Australian Journal of Basic and Applied Sciences, 7(7): 28-34, 2013
ISSN 1991-8178

A Semantic Analysis Based Climate Prediction Using Soa

[1]K.R. SEKAR, [2]K. ARCHANA, [3]K.S. RAVICHANDRAN, [4]J.SETHURAMAN

School Of Computing, SASTRA University, Tirumalaisamudram ,Thanjavur, TamilNadu, India – 613402.

Abstract: Prediction is the one of the factors which is always sought after in all the fields. A negative prediction might not yield the right result. This study is made to increase the probability to the maximum extent in prediction with the use of methodologies like Bayes' theorem and tree-induction. Acquiring a set of records from the host databases in websites is the major work required to be done for a valuable prediction. The training set is to be studied in detail and is packed together for predicting values and the same is stored in the data warehouse for predicting future patterns. This paper mainly concentrates on predicting the weather report for a given place using semantics and ontology in web structures and thereby developing a service oriented architecture.

Key words: Prediction, patterns, semantics, ontology and service oriented architecture

INTRODUCTION

Prediction techniques are the most desired with the evolution of technologies. The practicability of data mining makes it possible to increase the accuracy in predictions and henceforth is applied for the problem on prediction.

High possibilities for errors can be met during the course of prediction for which we acquire the data from various websites and store in the local database and are called the training set. The data contained in the training set are reliable as they are acquired from the database of reliable weather websites. The abstract values are used to predict a given pattern. Specifically we make use of the Bayes' theorem to increase the accuracy in prediction.

Bayes' theorem is a probability theorem. In Bayesian probability, the degree of belief is linked before and after in proposition with the actual evidence. The probability is then used to predict the result, and in our case, the result is the pattern of weather.

The concentration is given high for predicting the weather condition for a place with given attributes. The weather condition changes from region to region.

A slight change in any of the weather attributes will have a greater impact on the weather pattern. Each attributes that are taken into account have got equal priority.

The attributes are compared with that of the trained and acquired data with which the prediction is made possible with least errors. The domain that is used is the service oriented architecture. SOA is the collection of discrete software modules. Each of the modules performs a unique service. All the services together provide the functionality of larger software to solve complex problems. The SOA is used along with web semantics. This service oriented architecture is built using the ontology in web structure.

Related Work:

A. This paper deals with a testing methodology and in addition provides description about the philosophy and domain-knowledge required for a SOA tester (Srikanth Inaganti *et al*., 2008).

B. This paper has discussions about the design of architecture for establishing a semantic service for the use of a model being managed in distributed system. (Omar-el-Gayar *et al*., 2012)

C. This paper has a study about having a cloud service providing data mining services (Software as a Service (SaaS)) for non-expert data miners. (Omar-el-Gayar *et al*., 2012)

D. This paper is based on the research work on Network Enabled Capability through Innovative Systems Engineering (NECTISE) project to model and simulate SOA for dependable and stable military capability with extension to regional surveillance. (Lu liu, Duncan Russell et.al., 2009).

E. Addressing the existing problem of finding the service providers who minimize the execution cost of the business process with respect to the constraints like execution period and cost, this paper proposes an algorithm to reduce the effort of search. It is done by delivering solution to moderate size problems. (Daniel A. Menascé *et al*., 2009)

F. Quality of Service (QoS) management in SOA is focused in this paper in which Service Providers (SP) delivers a collection of related services to the consumers. QoS broker intermediates between SPs and consumers about QoS negotiations with five important contributors in the process. ((Daniel A. Menascé *et al*., 2007).

Corresponding Author: K.R. Sekar, School Of Computing, SASTRA University, Tirumalaisamudram ,Thanjavur, TamilNadu, India – 613402.

G. The paper aims at delivering a user-friendly, Hardware and platform independent Industrial Automation System utilizing FOSS and COTS hardware. (S. Veera Raghvan *et al*., 2012).

H. This paper aims to bring out the efficiency network services and its usage in the Web 2.0 and SOA applications and bringing them to be the upcoming “disorderly force” of innovation with the extensive utilities provided by SONA (Service-Oriented-Network-Architecture) of CISCO. (Qing Gu *et al*., 2009).

Proposed Methodology:

The class in our study considers the weather condition for a given pattern in a region. The classes dealt with are five in number and are cloudy, rainy, fog, clear and sunny. The code that is used for obtaining the training set is shown below in fig.: 1. the training set is shown in Table: 1.

```
curl_setopt($ch, CURLOPT_RETURNTRANSFER, 1);
$content=curl_exec($ch);
$htmlcontent = htmlspecialchars($content);
curl_close($ch);
$12=strlen("wx-details");
$ps=strpos($htmlcontent,"wx-details");
$ps1=substr($htmlcontent,($ps+$12+490),80);
$ps2= trim($ps1);
$ps3=substr($htmlcontent,($ps+$12+770),20);
preg_match_all('!\d+!', $ps3, $match);
$humd= implode('',$match[0]);
$pre=strpos($htmlcontent,"wx-data");
$pre1=substr($htmlcontent,($pre+$12+490),80);
$prep= trim($pre1);
$pst=strpos($htmlcontent,"Tonight");
$13=strlen("Tonight");
$ps4=substr($htmlcontent,($pst+$13+480),8);
$low = trim($ps4);
$pt=strpos($htmlcontent,"Forecast For Today");
$14=strlen("Forecast For Today");
$ps5=substr($htmlcontent,($pt+$14+2190),56);
preg_match_all('!\d+!', $ps5, $match1);
$hi= implode('',$match1[0]);
$find=strpos($htmlcontent,"wx-phrase");
$15=strlen('"wx-phrase"');
$ps6=substr($htmlcontent,($find+$15+88),10);
mysql_query("UPDATE CHENNAI SET Mintemp='$low',Maxtemp='$hi',Wind='$ps2',Humidity='$humd',
    Precipitation='$prep',Class='$ps6' WHERE LINK LIKE '%$lower%'",$conn);
```

This code extratcts data from reliable websites for creating the training set

Fig. 1: Code to generate training set for each city

Table 1: Training set for Chennai

Mintemp	Maxtemp	Wind	Direction	Humidity	Precipitation	Class
27	31	14	NE	79	0.0	Cloudy
29	32	5	NE	63	0.0	Clear
27	32	4	NE	62	0.0	Clear
26	29	22	ENE	70	0.0	Cloudy
26	30	15	E	78	0.2	Rain
27	29	16	ENE	64	0.0	Cloudy
24	29	13	NE	77	0.6	Rain
25	30	24	NE	69	0.0	Cloudy
30	33	16	SSE	63	0.0	Sunny
24	30	15	ENE	73	0.4	Rain
28	35	10	E	64	0.1	Sunny
25	31	21	SSE	68	0.1	Cloudy
29	37	15	SSE	56	0.0	Sunny
29	38	11	SE	54	0.0	Sunny
30	40	7	SE	44	0.0	Sunny
26	30	15	ENE	74	0.8	Rain
28	30	3	ESE	52	0.0	Clear
29	32	2	SE	58	0.0	Clear
25	28	18	NNE	77	0.2	Rain
25	27	14	NNE	81	0.1	Fog
24	27	16	NNE	80	0.1	Fog
26	28	12	NE	87	0.0	Fog
25	27	10	NW	86	0.1	Fog
25	28	9	ENE	83	0.1	Fog
29	33	7	NE	60	0.0	Clear

The set is provided with 7 fields in which 6 fields are reserved for the attributes and one filed is for the class declaration. For each attribute field, the mean value for the provided attributes is identified and the entire field is divided into two classes for optimized prediction. The following illustrates the process.

The process begins with the following determination

1. The probability for each class to happen is identified

P (class) =sample space of the required class / sample space (1.1)

2. In case of quantified data, the mean value for each field is identified such that the evaluation will optimize the determination.

3. With the above values the occurrence of a particular class for the provided attributes' value is identified. The calculations are shown below.

4. The mean value for the minimum temperature *(Mintemp)* field is evaluated to be 24.7 units. The values are now distinguished as values above mean and values below mean and the same process is applied to the rest of the fields.

5. The next step is to identify the probability of occurrence of a class for the distinguished partitions.

For example,

Probability of the day being sunny for all temperatures below the mean (*mintemp*) and above the mean value is to be identified.

P (*mintemp* < 24.7 / *sunny*) = 0.4

P (*mintemp*>=24.7 / sunny) = 0.6

The same procedure is followed for the rest of the classes which will end up with 5(*classes*) x 2 (*partitions*) x 5(*attributes*) = 50 values

For a given pattern, the attributes are taken as individual requests and will be compared with the attributes of the training set to increase the probability of success in prediction. The values pertaining to a particular class are multiplied and the eventual values of each class are compared and the class with the maximum value is returned to be the condition of weather for the given pattern and for the specified region.

Worked Example:

A. Problem Definition:

Using bayes finding a given city's climatic conditions based on a set of data attributes like Min. Temperature, Max. Temperature,Wind, Precipitation, Humidity from various online web resources.

B. Proof:

Variables used:

Minmean: mean value of minimum temperature in the training set

Maxmean: mean value of maximum temperature in the training set

Hmean : mean value of humidity in the training set

Wmean: mean value of wind in the training set

Pmean: mean value of precipitation in the training set

Min: minimum temperature value given by the user

Max: maximum temperature value given by the user

Win: wind value given by the user

Hum: humidity value given by the user

Prep: precipitation value given by the user

R1: Probability of Rain pattern calculated based on user's input using the minimum temperature values in the trained set

CL1: Probability of Cloudy pattern calculated based on user's input using the minimum temperature values in the trained set

S1: Probability of Sunny pattern calculated based on user's input using the minimum temperature values in the trained set

F1: Probability of Fog pattern calculated based on user's input using the minimum temperature values in the trained set

C1: Probability of Clear pattern calculated based on user's input using the minimum temperature values in the trained set

Similarly *R2,CL2,S2,F2,C2 are probabilities of Rain, Cloudy, Sunny, Fog, Clear patterns calculated based on user's input using maximum temperature values in the trained set.*

R3, CL3, S3, F3, C3 are probabilities of

Rain, Cloudy, Sunny, Fog, Clear patterns calculated based on user's input using wind values in the trained set.

R4, CL4, S4, F4, C4 are probabilities of Rain, Cloudy, Sunny, Fog, Clear patterns calculated based on user's input using humidity values in the trained set.

R5,CL5,S5,F5,C5 are probabilities of Rain, Cloudy, Sunny, Fog, Clear patterns calculated based on user's input using precipitation values in the trained set.

Chennai Weather Table:

Total number of rows in the training set for Chennai : 25
Maxmean : 31.04
Minmean : 26.72
Hmean : 68.88
Wmean : 12.56
Pmean : 0.112

Classes Category:

Class1 : Rain
Class2 : cloudy
Class3 : sunny
Class4 : fog
Class5 : clear

Given Pattern:

Min: 24 Max: 29 Wind: 13 Hum: 77 Prep: 0.6
Min < Minmean

R1=0.133 Cl1=0.069 S1= 0 F1=0.133 C1=0
Max < Maxmean

R2=0.1020 Cl2=0.1020 S2=0 F2=0.1020 C2=0.0163
Wind >= Wmean

R3=0.1219 Cl3=0.1219 S3=0.0399 F3=0.0399 C3=0
Hum >= Hmean

R4=0.133 Cl4=0.069 S4=0 F4=0.133 C4=0
Prec >= Pmean

R5=1 Cl5=0 S5=0 F5=0 C5=0
*Rain ->0.133*0.1020*0.1219*0.133*1 => 0.000219*
*Cloudy->0.069*0.1020*0.1219*0.069*0 =>0*
*Sunny ->0*0*0.0399*0*0 =>0*
*Fog ->0.133*0.1020*0.0399*0.133*0 =>0*
*Clear ->0*0.0163*0*0*0 =>0*
Result: ***Rainy day***

Implementation:

The Bayesian concept has been used to predict the given pattern for any class result. The snapshots below (Fig.:1) explain the implementation of the concept with the values learnt from the training set and stored in the data warehouse.

```
$minc1=$i1+$i2+$i3+$i4+$i5;
$minc2=$j1+$j2+$j3+$j4+$j5;
$r01=($i1/$minc1)*(($i1+$j1)/$c))/((($i1/$minc1)*(($i1+$j1)/$c))+(($minc1-$i1)*(1-(($i1+$j1)/$c))));
$c01=($i2/$minc1)*(($i2+$j2)/$c))/((($i2/$minc1)*(($i2+$j2)/$c))+(($minc1-$i2)*(1-(($i2+$j2)/$c))));
$s01=($i3/$minc1)*(($i3+$j3)/$c))/((($i3/$minc1)*(($i3+$j3)/$c))+(($minc1-$i3)*(1-(($i3+$j3)/$c))));
$f01=($i4/$minc1)*(($i4+$j4)/$c))/((($i4/$minc1)*(($i4+$j4)/$c))+(($minc1-$i4)*(1-(($i4+$j4)/$c))));
$cl01=(($i5/$minc1)*(($i5+$j5)/$c))/((($i5/$minc1)*(($i5+$j5)/$c))+(($minc1-$i5)*(1-(($i5+$j5)/$c))));

$r11=($j1/$minc2)*(($i1+$j1)/$c))/((($j1/$minc2)*(($i1+$j1)/$c))+(($minc2-$j1)*(1-(($i1+$j1)/$c))));
$c11=($j2/$minc2)*(($i2+$j2)/$c))/((($j2/$minc2)*(($i2+$j2)/$c))+(($minc2-$j2)*(1-(($i2+$j2)/$c))));
$s11=($j3/$minc2)*(($i3+$j3)/$c))/((($j3/$minc2)*(($i3+$j3)/$c))+(($minc2-$j3)*(1-(($i3+$j3)/$c))));
$f11=($j4/$minc2)*(($i4+$j4)/$c))/((($j4/$minc2)*(($i4+$j4)/$c))+(($minc2-$j4)*(1-(($i4+$j4)/$c))));
$cl11=(($j5/$minc2)*(($i5+$j5)/$c))/((($j5/$minc2)*(($i5+$j5)/$c))+(($minc2-$j5)*(1-(($i5+$j5)/$c))));
```

Finding the probability foreach class (Rain,Sunny,Cloudy,Fog,Clear) here

Fig. 2: Code that performs Bayesian theorem

Findings And Discussions:

The study as explained above has been incorporated in a single field for prediction and the result being the prediction of weather for a given pattern has been achieved. The implementation is also explained and illustrated in fig.:3, Fig.:4 and Fig.:5.

Fig. 3: first page of the project where the city name is to be selected

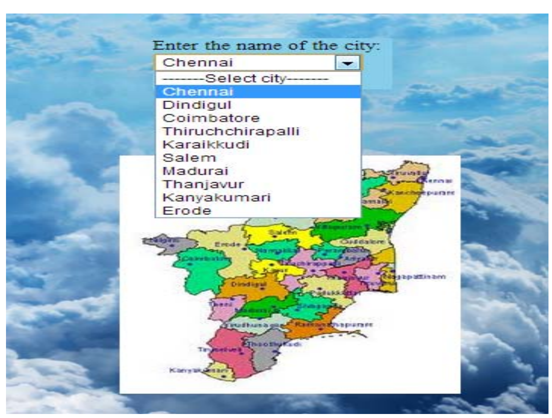

Fig. 4: City name is being selected and "Find" button is clicked

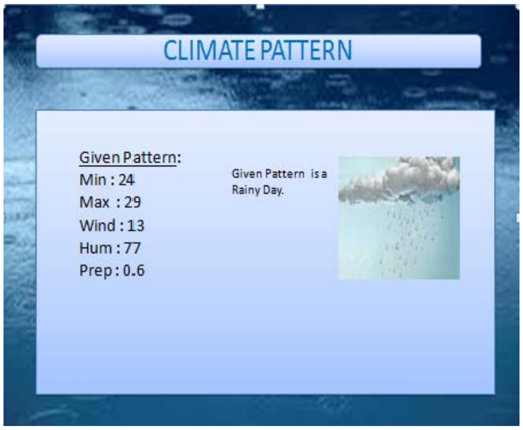

Fig. 5: User enters the input values for minimum, maximum temperatures, wind, humidity and precipitation. Based on the given input, the pattern of the weather is predicted and is displayed.

Conclusion:

This paper has an extensive application of Bayesian concept which can be extended to further more levels. The training set, which is being achieved in due course of the proceedings can be used in the detection of climatic changes over the seasons in any particular region and can also be used to predict the future climatic condition in any region.

ACKNOWLEDGEMENT

We greatly appreciate the following final year students Ms.G.Janani, Ms.B.Ishwarya, Ms.S.Divya of SASTRA UNIVERSITY 2009-13 batch for their contribution and tireless work in implementing this novel idea and bringing out this conception as a fruitful product

REFERENCES

Cisco Service-Oriented Network Architecture: Support and Optimize SOA and Web 2.0 Applications., Cisco white paper

Daniel A. Menascé., Emiliano Casalicchio b and Vinod Dubeyc., 2009. On optimal service selection in Service Oriented Architectures. D.A. Menascé *et al*./ Performance Evaluation., 67(2010): 659-675.

Daniel, A., Menasce., Honglei Ruan and Hassan Gomma, 2007. QOS management in SOA. Performance Evaluation., 64: 646–663.

Daniel, A., Menascé., Emiliano Casalicchio and Vinod Dubey, August 2010. On optimal service selection in SOA. 67(8): 659-675.

Haluk Demirkan., Robert J. Kauffman., Jamshid A. Vayghan., Hans-Georg Fill., Dimitris Karagiannis and Paul P. Maglio, Winter, 2008. Service-oriented technology and management: Perspectives on research and practice for the coming decade. 7(4): 356-376.

Votis, K., C. Alexakos, B. Vassiliadis and S. Likothanassis, 2008. An ontologically principled service-oriented architecture for managing distributed e-government nodes., 31(2): 131-148.

Lu liu and Duncan Russell *et al*., 18 June 2009. Modelling and simulation of Network Enabled Capability on Service-oriented Architecture.

Marcos Palacios., José García-Fanjul and Javier Tuya,2011. Testing in Service Oriented Architectures with dynamic binding: A mapping study., 53(3): 171-189.

Omar-el-Gayar and Amit Deokar, 2012. A Semantic service-oriented architecture for distributed mode management systems, pp: 4-6.

Omar El-Gayar and Amit Deokar, 2012. A semantic service-oriented architecture for distributed model management systems. Article In press.

Qing Gu and H. van Vliet, 2009. SOA decision Making - what do we need to know, in: IEEE Workshop on Sharing and Reusing. Architectural Knowledge.

Rajesh Karunamurthy., Ferhat Khendek and Roch H. Glitho, 2012. A novel architecture for Web service composition., pp: 787-802.

Saulo Barbará de Oliveira., Antonio José Balloni., Felipe Nogueira Barbará de Oliveira and Favio Akiyoshi Toda, 2012. Information and Service-Oriented Architecture & Web Services: Enabling Integration and Organizational Agility, 5: 141-151.

Stein Desmet., Bruno Volckaert and Filip De Turck, 2012. Design of a service oriented architecture for efficient resource allocation in media environments. 28(3): 527-532.

Veera Raghvan., S., Ibrahim Kusumah Kusnantoa and Velappa Ganapathy, 2012. Service Oriented Framework for Industrial Automation Systems, Procedia Engineering, 41: 716-723.

Srikanth Inaganti and Sriram Aravamudan, 2008. SOA Maturity Model. pp: 5-23.

Tristan Glatard., Johan Montagnat., David Emsellem and Diane Lingrand, 2008. A Service-Oriented Architecture enabling dynamic service grouping for optimizing distributed workflow execution. 24(7): 720-730.

Zongsheng Zheng., Dongmei Huang., Jianxin Zhang., Shengqi He and Zhiguo Liu. A SOA-based decision support geographic information system for storm disaster Assessment.