Climate Modeling System with Adaptation of Neural Network and AI Data Mining Techniques

Ahmed Mateen	Salman Afsar	Amir Waheed	Zulfiqar Ali
Department of Computer Science University of Agriculture Faisalabad, Pakistan	Department of Computer Science, University of Agriculture Faisalabad, Pakistan	Department of Computer Science, University of Agriculture Faisalabad, Pakistan	Department of Computer Science, University of Agriculture Faisalabad, Pakistan

ABSTRACT
The purpose of this study is to develop a climate modeling system by using data mining techniques which are the process of extracting needed information's from the large database. Thus, the fetching information can be used into practical knowledge for future prediction in climate scenario. It is a powerful and new technology which helps to analyze the hidden predictive information and define the rules for different group of peoples which are working like flood management, hurricane experts, scientists, farmers, weather belonging, social networking etc. to properly manage their needs according to spatial data analysis and plan their coming goals accordingly. In this paper, data mining procedures are used with generalized Neural Network technique which is useful for weather forecasting quickly with the help of data clustering and screening. Giving that investigative instrument to view and utilize this information for decision making processes by taking examples from real life. It is difficult to manage and handle huge data manually. This study gives us number of facilities for inserting, deleting, editing and saving data. This research improves the system performance and information search services which can enhance the quality by regularities in the behavior analysis with respect to time and seasonal data management with Artificial Intelligence with machine learning and pattern analysis.

Keywords
Data Mining, Forecasting, Prediction, Neural Network, Artificial Intelligence, Machine learning

1. INTRODUCTION
The use of computer has happened to much admired and it has took activist alterations even that is not only in the pasture of knowledge although it is in every subject of individual doings. In Science and Technology Pakistan is developing. In the running age of life it is the age of information technology. For the reason is that of information technology, here, in 21st century there are lots of revolutions, by which modern man have distorted the existence style. The use of Computer and information technology is cooperative in process of continuing mills and also for manufacture of parts of machines. Shorten delivery time improve quality, easy way of communication and generating report. The computer based system will be ideal for solving complicated problems [1].

Propose Patterns, that give us conceptual solutions to generally frequent design troubles, widely used in the field of software and hardware have been. Non-functional requirements in the design of safety-critical embedded systems as a key aspect, it is an existing design pattern concept is focused on the integration of passive effects. Perceived safety requirements, reliability, modifiability, cost, and includes the execution time [2].

Some time ago, teaching the use of technology in the information stream, and part was relatively easy. Technology has become an integral part of business processes, so to develop as a more integrated world, the information transfer process has become more complex. Now, the challenges educational institutions face explosive growth of data and maintenance data to improve the quality of decisions to use. Data mining techniques important knowledge from large data sets can be used to carry out the investigative tools [3].

Wind farm forecasting model are five different data mining algorithms. Accuracy is to develop structural analysis. Therefore, by creating a neural network model for both short and long-term forecasts outperforms the other models. Two basic methods are presented forecast: Weather forecast electricity are generated prediction data is generated directly under the direct forecasting models, and wind speed The weather forecast data which is generated with the integrated forecasting model predicts, and then the power of the wind speed. Directly integrated forecasting model forecast model is better than the forecast performance [4].

1.1 Objectives
Data mining procedures are used with generalized Neural Network technique which is useful for weather forecasting quickly with the help of data clustering and screening which is helpful for machine learning and narrate rules for climate predications.

1.2 Ease of Use
Templates and text styles to format your paper is used. All margins, column widths, line spaces, and text fonts are fixed; please do not change them. You can note characteristics. For example, it is customary template measures the margin ratio. This measure and other prospects as a part of the action that can evaluate this topic using the specifications, and as an independent document.

1.3 Developed System
Weather forecast for the developed system uses data mining techniques. The first part of the system based on past experiences in training for the future trend of search data to find trends that have a data mining techniques. For the purpose of this message interacts with the system via an interface. The system that the future trend of mining inputs used to search for factors that desired values are in the form obtained.

1.4 Scope

- Searching particular information takes minor time.

- Decision making is good due to efficient and fast reporting.

- This computerized system made possible to meet the challenges of current time.

- Time to time need of up gradation in the award list submitting system is easily possible now.

- It is possible to reduce the chance of errors using this system.

- The developed system gives cost effective and time saving solutions.

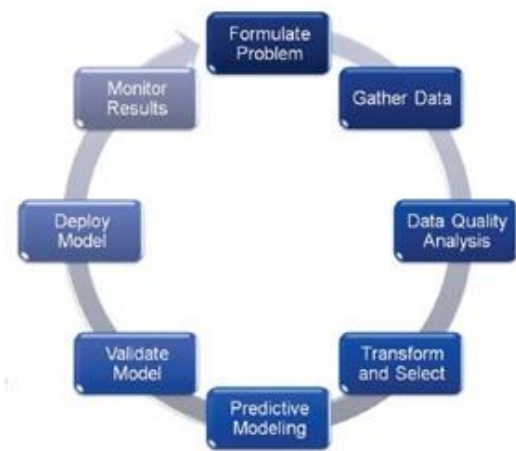- Security of data is possible in the automated system.
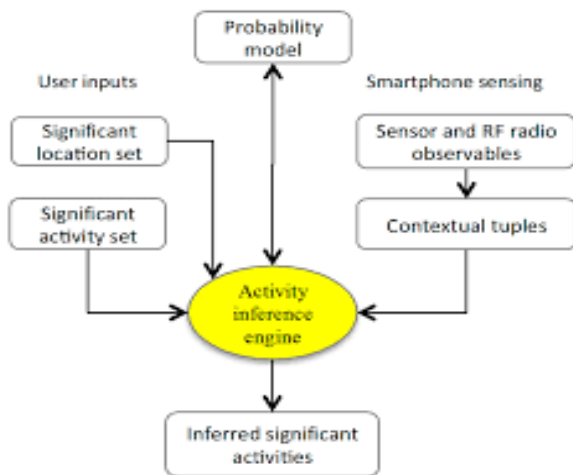

Fig. 1: Data mining process

Fig. 2: Probability model

2. REVIEW AND LITERATURE

Relationship and connection based measures (e.g., the coefficient of determination) have been generally used to assess the "decency of-fit" of hydrologic and hydro climatic models. These measures are oversensitive to compelling qualities (anomalies) and are unfeeling to added substance and corresponding contrasts between model forecasts and perceptions. Due to these confinements, relationship based measures can demonstrate that a model is a decent indicator, notwithstanding when it is definitely not. In this paper, helpful option integrity of-fit or relative mistake measures (counting the coefficient of productivity and the file of assertion) that conquer a significant number of the confinements of connection based measures are talked about. Changes to these insights to help in translation are exhibited. It is reasoned that connection and relationship based measures ought not be utilized to survey the decency of-attack of a hydrologic or hydro climatic model and that extra assessment measures, (for example, rundown insights and outright mistake measures) ought to supplement model assessment devices.

Field estimations in the preservation sciences still make sure of on the portrayal and cloth scrape cushion for data aggregation. Though effective, this procedure is work raised and defenseless against recording and dereferencing botches in the midst of translation. Late advances in convenient enlisting and remote correspondences allow the geoscientist to plan and transmit data while still in the field, in this way minimizing human bungles and time delays. To depict a consolidated system made for biological and geo location data getting that is relied upon to streamline the social event strategy.

The system involves programming applications and hardware parts that engage remote, versatile and Internet figuring in the midst of field campaigns. So accumulated data is refined between the field site and a remote range, a thought suggested as field data spouting. A model system has been attempted in field trials in capability and precision achieved with the use of the field experiments.

Potential applications join concentrated field testing. Field estimations in the regular sciences still depend on the pencil and paper scratch pad for data gathering. But capable, this methodology is worked concentrated and unprotected to recording and geo referencing botches in the midst of translation. Late advances in flexible enlisting and remote trades allow the geoscientist to handle and transmit data while still in the field, thusly minimizing human goofs and time delays.

To portray a consolidated structure delivered for environmental and geo area data acquisition that is proposed to streamline the social affair process. The structure includes programming applications and gear fragments that engage remote, flexible besides, figuring in the midst of field campaigns. So data is proficient between the planed field area, a thought insinuated as field data spouting. A model system has been attempted in field trials for investigation.

Before full-scale field sending, the consolidated data gathering system was pilot-attempted in an investigation of the researchers. They analyzing bunches attempted distinctive system sections numbering the mobile phone programming, geo positional sensor, water quality test and remote data transmission. A meandering van furnished with the field convenient workstation, remote outside switch, amplifier and radio wire served as the central data server. This trial study verified that biological field data could be successfully assembled, secured and transmitted utilizing the organized structure.

3. EXPERIMENT DESIGN

Following step by step procedures is implemented for Machine learning AI system.

- Data Collection

- Validation and Verification

- Transformation

- Analysis(Generalized Algorithm)

- Output(Future Forecasting)

With the help of multi-tier architecture all these steps processed which is very helpful in machine learning and detail model result is suggest/proposed for assembling data and software for parallel view and reordering the data for optimized result in future prospective.

Below figure explain the complete process of multitier architecture [5] of ware hose
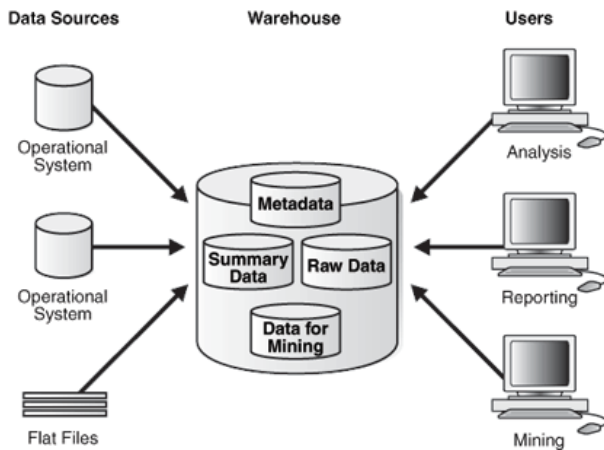
Fig.3: Architecture of data warehouse

Given below statically analysis are applied [6] on the sample data set.

- **Regression Structure**
A Simple Regression [7] equation in the form Y=A+BX (1)

- **Multiple Linear Regression**
In multiple regressions [8] there are more than two variables among one is dependent variable and all others are independent variables and equation look like this:

$Yi = B0+B1Xi1+B2Xi2+B3Xi3--------BnXin$ (2)

The linear correlation coefficient sometimes called Pearson's correlation coefficient. The mathematical formulae for r [9] is given as

$$ r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\left(\sum x^2\right) - \left(\sum x\right)^2}\sqrt{n\left(\sum y^2\right) - \left(\sum y\right)^2}} $$

- **Correlation Coefficient:** This measures the statistical correlation between the predicted and actual values. A higher number means a better model.

- **Mean Squared Error:** Mean-squared error is one of the most commonly used measures of success for numeric prediction [10]. This value is computed by taking the average of the squared differences between each computed value and its corresponding correct value.

- **The Mean-squared Error** is simply the square root of the mean-squared-error [10]. The mean-squared error gives the error value the same dimensionality as the actual and predicted values.

Following are the attributes [10] of the data set.

Table 1. Parameter of data

Attribute	Type	Description
Year	Numerical	Year considered
Month	Numerical	Month considered
Min Temp	Numerical	Minimum Temperature
Max Temp	Numerical	Maximum Temperature
Rainfall	Numerical	Total monthly rainfall

In the multitier architecture all the stored procedures are hidden in business logic with generalized Neural Network Algorithmic approach and verify and validate output analysis in this regard. Front end view for user is available to manipulate the options and choices regarding data set, so from the data base desired values are manipulate and show the results to the users after filtering the result with data Mining AI procedure satisfactory result are shown which will show that climate prediction using this approach is effective and accurate to help the scientist in research prospective for security reason lock the dataset and record and create users with assigning roles in the reporting environment.

Table 2. Average values Analysis

	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	SDV
Average maximum temperature	19.28	22.19	27.03	33.85	39	40.7	37.61	36.42	35.98	33.2	27.41	21.65	7.3825
Average minimum temperature	5.62	7.99	12.79	18.51	23.68	27.54	27.59	26.99	24.43	18.22	11.31	6.76	8.4621
Average rainfall	11.2	22.4	24.6	24.8	13.6	38.9	112.5	105.3	40	4.4	3.4	6.5	37.073

Above table explain the analysis measure [11] of temperature in 1975-2004 and its SDV and the below is its graphical representation of data.
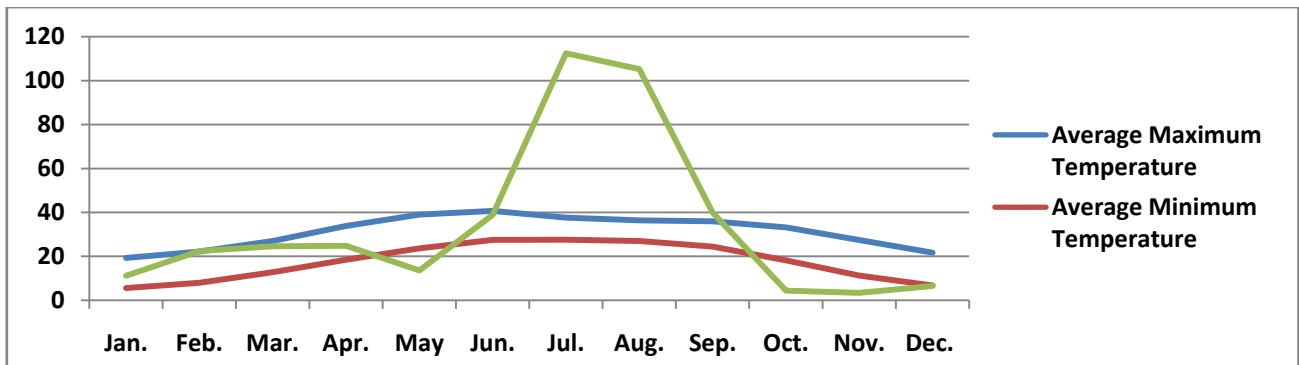
Fig 4: Graph shows that average Max, Min temperature and average rainfall in different months

4. CONCLUSION

In this work the data set of 1975-2004 was used with integration of statical analysis and .net frame work all these are assembled with software and hardware infrastructure with stored and predicted AI& ANN data set when the user select its option from GUI interface these procedure are automatically invoke and climate forecasting is display with 90% of accuracy and auto-reordering of temporary preference is also embed in data set next time corresponding search applied, so predication is show is parallel grid view initially invoke the auto created feature so time and memory is properly managed. beside this the system is also provide the user option like change the parameter so the process component and memory elements is properly utilized with hidden layers which are placed in business logic of multi-tier architectural approach. In future it can be enhance with cloud clustering and web semantic with K-map and Decision tree so the query processing will be further optimized and future trends will be introduced.

5. 6. REFERENCES

[1] Camilli, R., and Vivoni, E. R., 2003. Real-time streaming of environmental field data. Journal of Computers and Geosciences, 29(4): 457-468.

[2] Armoush, A., Salewski, F., and Kowalewski, S., 2009. Design Pattern representation for Safety-Critical Embedded Systems. Journal of Software Engineering & Applications, 2(1): 1-12.

[3] Halder, R., Pal, S., and Cortesi, A., 2010. Watermarking Techniques for Relational Databases. Journal of Universal Computer Science, 16(21): 3164-3190.

[4] Kusiak, A., Zheng H., and Song, Z., 2009. Wind farm power prediction: a data-mining approach. Journal of Wind Energy, 12(3): 275-293.

[5] https://docs.oracle.com/cd/B10500_01/server.920/a96520/concept.htm

[6] Sreehari1, E., Velmuruganv, J., and Venkatesan, M., 2016. A survey paper on climate changes prediction using data mining.

[7] http://www.oxfordjornals.org/our_journals/tropej/online/ma_chap2.pdf

[8] Sethi, N., and Garg, K., 2014. Exploiting data mining technique for rainfall prediction. International Journal of Computer Science and Information Technologies. Vol. 5 (3) , 3982-3984

[9] Olaiya, F., and Adeyemo, A. B., 2012. Application of data mining techniques in weather prediction and climate change studies. International Journal of Information Engineering and Electronic Business, 1: 51-59 MECS (http://www.mecs-press.org/) DOI: 10.5815/ijieeb.2012.01.07.

[10] Kumar, V., and Chadha, A., 2011. An Empirical Study of the Applications of Data Mining Techniques in Higher Education. International Journal of Advanced Computer Science and Applications, 2(3): 80-84.

[11] Muhammad,A,C., Muhammad,F., Rashid,A and Hassan M., 2006. Climatic trends in Faisalabad (Pakistan) over the last 60 years. Journal of Agricultural Sociology and Sciences, 2(1): 42-45.