Telecare products and systems, and digital participation services.
- E-Insurance: E-Governance aims at facilitating customer services, automating grievance redressal mechanism and, creating a holistic database of insurance users.
- E-Tax: Various important e-services being offered include facility for downloading of various forms, online submission of applications for PAN and TAN, query-based services for allotment of PAN and TAN, e-filing of Income Tax Returns, e-filing of TDS returns, online payment of Taxes, issue of refunds through Electronic Clearance Scheme (ECS) and Refund Banker, etc.
- Gram Panchayats: The projects aim at improving governance at the grass roots and providing various e-services at the Panchayat level.
- E District: The MMP aims at delivery of high volume, citizen-centric services at the District level such as issue of birth/death certificate, income and caste certificate, old age and widow pension.
- Employment Exchanges: This project aims at providing e-services to employment seekers and employers.

4. PROPOSED WORK AND ITS IMPLEMENTATION

Proposed methodology will include: Question answering, information retrieval (IR), information extraction (IE), topic tracking, summarization, categorization, concept linkage, and information visualization. Fig 1 shows the various phases of opinion mining:

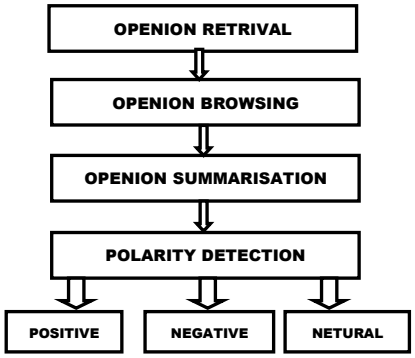

Fig 1: Opinion Mining

We collected the opinions of different people regarding the projects. For convenience we have categorised the people into three domains:

Upper Class: These people are aware of almost all the projects and policies of E-governance. They are mostly active internet users, so E-Governance is very helpful to them. These facilities are very time efficient. Example: E-Visa.

Middle Class: They also know about some of the projects that fall in their category. They find it better and simpler to use. They may avail the facilities of Internet. Example: E-Tax, E-Passport Services etc.

Lower Class: This class of people is mostly deprived of E-Governance Services. This is because neither they know about the projects nor they use Internet facilities.

The opinions by the different classes of people are totally based on the concept of smart city and E- governance. Our main concern is to know:

- Whether they are comfortable in using internet or not.
- Whether they know about some E-Governance Project.
- Whether people know about the concept of smart city
- Whether they are happy with the current scenario of Governance mainly on current city projects.
- If yes, what do they like most.
- If not, what are the desired changes?

- Would they prefer E-Governance or not.

The implementation of the crucial steps techniques are discussed briefly below to enable better understanding of opinion and sentiment analysis:

a) Extraction of information: Here suitable Information extraction algorithms will be used which will identify the relationships between all the acknowledged sequences to provide the user with significant information. This technology can be very helpful when dealing with sheer volumes of information.

b) Categorizing the opinions: Categorization deals with identifying the main idea of a document by placing the document into a pre-defined set of topics. It does not effort to process the actual information but it attempts to process the recognized opinions. It only counts words that appear in the text and, from the counts, identifies the main theme that the document covers. Categorization often depends on a glossary for which topics are predefined, and relationships are identified by looking for wider terms, narrower terms, synonyms, and related terms.

c) Clustering: Clustering is a technique which is used to group documents which are similar in some manner, instead of the use of predefined topics as in opinion categorization. A basic clustering algorithm creates a various topics for each document and measures how well the document fits into each cluster.

d) Tracking of topic: A topic tracking system involves keeping track of user profiles which is based on the documents the user views, predicts other documents of interest to the user. By using suitable text mining tools this system can even automatically infer the user's interests based on his/her reading history and click-through information.

e) Summarization: Text summarization is very crucial for figuring out whether or not a lengthy document meets the user's requirements. Basic use of summarization is to lessen the length and detail of a document while retaining its main points and overall meaning. In opinion summarization only those features of products are mined on which the users state their opinions.

For evolving the better sentiments of people we use the concepts of Naive bayes classifier which involves a simplifying conditional independence assumption. That is given a class (positive or negative); the words are conditionally independent of each other.

In our case, the maximum likelihood probability (P(xi|c)) of a word belonging to a particular class is given by the expression:

$$= \frac{\boldsymbol{Count\ of\ xi\ in\ documents\ of\ class\ c}}{\boldsymbol{Total\ no.\ of\ words\ in\ documents\ of\ class\ c}}$$

The frequency counts of the words are stored in hash tables during the training phase. As the proposed methodology is based on positive and negative aspect, so we calculate the probability of data to be positive or negative. So probability for sentiments of opinion of different class will be given by

$$\boldsymbol{P(S1|S2)} = \frac{\boldsymbol{P(S2|S1)P(S1)}}{\boldsymbol{P(S2)}}$$

Where S1 is the sentiment for any sentence S2. Now using the natural language processing we will find the maximum entropy. The entropy will give the accuracy of average information of any sentiments and will be calculated as:

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{Tp + Tn}}{\boldsymbol{Tp + Tn + Fp + Fn}}$$

Where ***Tp*** is true positive, ***Tn*** is true negative, ***Fp*** is false positive and ***Fn*** is false negative.

5. RESULT

The opinions collected were different for different classes of people (categorised as upper class, middle class, and lower class). We have used the concept of naive bayes classifier to analyze the sentiments. Some people are comfortable in using internet while some are not. Internet is not accessible to all classes of people. The problem associated with the current scenario varies from person to person. There are people who would prefer E- Governance as it is time efficient and easy to use for the people who use internet services. It is the use of a range of modern Information and Communication Technologies such as Internet, LAN, cell phones etc. by Government to improve the efficiency, competence, service delivery and to encourage democracy. The key challenges with electronic governance are not technology or internet issues but organizational issues like redefining policy and actions, legal issues, infrastructure, ability and consciousness, access to right information and tendency to resist the change in work culture.

E-Governance can renovate citizen service, provide access to information to empower citizens, facilitate their participation in government and enhance citizen economic and social opportunities, so that they can make better lives, for themselves and for the next generation.

6. CONCLUSION

Opinion mining is an rising field of data mining used to extract the knowledge from huge volume of online resources like consumer interpretation, feedback and reviews on any product or topic etc. A lot of work has been conducted to mine opinions in form of document, sentence and feature level sentiment analysis. Sentiment analysis helps in determining the attitude of people regarding various contexts, products etc. This paper basically discusses the sentiments of the citizens regarding e-governance projects like smart city undertaken by the Government of India. The E-Governance plans benefit at wider level from villages to cities. We have discussed how opinion mining techniques help in recovery of information and relationships from textual data sources, thereby support policy makers in discovering relations between policies and citizens' opinions uttered in electronic public forums and blogs etc. We present here, methodology of using opinion mining for e-governance decision support. We have tried to find out the reviews of different people regarding the upcoming projects of smart city that is to be undertaken by the government. Our concern was to know whether the E-Governance is acceptable by the mass or not, and tried to know their expectations in this field.

7. REFERENCES

[1] E-Governance Project Life cycle, printed and published by National Institute for Smart Government, www.nisg.org on behalf of the Department of Electronics & Information Technology, Government of India, April 2012.

[2] Dr. A. Padmapriya, E-Governance: A move towards paperless Administration in India, International Journal

of Computer Trends and Technology- volume4Issue3-2013 ISSN: 2231-2803 Page 404.

[3] Bing liu,Lei zhang,Chapter 1 A survey of opinion mining and Sentiment analysis, Nov 4, 2010.

[4] G.Vinodhini and R.M. Chandrasekaran, Sentiment Analysis and Opinion Mining: A Survey, Published in International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 6, June 2012.

[5] David Osimo and Francesco Mureddu, Research Challenge on Opinion Mining and Sentiment Analysis, June 2012.

[6] Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. Learning Word Vectors for Sentiment Analysis. The 49th Annual Meeting of the Association for Computational Linguistics (ACL 2011).

[7] Vishal Gupta, Gurpreet S. Lehal, A Survey of Text Mining Techniques and Applications, Journal of Emerging Technologies in Web Intelligence, Vol. 1, No. 1, August 2009, Academy publisher, pp 60-76.

[8] Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012.

[9] Bo Pang and Lillian Lee, Opinion Mining and Sentiment Analysis, Foundations and Trends in Information Retrieval, Vol. 2, Nos. 1–2 (2008) 1–135.

[10] S Padmaja and Prof. S Sameen Fatima, Opinion Mining and Sentiment Analysis –An Assessment of Peoples' Belief: A Survey, International Journal of Ad hoc, Sensor & Ubiquitous Computing (IJASUC) Vol.4, No.1, February 2013.

[11] Erik Cambria , Björn Schuller, Yunqing Xia, Catherine Havasi, New Avenues in Opinion Mining and Sentiment Analysis, IEEE Computer Society 2013.