GROUND WATER LEVEL PREDICTION USING HYBRID RANDOM FOREST AND DCNN

[1] D. SUMA, [2] R. LAXMITHRISHA, [3] P SHIRISHA, [4] M. RAKESH, [5] M. RAMESH

[2,3,4,5] U.G. Scholor, Department of DS, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.

[1]Assistant Professor, Department of DS, Sri Indu College Of Engineering & Technology, Ibrahimpatnam, Hyderabad.

Abstract— **In recent years, the growth of the economy has led to the increasing exploitation of water resources and groundwater. Due to heavy abstraction of groundwater its importance increases, with the requirements at present as well as in future. Accurate estimates of groundwater level have a valuable effect in improving decision support systems of groundwater resourcesexploitation.Thispaperinvestigatestheability ofahybridmodelofartificialneuralnetwork(ANN)and geneticalgorithm(GA)inpredictinggroundwaterlevels in an observation well from Udupi district. The ground waterlevelforaperiodoftenyearsandrainfalldatafor the same period is used to train the model. A standard feed forward network is utilized for performing the prediction task. A groundwater level forecasting model isdevelopedusingartificialneuralnetwork.TheGenetic Algorithmisusedtodeterminetheoptimizedweightsfor ANN. This study indicates that the ANN-GA model can be used successfully to predict groundwater levels of observation well. In addition, a comparative study indicates that the ANN-GA hybrid model performs better than the traditional ANN back-propagation approach...**

Index Terms—**Groundwater Level Prediction, Artificial Neural Network (ANN), Genetic Algorithm (GA), Hybrid Modeling, Water Resources Management, Hydrology, Time Series Forecasting, Optimization Techniques**

I. INTRODUCTION

Groundwaterisoneofthemajorsourcesofsupplyfor domestic, industrial and agricultural purposes. Estimation of groundwater level is very important in hydrogeology studies and aquifer management. In many cases, groundwater level fluctuations have resulted in damage to engineering structures. With considerable amounts of these fluctuations, appropriate decisions can be presented in terms of hydrogeology,waterqualityandits management.For this, a constant monitoring of the groundwater levels isextremelyimportant.Thewaterlevels,ifforecast wellinadvance,helpsadministratorstobetterplanthe groundwater utilization. A continuous forecast of groundwater levels is required to effective use of any simulation model for water management and overall development.Inthisregard,itisimportanttodevelop afastandcost-effectivemethodforaquifersimulation withanacceptableaccuracy.Towardsthisgoal,many researchers have used intelligent systems including, Coulibaly et al., Daliakopouloset al., Lallahem et al., Dogan et al., Nourani et al, Yang et al., Sreekanth et al. [5,8,6,9,10,11,2 These researchers used ANN for aquifer modelling in a variety of basins. ANN is an information-processing paradigm, that is inspired by thewaybiologicalnervoussystems,suchasthebrain, processes information. It determines the relationship between inputs and outputs of physical systems by a network of interconnecting nodes adjusted by connectingweightsbasedonthetrainingsamples,and extracts patterns and detects trends that are too complex to be noticed by either humans or other computational techniques. Neural networks take a different approach to problem solving than that of conventional computers. It has remarkable ability to learn and derive meanings from complicated and imprecise data.Ithas an ability tolearnandapply the knowledge based on the data given for training or initial experience

II. LITERATURESURVEY

2.1INTRODUCTION:

1. Title:"Estimationofgroundwaterlevel usinga hybrid genetic algorithm Neural network" Reference: https://jpoll.ut.ac.ir/article_52176_2d68fe031d5a16632 14492a904b999f5.pdf

Year:2015

Authors:HosseiniZ.andNakhaeiM.,

Description : In this paper, we present an application ofevolvedneuralnetworksusingarealcodedgenetic algorithm for simulations of monthly groundwater levels in a coastal aquifer located in the Shabestar Plain, Iran. After initializing the model with groundwater elevations observed at a given time, the developedhybridgeneticalgorithm-backpropagation (GA-BP) should be able to reproduce groundwater level variations using the external input variables, including rainfall, average discharge, temperature, evaporation and annual time series. To achieve this purpose, the hybrid GA-BP algorithm is first calibrated on a training dataset to perform monthly predictions of future groundwater levels using past observed groundwater levels and additional inputs. Simulations are then produced on another data set by iteratively feeding back the predicted groundwater levels, along with real external data. This modelling algorithmhasbeencomparedwiththeindividualback propagation model (ANN-BP), which demonstrates the capability of the hybrid GA-BP model. The later provides better results in estimation of groundwater levels compared to the individual one. The study suggests that such a network can be used as a viable alternative to physical-based models in order to simulate the responses of the aquifer under plausible future scenarios, or to reconstruct long periods of missing observations provided past data for the influencing variables is available.

III. SYSTEMANALYSIS

EXISTINGSYSTEM:

- In the existing system, we use FNN with the gradient descent method, its algorithms, easily become stuck in local minimum and often need a longer training time.
- Thestochasticoptimizationmethod(GA)totraina FNN; therefore, numerical weights of neuron connections and biases represent the solution components of the optimization problem.
- Thestochasticoptimizationmethod(GA)totraina FNN; therefore, numerical weights of neuron connections and biases represent the solution components of the optimization problem.

DISADVANTAGESOFEXISTINGSYSTEM:

- The existing system model with Logistic regression fails to predict a continuous outcome.
- The existing system model with Logistic regression maynotbe accurateif thesample size is too small.
- The existing system may lead to overfitting problem.
- The existing system accuracy depends on the quality of the data.

PROPOSEDSYSTEM:

- In this project we are using Crow Search with Genetic Algorithm and Grey Wolf with Genetic Algorithm to optimize ground water level features.
- This optimized features will be input to ANN (artificial neural networks) algorithm to train Ground water level prediction.
- ANNtrainedmodelappliedontestdatatopredict water level and then calculate MSE between predicted and test data. MSE refers to difference between actual test data values and predicted values so the lower the MSE.

3.2.1ADVANTAGESOFPROPOSEDSYSTEM:

- The proposed system reduces overfitting in decision trees and helps to improve the accuracy.
- It is flexible to both classification and regression problems.
- Itworkswellwithbothcategoricalandcontinuous values.
- It automates missing values present in the data. ϖNormalizingofdataisnotrequiredasitusesarule-based approach.
- Decreased the total number of verification steps and measures

SYSTEMREQUIREMENTS

HARDWAREREQUIREMENTS(minimum):

- System :Pentiumi3Processor
- HardDisk :500GB
- Ram :4GB.

SOFTWAREREQUIREMENTS:

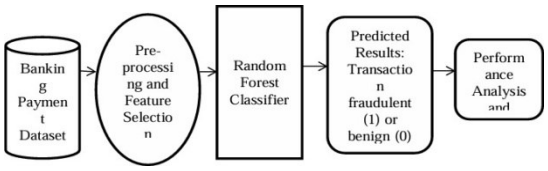
Operating system	:	WindowsXP/7.
Coding Language	:	Python
WebFramework	:	Flask

MODULES:

- Tensorflow
- NumPy

- pandas
- Matplotlib
- Scikit-learn

IV. SYSTEMARCHITECTURE

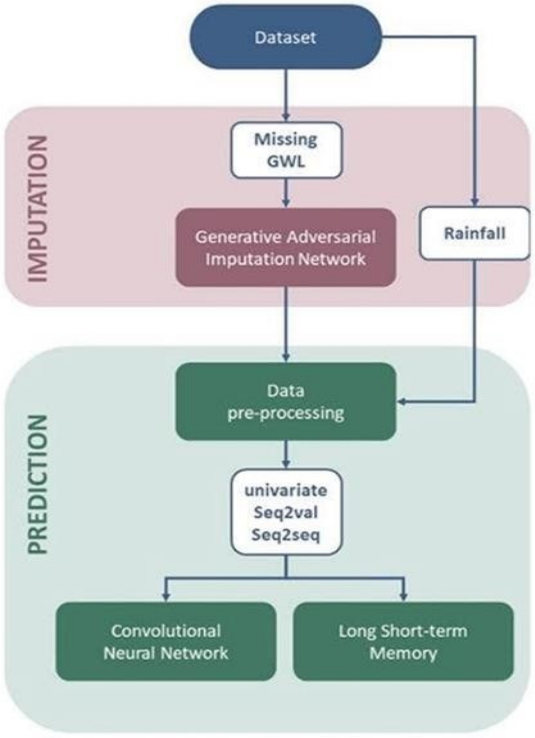


V. SYSTEMDESIGN:DATAFLOWDIAGRAM

DATAFLOWDIAGRAM:

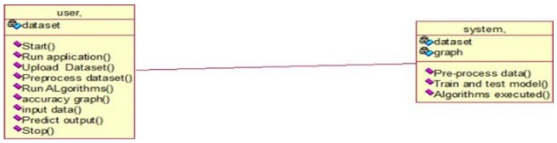
DATAFLOWDIAGRAM:

1. The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.
2. The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components.



CLASSDIAGRAM:

Class diagrams area unit the foremost common diagrams employed in UML. Category diagram consists of categories, interfaces, associations and collaboration. Category diagrams primarily represent the thing directed read of a system that is static in nature.Activecategoryisemployedinaverycategory diagram to represent the concurrency of the system.



VI. SOFTWAREENVIRONMENT

WhatisPython:-BelowaresomefactsaboutPython.
Python is currently the most widely used multi-purpose, high-level programming language. Python allows programming in Object-Oriented and Proceduralparadigms.Pythonprogramsgenerallyare smaller than other programming languages like Java. Programmers have to type relatively less and indentation requirement of the language, makes them readableallthetime.Pythonlanguageisbeingusedby almost all tech-giant companies like – Google, Amazon,Facebook,Instagram,Dropbox,Uber…etc.

AdvantagesofPython:-

Let us see how Python dominates over other languages.

1. ExtensiveLibraries

Python downloads with an extensive library and it contain code for various purposes like regular expressions, documentation-generation, unit-testing, web browsers, threading, databases, CGI, email, image manipulation, and more. So, we don't have to write the complete code for that manually.

2. Extensible

As we have seen earlier, Python can be extended to other languages. You can write some of your code in languages like C++ or C. This comes in handy, especially in projects.

3. Embeddable

Complimentarytoextensibility,Pythonisembeddable aswell. You can putyour Python codein your source

codeofadifferentlanguage,likeC++.Thisletsusadd scriptingcapabilitiesto ourcode inthe otherlanguage.

VII. SYSTEMIMPLEMENTATION

Samplecode:

```
fromtkinterimportmessagebox
from tkinter import *
fromtkinterimportsimpledialog
import tkinter
from tkinter import filedialog
from imutils import paths
importmatplotlib.pyplotasplt
import numpy as np
fromtkinter.filedialogimportaskopenfilename import
numpy as np
importpandasaspd
fromsklearn.model_selectionimporttrain_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn import linear_model
fromsklearn.metricsimportaccuracy_score
fromsklearn.model_selectionimporttrain_test_split
from keras.models import Sequential
fromkeras.layersimportDense,Dropout,Activation
from keras import optimizers
fromgenetic_selectionimportGeneticSelectionCV
import webbrowser
fromsklearn.metricsimportmean_squared_error main
= tkinter.Tk()
main.title("GroundwaterLevelPredictionUsing Hybrid
Artificial Neural Network with
Genetic Algorithm")
main.geometry("1300x1200")
global filename, dataset
globalX,Y,X_train,X_test,y_train,y_test,Y1
global mse, text, pathlabel
```

VIII SYSTEMTESTING

SYSTEMTESTING

Thepurposeoftestingistodiscovererrors.Testingis the process of trying to discover every conceivable faultorweaknessinaworkproduct.Itprovidesaway to check the functionality of components, sub-assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail inanunacceptablemanner.Therearevarioustypesof tests. Each test type addresses a specific testing requirement.

TYPESOFTESTS

Unittesting

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. Thisisastructuraltesting,thatreliesonknowledgeof its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly definedinputsandexpectedresults.Integrationtesting

Integration tests are designed to test integrated softwarecomponentstodetermineiftheyactuallyrun as one program. Testing is event driven and is more concernedwiththebasicoutcomeofscreensorfields.

Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testingisspecifically aimed at exposingtheproblems that arise from the combination of components.

Functionaltest

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

IX SCREENSHOTS

INDEX PAGE:

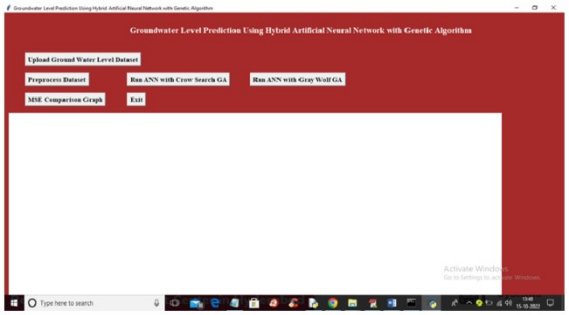


Fig 9.1: Index Page

Uploading Data Sets:

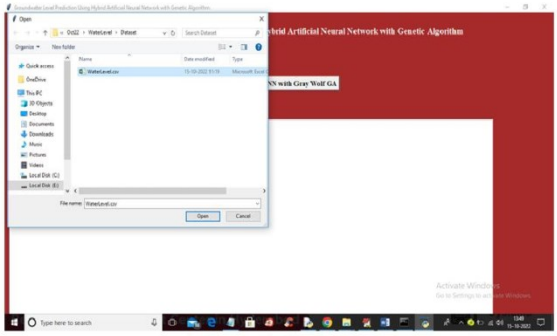


Fig 9.2: Uploading Data Sets

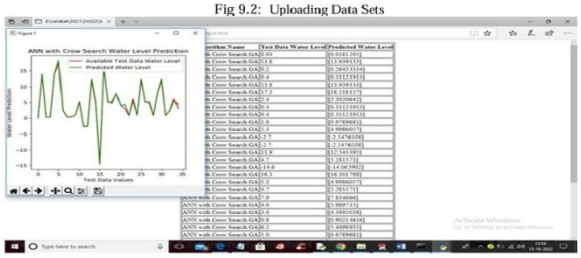


CHART:

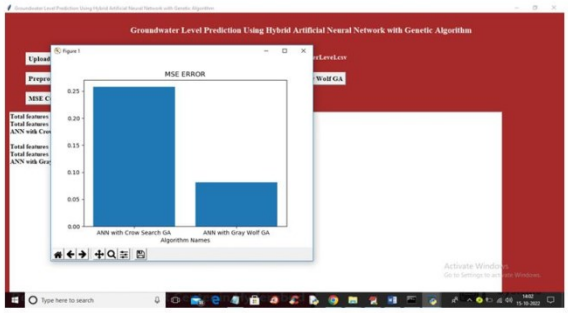


Fig 9.8: Chart

X.CONCLUSION

In this paper, two approaches of soft computing have beendevelopedforpredictinggroundwaterlevelinan observation well identified in Udupi district. Initially ANN modelling was carried out using feed forward neuralnetworkarchitecturetopredictgroundwater level. The inputs of the ANN model were monthly rainfall record and water level for period of 10 years. The hybrid ANN-GA model was developed and the results are compared with the ANN gradient descent algorithm. The performance of ANN and ANN-GA algorithms was evaluated. It is observed that the performance of ANN-GA is considered superior than ANNmodel.Thus,ANN-GAhybridalgorithmcanbe usedforpredictinggroundwaterlevelsoverthestudy area. Further, more investigations needed on the field generated data in groundwater level forecasting to have a precise statement

REFERENCE

[1] Sreekanth P.D., Geethanjali N., Sreedevi P.D, Shakeel Ahmed, Ravi Kumar N. and Kamala Jayanthi P.D., Forecasting groundwater level using artificial neural networks., Journal of Current Science, Vol. 96, No. 7, 2009.

[2] Hosseini Z. and Nakhaei M., Estimation of ground water level using a hybrid genetic algorithm Neural network., Journal of pollution, Vol 1(1), Winter, pp.9-12, 2015.

[3] Banerjee P., Prasad R.K. and Singh V.S., Forecasting of groundwater level in hard rock regionusingartificialneuralnetwork.,Journalof Environmetal Geology, 58(6),pp 1239-1246, 2009

[4] S.M.Nasseri,K.AsghariandM.J.Abedini, ,Optimizedscenarioforrainfallforecastingusing GA coupled with artificial neural network, Science direct, Elsevier, Vol 35, pp. 1415-1421,2008.

[5] Coulibaly P, Anctil F, Aravena R, Bobee B., Artificialneuralnetworkmodelingofwatertable depth fluctuations, Water Resources Research, 37(4): 885-896,2001.

[6] Lallahem, S., and Mania, J., Evaluation and forecastingofdailygroundwaterinflowinasmall chalky watershed., Hydrological Process., 17(8), 1561- 1577,2003.

[7] Chau, K. W., Application of a PSO-based neural network in analysis of outcomes of construction claims. Automation in Constructions. 16(5), 642 646,2007

[8] Daliakopoulos,I.,Coulibaly,P.,andTsanis,I.K., Groundwater level forecasting using artificial neural networks., Journal of Hydrology., 309, pp.229-240,2005.