

Automated Customer Query Resolver using Data Mining

Ashwini Jadhav
Department of Computer
Engineering
RMD Sinhgad School of
Engineering
Pune, India

Parth Wadekar
Department of Computer
Engineering
RMD Sinhgad School of
Engineering
Pune, India

Bhakti Wani
Department of Computer
Engineering
RMD Sinhgad School of
Engineering
Pune, India

Nikita Joshi
Department of Computer Engineering
RMD Sinhgad School of Engineering
Pune, India

Parth Sagar
Department of Computer Engineering
RMD Sinhgad School of Engineering
Pune, India

ABSTRACT

In an enterprise service centre, the biggest problem is to provide accurate information to customers and solve their queries. In this case, financial as well as human resources are consumed to a greater extent. In order to reduce this problem, there should be an efficient solution. There are some existing technologies which are used by modern enterprise centres. In an enterprise service centre, when the customer places his query, the frequently asked questions are displayed first. If the customer is not satisfied with the solution or if the required content is not available, then the call will be transferred to the enterprise service centre. As the call is placed, the human interaction between the customer and the enterprise service centre will be substantially increased. In this paper, we propose a system which reduces human interaction and provides automation for resolving queries. For this purpose we use the concept of enterprise mobility. Mobility provides exciting opportunities to interact with your customers, partners and suppliers, empower your employees and connect things to your business.

General Terms

Machine learning, business rules, ontology

Keywords

Knowledge management service, semantic web, data mining

1. INTRODUCTION

Undertaking administrations focuses and IT counseling administrations are a developing business in the present quick paced advertise put[11]. They give an essential approach to ventures to between act with their clients. Administration focuses get an extensive number of administration demands from clients and accomplices. IT counseling administrations are likewise popular as companies are under strain to keep up a specialized favorable position in the present hypercompetitive market. Precise and opportune conveyance of related data to aid benefit rejourney counteractive action and determination is basic for giving the most elevated amounts of administration to clients. This data can be overhauled as a curated learning vault what's more, can be utilized to imbue the administration ask for with information on the most proficient method to unravel the issue. What's more, the information can be coded into business decides that can be utilized as a part of the type of computerized occasion preparing to proactively settle or even anticipate issues in other client networks with similar

gadgets/programming pictures, in this manner maintaining a strategic distance from benefit asks for all together.

Data mining is the figuring procedure of finding designs in extensive informational indexes including strategies at the crossing point of machine learning, insights, and database frameworks. It is a basic procedure where wise strategies are connected to extricate information designs. The general objective of the information mining process is to extricate data from an informational collection and change it into a reasonable structure for additionally utilize. Data recovery (IR) is the action of getting data assets pertinent to a data require from a gathering of data assets. Pursuits can be founded on full-content or other substance based ordering.

1.1 Applications

1.1.1 Future Healthcare

Information mining holds incredible potential to enhance wellbeing frameworks. It utilizes information and investigation to recognize best practices that enhance mind and diminish costs. Mining can be utilized to foresee the volume of patients in each classification.

1.1.2 Market Basket Analysis

Market basket examination is a demonstrating system in view of a hypothesis that on the off chance that you purchase a specific gathering of things you will probably purchase another gathering of things. This method may enable the retailer to comprehend the buy conduct of a purchaser.

1.1.3 Search Engines

Information mining can be utilized as a part of the inquiry engines. User enters the watchword in the web index and every one of the information in the database coordinating with the catchphrase is come back to the client as an answer.

2. DATA MINING

Data Mining is the figuring procedure of finding designs in extensive information sets including techniques at the crossing point of machine learning, insights, and database frameworks. The information mining process separates into five stages. To start with, associations gather information and load it into their information distribution centers. They store and oversee the information, either on in-house servers or the cloud. At that point, application programming sorts the information in light of the client's outcomes, and at last, the end client displays the information in a simple to-share design. Information mining enables showcasing organizations to construct models in light of verifiable information to enhance nature of administration.

2.1 Enterprise Ontology

Enterprise Ontology is about the need to create authoritative models on an abnormal state of reflection, keeping in mind the end goal to have the capacity to create compelling and proficient, purported, between and intra-endeavor data frameworks. These models should be with the goal that it is comprehended both by agents, who are characterizing their usefulness, and programming engineers, who are building and executing the product frameworks that understand the framework's[5][6] usefulness. The possibility of business parts for demonstrating data frameworks is extremely important since they straightforwardly mirror the business decides and the requirements that apply to the endeavor space.

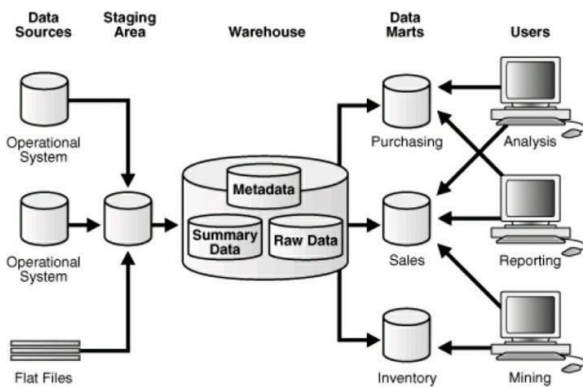


Fig. 1: Simple architecture of Data Mining

2.2 Web Mining

Web mining is the utilization of information mining procedures to find designs from the World Wide Web. As the name proposes, this is data accumulated by mining the web. It makes usage of mechanized devices to uncover and remove information from servers and web2 reports, and it licenses associations to get to both composed and unstructured data from program exercises, server logs, site and connection structure, page content and distinctive sources. Web mining can be separated into three unique sorts – Web use mining, Web content mining and Web structure mining.

In client relationship administration (CRM), Web mining is the joining of data assembled by conventional information mining strategies and procedures with data accumulated over the World Wide Web. (Mining implies separating something helpful or important from a baser substance, for example, mining gold from the earth.) Web mining is utilized to comprehend client conduct, assess the adequacy of a specific Web webpage, and help measure the accomplishment of a promoting effort.

2.2.1 Classes of Web mining

Web content mining — This is the way toward mining valuable data from the substance of Web pages and Web archives, which are generally content, pictures and sound/video records.

Web structure mining — This is the way toward breaking down the hubs and association structure of a site using chart hypothesis.

Web utilization mining — This is the way toward extricating examples and data from server logs to pick up knowledge on client movement including where the clients are from, what number of clicked what thing on the webpage and the sorts of exercises being done on the website.

2.3 Intellectual Capital Mining:

The proposed classifier can group data to a organized configuration that can be effectively comprehended and ingested by benefit focus specialists or clients. The organized data is called Intellectual Capital[11], or IC for short. IC can be utilized as a part of the type of business decides that can be utilized by a generation lead framework to encourage induction what's more, reuse.

The IC mining model offers better arrangement of administration ask for determination information alongside progressed determination and coordinating techniques. It incorporates rich semantics, propelled seek with information mining and machine learning advancements.

2.4 Threats of Data Mining:

2.4.1 Privacy issues

The Threats about the individual security have been expanding hugely as of late particularly when the web is blasting with interpersonal organizations, internet business, discussions, website. In light of protection issues, individuals fear their own data is gathered and utilized as a part of an exploitative way that conceivably causing them a ton of inconveniences.

2.4.2 Security issues

Security is a major issue. Organizations claim data about their workers and clients including government managed savings number, birthday, finance and so forth. Businesses own information about their employees and customers including social security number, birthday, payroll and etc. However how properly this information is taken care is still in questions.

2.4.3 Misuse of information/inaccurate information

Data is gathered through information digging proposed for the moral purposes can be abused. This data might be abused by exploitative individuals or organizations to take advantages of defenseless individuals or oppress a gathering of individuals.

2.5 Evaluation

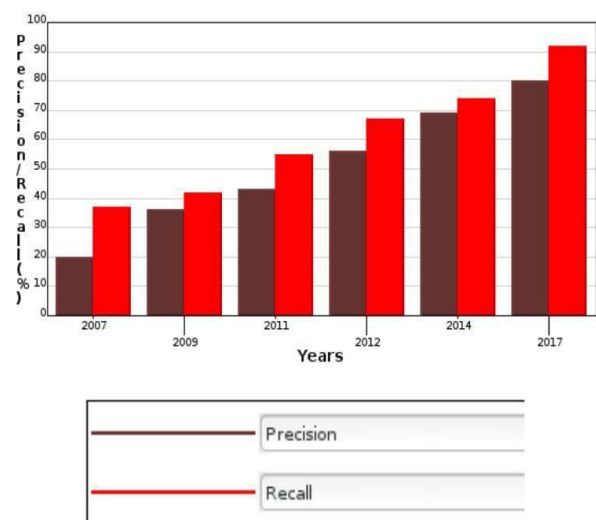


Fig. 2: Comparative study of accuracy achieved in Data Mining

3. LITERATURE SURVEY

[1] Semantics-enhanced Online Intellectual Capital Mining Service for Enterprise Customer Centers (Juan Li, Nazia Zaman Ammar Rayes Ernesto Custodio-2017): This paper propose an effective knowledge mining solution to improve the quality of service request resolution in enterprise service center.

[2] Call-type classification and unsupervised training for the call center domain (Tang, Min Bryan Pellom Kadri Hacioglu-2015): This paper describes recent call type classification and acoustic mod-eling

for speech recognition in the call center domain.

[3] A semantic-based text classification system(Bawakid, Abdullah Mourad Oussalah-2015): This paper proposes a system that automates semantic based text classification. A series of induced methods were implemented that extract semantic features from text and use them to decide how similar a document is to different topics.

[4] Exploratory undersampling for class-imbalance learning (Liu, Xu-Ying Jianxin Wu Zhi-Hua Zhou-2013): The main deficiency of undersampling is that many major class examples are ignored. To overcome this deficiency mainly two algorithms are proposed: Easy Ensemble and balance cascade.

[5] Knowledge Management in Call Centres(Rasooli, Pooya Amir Albadvi- 2012): This paper focuses on five main roles of knowledge management namely: knowledge acquisition, utilization adaptation, dissemination and generation while knowledge management is achieved by identifying and managing these roles efficiently in an organization.

[6] A Framework for Ontology Based Knowledge Management (Abburu, Sunitha, G. Suresh Babu-2012): This paper proposes a framework focusing knowledge reuse. This paper presents the method of knowledge capture, and introduces the mechanism of knowledge push and knowledge pull and also provides a practical example of KMSRP.

[7] Ontology-based semantic online classification of documents: Supporting users in searching the web(Luca, De, Ernesto William, Andreas Nrnberger, O.von- Guericke 2011): This paper describe first results of research on the categorization of texts in order to disambiguate user queries. The discussed methods are based on a combination of indexing and ontology-based information retrieval techniques in an interactive retrieval system.

[8] Design and implementation of domain ontology based oilfield non-metallic pipe information retrieval system.(Rong,Guo,Wu Jun-2010): In order to solve the problems of poor retrieval performance and low retrieval quality in traditional retrieval method, a non-metallic pipe retrieval system based on ontology was designed. Compared with the traditional retrieval method, the domain ontology based non-metallic pipe information retrieval system can significantly improve the recall and precision rate.

[9] Semantics enhanced privacy recommendation for social networking sites(Li, Qingrui, Juan Li, Hui Wang, Ashok Ginja- 2009): An intelligent semantics based privacy configuration system which is named SPAC is designed, to automatically recommend privacy settings for Social Networking Sites users. SPAC learns users privacy configuration patterns and make predictions by utilizing machine learning techniques on users profiles and privacy setting history.

[10] Extended approximate string matching algorithms to detect name aliases (shaikh Muniba, Nasrullah Memon, Uffe Kock Wiil- 2007): This paper mainly focuses on the problem of alias detection based on orthographic variations of Arabic names. Alias detection is the process to identify different variants of the same name. To detect aliases based on orthographic variations the approximate string matching (ASM) algorithms are widely used that measure the similarities between two strings.

[11] Knowledge Extraction and Reuse within Smart Service Centers(Wang, Chunye, Ram Akella, Srikant Ramachandran, David Hinnant-2007): This paper describes the initial version

of a text analytics system under development and use at Cisco, where the objective is to optimize the productivity and effectiveness of the service center.

4. EXISTING SYSTEM

These days current administration focuses have been working on enhancing effectiveness by building knowledge base arrangements. Information administration pointedly diminishes the requirement for acceleration inside and past an administration focus. Frequently an administration ask for being asked to an administration focus, has normally been asked some time recently, and probably will be asked once more. Along these lines, most administration focuses attempt to catch answers to already postured demands and manufacture organized information from this experience. After accepting an administration ask, the framework will coordinate the administration ask for with comparable cases which have been settled some time recently.

This sort of learning contributed by gifted designers and based upon genuine encounter, can be introduced as a learning store or imbued into the genuine administration ask for speedier access and to encourage proficient administration ask reaction. To address the previously mentioned issues, we propose an online learning mining framework, which can help clients find the most forward and pertinent data related to benefit solicitations or client engagement, regardless of whether the clients' solicitations are new to the framework. To get breakthrough data identified with specific points, we swing to the wealthiest sources on the planet– the Internet and the venture's intranet. We actualize a semantics- extended web crawler, which can seek data in light of the semantics instead of sentence structure. To expel the huge measure of clamor come back from the web search tool and shape the data into a capable portrayal, we propose and actualize a semantics- improved multi-level[2] arrangement instrument. The proposed classifier can arrange data to a organized configuration that can be effortlessly comprehended and ingested by benefit focus architects or clients. The organized data is called Intellectual Capital, or IC for short. IC can be utilized as a part of the type of business decides that can be utilized by a creation manage framework to encourage induction what's more, reuse. The proposed IC mining[11] model offers better order of administration ask for determination information alongside moved forward particular and coordinating systems. The proposed work incorporates rich semantics, propelled look with information mining and machine learning innovations. The objective of this work is to understand a usable, savvy, and successful structure for IC mining. Specifically, the commitments of this paper are outlined as takes after: is to understand a usable, savvy, and successful structure for IC mining. Specifically, the commitments of this paper are outlined as takes after:

There is an online pursuit and order model to mine IC. This approach defeats the current issues of information disclosure in benefit focuses, in particular frosty begin, i.e., unfit to unravel the never-seen issues, and hard to incorporate forward new data. Calculations to use the endeavor's philosophy to manage inquiry and information examination prompting better execution.

5. CONCLUSION

To help undertaking client focuses resolve benefit ask for and speed up the time to determine cases utilizing on the web information ,we propose a productive learning administration module to change the heap of information into reusable knowledge or Intellectual Capital(IC). The framework adequately enhances the administration engineers execution and expands the measure of reusable learning.

Thus,we propose a framework which diminishes human collaboration[1] and gives computerization for settling queries. For this reason, we utilize the idea of venture portability which gives

energizing chances to associate with your clients, accomplices and providers, enable your workers and connect things to your business.

6. REFERENCES

- [1] Tang, Min, Bryan Pellom, and Kadri Hacioglu. "Call-type classification and unsupervised training for the call center domain." *Automatic Speech Recognition and Understanding*, 2003. ASRU'03. 2003 IEEE Workshop on. IEEE.
- [2] Bawakid, Abdullah, and Mourad Oussalah. "A semantic-based text classification system." *Cybernetic Intelligent Systems (CIS)*, 2010 IEEE 9th International Conference on. IEEE.
- [3] Liu, Xu-Ying, Jianxin Wu, and Zhi-Hua Zhou. "Exploratory undersampling for class-imbalance learning." *Systems, Man, and Cybernetics, Part B: Cybernetics*, IEEE Transactions on 39.2 (2009): 539-550.
- [4] Rasooli, Pooya, and Amir Albadvi. "Knowledge Management in Call Centres." *Electronic Journal of Knowledge Management* 5, no. 3 (2007):323-332.
- [5] Abburu, Sunitha, and G. Suresh Babu. "A Framework for Ontology Based Knowledge Management." *International Journal of Soft Computing and Engineering (IJSCE)* Volume-3, Issue-3, 2013.
- [6] Luca, De, Ernesto William, Andreas Nu"mberger, and O. von-Guericke. "Ontology-based semantic online classification of documents:Supporting users in searching the web." *Proc. of the European Symposium on Intelligent Technologies (EUNITE 2004)*, Aachen. 2004.
- [7] Fern'andez, Alberto, Salvador Garc'ia, Mar'ia Jose' del Jesus, and Francisco Herrera. "A study of the behaviour of linguistic fuzzy rule based classification systems in the framework of imbalanced data-sets." *Fuzzy Sets and Systems* 159, no. 18 (2008): 2378-2398.
- [8] Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. "SMOTE: synthetic minority oversampling technique." *Journal of artificial intelligence research* 16, no. 1 (2002): 321-357.
- [9] Liu, Xu-Ying, Jianxin Wu, and Zhi-Hua Zhou. "Exploratory undersampling for class-imbalance learning." *Systems, Man, and Cybernetics, Part B: Cybernetics*, IEEE Transactions on 39.2 (2009): 539-550.
- [10] Banerjee, Arindam. "An Analysis of Logistic Models: Exponential Family Connections and Online Performance." *SDM*. 200
- [11] Juan Li, Nazia Zaman, Ammar Rayes, Ernesto Custodio"Semantics-enhanced Online Intellectual Capital Mining Service for Enterprise Customer Centers "2015