A Semantic Analysis Based Climate Prediction Using Soa

K.R. SEKAR, K. ARCHANA, K.S. RAVICHANDRAN, 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).
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.

```python
import requests
import json

response = requests.get('https://api.openweathermap.org/data/2.5/weather?q=Chennai&appid=YOUR_API_KEY&units=metric')
weather_data = response.json()

temp = weather_data['main']['temp']
humidity = weather_data['main']['humidity']
precipitation = weather_data['rain']['1h']
wind_speed = weather_data['wind']['speed']
wind_direction = weather_data['wind']['deg']

print(f'Temperature: {temp}°C, Humidity: {humidity}%, Precipitation: {precipitation}mm, Wind Speed: {wind_speed}m/s, Wind Direction: {wind_direction}°')
```

![This code extracts data from reliable websites for creating the training set](image)

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

**Table 1:** Training set for Chennai

<table>
<thead>
<tr>
<th>MinTemp</th>
<th>MaxTemp</th>
<th>Wind Direction</th>
<th>Humidity</th>
<th>Precipitation</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>31</td>
<td>NE</td>
<td>79</td>
<td>0.0</td>
<td>Cloudy</td>
</tr>
<tr>
<td>29</td>
<td>32</td>
<td>E</td>
<td>63</td>
<td>0.0</td>
<td>Clear</td>
</tr>
<tr>
<td>27</td>
<td>32</td>
<td>NE</td>
<td>62</td>
<td>0.0</td>
<td>Clear</td>
</tr>
<tr>
<td>26</td>
<td>29</td>
<td>ENE</td>
<td>70</td>
<td>0.0</td>
<td>Cloudy</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
<td>E</td>
<td>78</td>
<td>0.2</td>
<td>Rain</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>E</td>
<td>64</td>
<td>0.0</td>
<td>Cloudy</td>
</tr>
<tr>
<td>24</td>
<td>29</td>
<td>NE</td>
<td>77</td>
<td>0.6</td>
<td>Rain</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
<td>NE</td>
<td>69</td>
<td>0.0</td>
<td>Cloudy</td>
</tr>
<tr>
<td>30</td>
<td>33</td>
<td>SSE</td>
<td>63</td>
<td>0.0</td>
<td>Sunny</td>
</tr>
<tr>
<td>24</td>
<td>30</td>
<td>ENE</td>
<td>73</td>
<td>0.4</td>
<td>Rain</td>
</tr>
<tr>
<td>28</td>
<td>30</td>
<td>E</td>
<td>64</td>
<td>0.1</td>
<td>Sunny</td>
</tr>
<tr>
<td>27</td>
<td>31</td>
<td>SSE</td>
<td>68</td>
<td>0.1</td>
<td>Cloudy</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
<td>SSE</td>
<td>56</td>
<td>0.0</td>
<td>Sunny</td>
</tr>
<tr>
<td>29</td>
<td>33</td>
<td>SE</td>
<td>54</td>
<td>0.0</td>
<td>Sunny</td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td>NE</td>
<td>44</td>
<td>0.0</td>
<td>Windy</td>
</tr>
<tr>
<td>24</td>
<td>36</td>
<td>E</td>
<td>74</td>
<td>0.8</td>
<td>Rain</td>
</tr>
<tr>
<td>28</td>
<td>30</td>
<td>SE</td>
<td>52</td>
<td>0.0</td>
<td>Clear</td>
</tr>
<tr>
<td>23</td>
<td>32</td>
<td>SE</td>
<td>38</td>
<td>0.0</td>
<td>Clear</td>
</tr>
<tr>
<td>25</td>
<td>28</td>
<td>NW</td>
<td>77</td>
<td>0.2</td>
<td>Rain</td>
</tr>
<tr>
<td>27</td>
<td>28</td>
<td>NNE</td>
<td>81</td>
<td>0.1</td>
<td>Fog</td>
</tr>
<tr>
<td>24</td>
<td>27</td>
<td>NNE</td>
<td>80</td>
<td>0.1</td>
<td>Fog</td>
</tr>
<tr>
<td>28</td>
<td>28</td>
<td>NE</td>
<td>87</td>
<td>0.0</td>
<td>Fog</td>
</tr>
<tr>
<td>25</td>
<td>27</td>
<td>NW</td>
<td>85</td>
<td>0.1</td>
<td>Fog</td>
</tr>
<tr>
<td>28</td>
<td>28</td>
<td>ENE</td>
<td>83</td>
<td>0.1</td>
<td>Fog</td>
</tr>
<tr>
<td>29</td>
<td>33</td>
<td>NE</td>
<td>60</td>
<td>0.0</td>
<td>Clear</td>
</tr>
</tbody>
</table>

29
The set is provided with 7 fields in which 6 fields are reserved for the attributes and one field 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

\[ P(\text{class}) = \frac{\text{sample space of the required class}}{\text{sample space}} \]  

(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,
   \[ P(\text{MinTemp} < 24.7 \land \text{sunny}) = 0.4 \]
   \[ P(\text{MinTemp} \geq 24.7 \land \text{sunny}) = 0.6 \]

   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  C1 = 0.069  S1 = 0  F1 = 0.133  C1 = 0
Max < Maxmean
R2 = 0.1020  C2 = 0.1020  S2 = 0  F2 = 0.1020  C2 = 0.0163
Wind >= Wmean
R3 = 0.1219  C3 = 0.1219  S3 = 0.0399  F3 = 0.0399  C3 = 0
Hum >= Hmean
R4 = 0.133  C4 = 0.069  S4 = 0  F4 = 0.133  C4 = 0
Prec >= Pmean
R5 = 0  C5 = 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 * 0.0399 * 0 * 0.069 * 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.
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. 2: Code that performs Bayesian theorem

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