Web based Decision Support System with Data Mining Capabilities to Model the effect of Climate Change on Central Mexico

Alberto Hernández[1], Bruno Lara[1], Alberto Ochoa[2],

[1] Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Av. Universidad No. 1000, Col. Chamilpa.
62209 Cuernavaca, Morelos
{jose_hernandez, bruno.lara}@uaem.mx

[2] Insituto de Tecnología e Ingeniería, Universidad Autónoma de Cd. Juárez, México
alberto.ochoa@uacj.mx

Abstract. We present a prototype Web based Decision Support System with data mining capabilities. The purpose of the presented system is to analyze the effect of Climate Change on Mexico, the HoT-e (How Temperature effects) System was designed using Client-Server Technology. Original code was developed in Java for an intelligent agent to monitor main changes in the temperature of Central Mexico based on information obtained from weather websites on the Internet. We conducted an experiment using our prototype system applied to Cuernavaca, Morelos, a city located in Central Mexico. Preliminary results shows the system could be used to model temperature based on historical information and to identify critical hot days. This system can serve as a base for the development of a prediction model.

Keywords: Web based DSS, Decision Support System, Data Mining Weather, Climate Change, Central Mexico.

1 Introduction

People is increasingly concerned about global climate change and its regional impacts. High latitude areas have seen an increase in temperature in the past decade. In Fig. 1., one can see that Arctic sea ice cover is shrinking rapidly and sea level is rising at an accelerating rate of 3 mm/year. Globally, 8 of the 10 warmest years since 1860, when instrumental records began, occurred in the past decade. A mixture of long-term climate change and natural variability causes these effects. Their impacts are in some cases to result in damage to our environment, for example, increased coastal flooding, severe droughts, more extreme and frequent heat waves and weather events such as severe tropical cyclones [3].

Fig. 1. Arctic sea ice has been shrinking for decades.

Climate change will dramatically increase the number of hot, dry days in Mexico in the coming decades, while coastal regions like the Yucatán peninsula, in the southeast, will be swamped by sea levels half a meter higher than today [5].

According to the Oxford University study financed by the United Nations Development Programme (UNDP) [10], by 2030, Mexico's average daily temperature is likely to climb 1.4 degrees Celsius above what has been the average for the past 30 years. By 2090, this increase could rocket upwards by 4.1 degrees, virtually guaranteeing hot days and nights for 80 to 90 percent of the year.

Cold weather will become very rare in Mexico according to data from the Intergovernmental Panel on Climate Change (IPCC) [2], an umbrella organization of scientists from around the world and the preeminent authority on climate change.

According [6], Mexico is one area of the world where all the computer climate models agree, the highly complex models show Mexico will get much hotter and drier and this will happen sooner in the north and central regions of the country. Climate change has already made the country warmer by about 0.6 C since 1960. The number of hot days has increased by 36 per year between 1960 and 2003.

Morelos is a small State located in Central Mexico (18.45' degrees north, 99.10' degrees west) and in recent years has been experiencing hotter days and random patterns of rainfall which are affecting main economic activities like tourism, production of ornamental plants, sugar and rice production and activities related to aquaculture [9]. Specifically, on this paper we analyze Morelos Capital City, Cuernavaca, a city known worldwide by its "nice" weather and nicknamed as "the ever spring city".

Fig. 2. Morelos State. Capital Cuernavaca, Area: 4,941 km² , Altitude: 1,480m over sea level, Geographic coordinates: 18.45' degrees north, 99.10' degrees west [1].

1.1 Problem at hand

We would like to provide a Web based Decision Support System to model temperatures by means of weather historical information available on the internet and identify seasons of very hot days (up 32°C) to generate alerts for the population. The innovative approach of this research is the exploitation of inexpensive and trustworthy web based weather information by means of intelligent agents and a simplified visual model.

This paper is organized in the following way, Section two discusses the design of the HoT-e (How Temperature effects) System, describing the system architecture, performance and the web based data mining process. We will focus on the flow diagram and the Intelligent Agent's Graphic User Interface implemented on Java NetBeans on charge of retrieving weather information from top weather web sites, and the data mining experiment we carried out. On Section three, we present our preliminary results and discuss them. Finally, we present our future work and conclusions.

1.2 Related work

In Kotsiantis et al. [4] the efficiency of data mining techniques in estimating mean daily temperatures was investigated. They present a hybrid data mining technique to predict mean daily temperature values, results demonstrated it is possible to predict temperature values with satisfying accuracy (95%) using as input the temperatures of previous years.

Recently in [3] a data mining technique for analyzing time series salinity and temperature measurements was proposed, where informative salinity and temperature patterns are extracted. In this research, they designed a transformation method to convert spatial-temporal data to market-basket type data and then applied quantitative inter-transaction association rules mining algorithm to the transformed data set to get salinity and temperature variation patterns. Experimental evaluations show that the proposed algorithm achieves better performance than other inter-transaction association rule mining algorithms.

2 The How Temperature effects (HoT-e) System

Fig 3. shows the block diagram for our prototype system which consists of a three tier Client-Server Architecture based on guidelines from [7], the Client side consists of a Graphical User Interface (GUI) implemented on a Web browser, on this interface the user can introduce the city for which temperature will be modeled; the Server Side where data mining process is going to be executed; and a database where information retrieved from Internet and processed by the server is going to be stored.

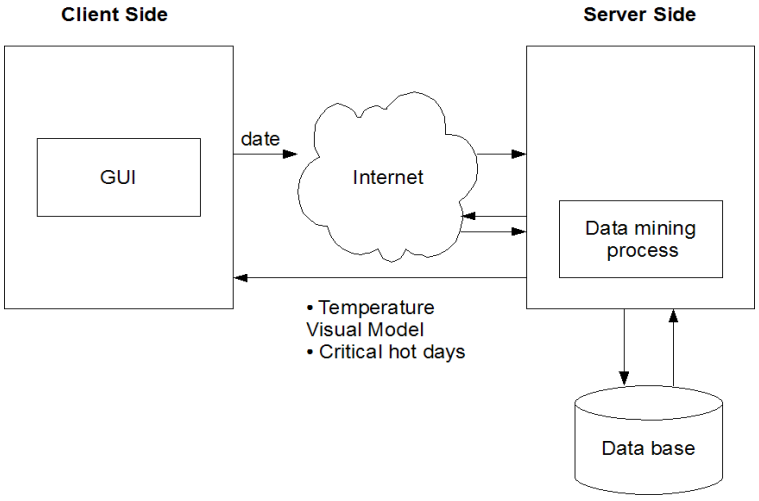

Fig. 3. Client-Server Architecture for the HoT-e System.

2.1 Data mining process

According to [8] knowledge discovering in databases is a process, a set of stages or tasks, among them are: establishing a relevant problem, selection of proper data to problem solution, exploration and cleaning of data, data processing and transformation, application of modeling techniques, generation and interpretation of models proposed, use of obtained knowledge, and generation of new data by means of its application on the real world. The Web based data mining process we used is shown in figure 4.

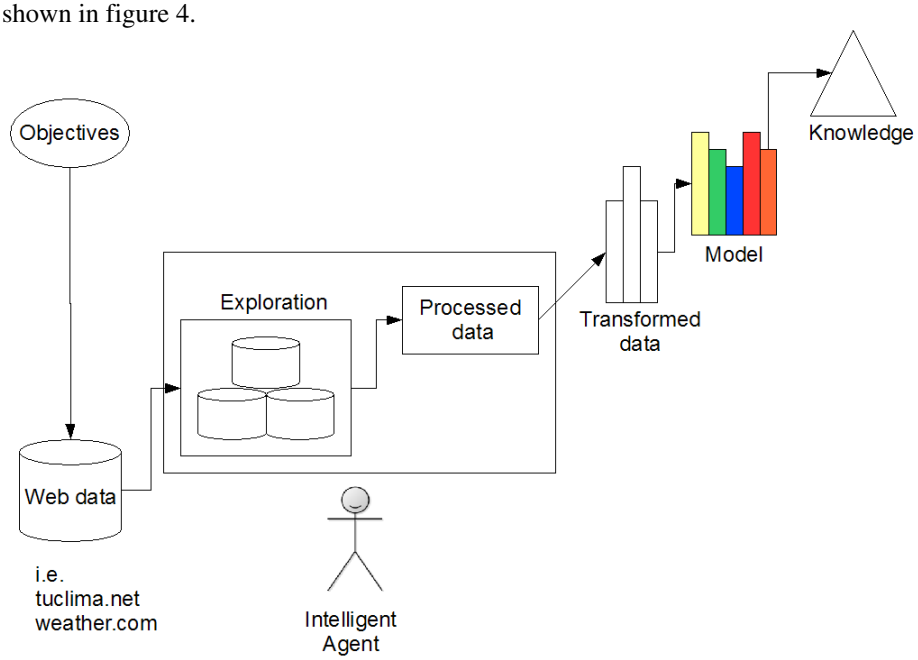

Fig. 4. Web based Data Mining Process.

Considering objectives, web based data is retrieved from top weather web sites by means of an intelligent agent which explores (executes queries on a web site) and process data (transforms meta data on HTML o XML code on valuable weather information), which is stored in a DB for further process. Later, the resulting data is transformed and modeled by means of a visual data mining tool (i.e. rapid miner, enterprise miner or WEKA), based on findings a visual model is generated and proved, finally resulting weather patterns become part of new knowledge.

2.2 Intelligent Agent

An Intelligent Agent is on charge of exploration, retrieving and preprocessing of Web based Information. Fig. 5 shows the complete chart flow indicating each one of the required stages to generate raw data.

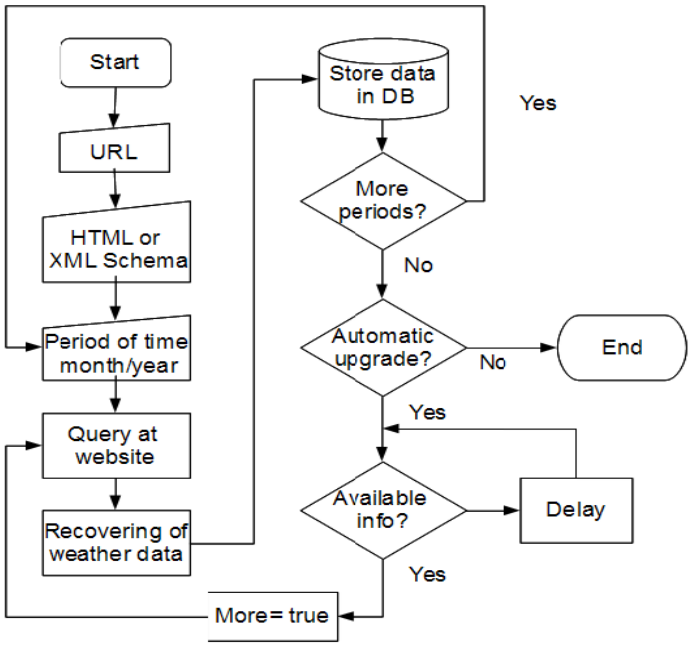

Fig. 5. Web based Data Mining Process Chart Flow.

Graphic User Interface (GUI).

By means of the Intelligent Agent GUI, the analyst or knowledge engineer will introduce the URL for the site where temperatures are contained, the meta data schema describing web page structure, and the range of dates to be retrieved. This request is executed over the Internet. During this process weather information is retrieved from Web sites, stored in a DB, and later, used as a raw material for data mining process. The interface was developed on Java by means of NetBeans 6.7.1.

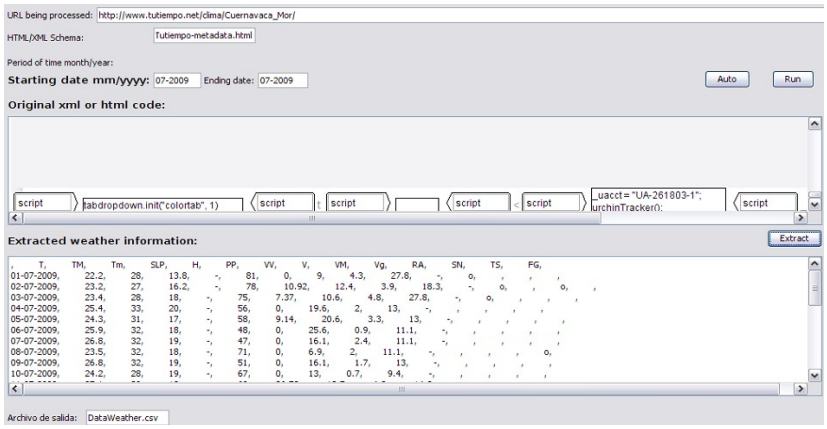

Fig. 6. Intelligent Agent. In the top section the URL, the XML or HTML meta data describing the structure of the information at a web page are introduced. In this same section the user can introduce the starting and ending dates indicating the period of time to be processed. In the middle section, the original HTML code is shown. Finally, the bottom section displays the extracted weather information.

2.3 Experiment

By means of our Intelligent agent, we automatically retrieved weather information for Cuernavaca, Morelos, considering the last decade, from January 2000 to July 2009; 3440 records of information were retrieved including average temperature, maximum temperature, minimum temperature, relative humidity and wind speed. We imported this information to Rapid Miner Software. Finally, the information is analyzed to generate a visual model.

3 Preliminary Results and Discussion

Cuernavaca has experienced a considerable increase of maximum temperatures, 3.8 ° Celsius average in last decade (see Fig. 7a.). Temperatures are dramatically increasing from early February to the third week of May. During this season, in the year 2000 maximum temperatures of 34° Celsius were observed (Fig. 7b) . Meanwhile for the same period of time, in 2008 almost 40° Celsius were recorded (Fig. 7c). The year 2009 reports still higher temperatures (Fig. 7d).

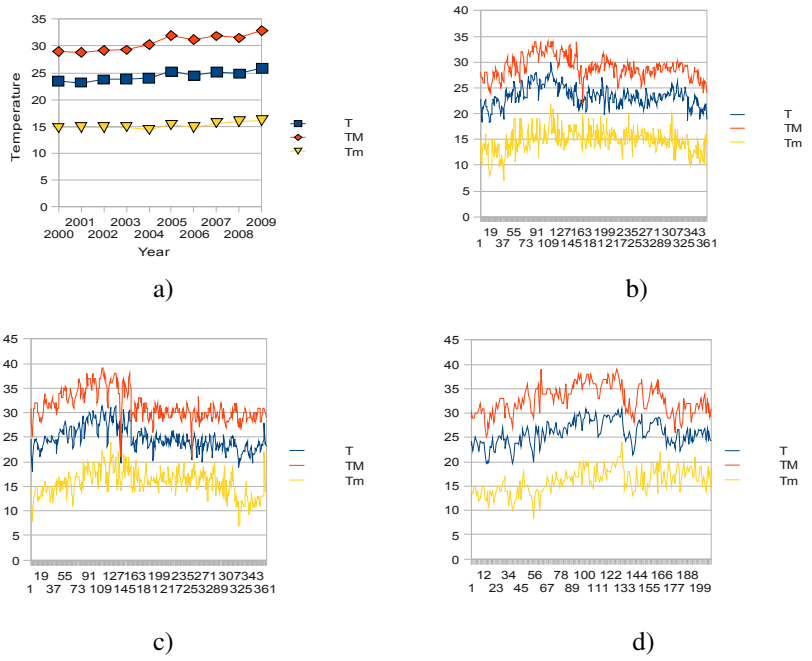

Fig. 7. Preliminary Results. Last decade comparative (a), year 2000 (b), year 2008 (c), year 2009 updated to July (d).

Preliminary results show that by using this visual model it is possible to support basic decisions like: “Use sun blocker on this season”, “Drink a lot of water on this season”, “Caution, season of extremely hot days”. Despite the fact that our system provides an interesting approach to understanding what is happening with climate, available information is not enough to support weather forecasting. In order to reach that goal additional information must be included and modeled providing so a better understanding of the climate change phenomena.

5 Conclusions and Future work

We would like to complete a model for the whole of the municipalities of the state of Morelos. As a second step we would like to consider the whole of Central México including: México City, Querétaro, Estado de México, Puebla and Tlaxcala. Based on our findings, we will like to identify climate patterns to support decisions on agriculture, tourism and water administration. The above described prototype system is a promising technology to support Web based decisions. Our original code developed for the Intelligent Agent can be reused to perform text mining over web pages with varying content. Weather temperature prediction is not a trivial task, a lot of additional work must be done to reach this goal.

Acknowledgments. Special thanks to Abarca, M., and Verduzco, R. for valuable recommendations on art work and constructive feedback on this research paper.

References

1. eMorelos, http://www.e-morelos.gob.mx/
2. Intergovernmental Panel on Climate Change, http://www.ipcc.ch
3. Kao, I.: Predicting Ocean Salinity and Temperature Variations Using Data Mining and Fuzzy Inference. PHD Dissertation, Tatung University, Korea (2008).
4. Kotsiantis, S.; Kostoulas, A., Lykoudis, S., Argiriou, A., Menagias, K. A Hybrid Data Mining Technique for Estimating Mean Daily Temperature Values. In: 64th Meeting of the EURO Working Group on Multicriteria Decision Aiding, September 28–30, Larissa, Greece. (2006).
5. Leathy, S. CLIMATE CHANGE: Hot Days and Nights in Mexico 2090, http://globalintel.net/wp/2008/11/21/climate-change-hot-days-and-nights-in-mexico-2090/
6. McSweeny, C., New, M., Lizcano, G. UNDP Climate Change Country Profiles, Mexico, http://country-profiles.geog.ox.ac.uk
7. Orfalli, R., Harkey, D., Edwards, J. Cliente/Servidor y objetos. Guía de Supervivencia Tercera Edición, Oxford University Press México, S.A. de C.V. México, (2002).
8. Pérez, C., Santín, D. Data Mining. Soluciones con Enterprise Manager. Alfa Omega Grupo Editor, S.A. de C.V., Mexico (2006)
9. Trujillo, T, Taboada, M.: Distribución y Duración de la Sequía Intraestival en el Estado de Morelos. Programa Nacional de Cambio Climático, http://www.atmosfera.unam.mx/cambio/pgmor.htm
10. UNDP Climate Change Country Profiles, http://country-profiles.geog.ox.ac.uk/