Ground Water Level Priduction Using Random Forest Algoritm &DCNN

M Jyothsna[1], K Yanadi[2]
[1]*PG Student, CSE, Priyadarshini College of Engineering & Technology*
[2]*Associate professor, CSE, Priyadarshini College of Engineering & Technology*

***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 resources exploitation. This paper investigates the ability of a hybrid model of artificial neural network (ANN) and genetic algorithm (GA) in predicting groundwater levels in an observation well from Udupi district. The ground water level for a period of ten years and rainfall data for 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 is developed using artificial neural network. The Genetic Algorithm is used to determine the optimized weights for 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

Groundwater is one of the major sources of supply for 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, water quality and its management. For this, a constant monitoring of the groundwater levels is extremely important. The water levels, if forecast well in advance, helps administrators to better plan the groundwater utilization. A continuous forecast of groundwater levels is required to effective use of any simulation model for water management and overall development. In this regard, it is important to develop a fast and cost-effective method for aquifer simulation with an acceptable accuracy. Towards this goal, many researchers have used intelligent systems including, Coulibaly et al., Daliakopoulos et 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 the way biological nervous systems, such as the brain, processes information. It determines the relationship between inputs and outputs of physical systems by a network of interconnecting nodes adjusted by connecting weights based on the training samples, 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. It has an ability to learn and apply the knowledge based on the data given for training or initial experience

II. LITERATURE SURVEY

2.1 INTRODUCTION:

1. Title: " Estimation of ground water level using a hybrid genetic algorithm Neural network" Reference: https://jpoll.ut.ac.ir/article_52176_2d68fe031d5a1663214492a904b999f5.pdf

Year: 2015

Authors: Hosseini Z. and Nakhaei M.,

Description : In this paper, we present an application of evolved neural networks using a real coded genetic 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 developed hybrid genetic algorithm-back propagation (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 algorithm has been compared with the individual back 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. SYSTEM ANALYSIS

1.1 EXISTING SYSTEM:

- 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.
- The stochastic optimization method (GA) to train a FNN; therefore, numerical weights of neuron connections and biases represent the solution components of the optimization problem.
- The stochastic optimization method (GA) to train a FNN; therefore, numerical weights of neuron connections and biases represent the solution components of the optimization problem.

3.1.1DISADVANTAGES OF EXISTING SYSTEM:

- The existing system model with Logistic regression fails to predict a continuous outcome.
- The existing system model with Logistic regression may not be accurate if the sample size is too small.
- The existing system may lead to overfitting problem.
- The existing system accuracy depends on the quality of the data.

1.2 PROPOSED SYSTEM:

- 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.
- ANN trained model applied on test data to predict 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.

3.2.1ADVANTAGES OF PROPOSED SYSTEM:

- The proposed system reduces overfitting in decision trees and helps to improve the accuracy.
- It is flexible to both classification and regression problems.
- It works well with both categorical and continuous values.
- It automates missing values present in the data. ϖ Normalizing of data is not required as it uses a rule-based approach.
- Decreased the total number of verification steps and measures

3.3.SYSTEM REQUIREMENTS

3.3.1.HARDWARE REQUIREMENTS(minimum):

- System : Pentium i3 Processor
- Hard Disk : 500 GB
- Ram : 4 GB.

3.3.2. SOFTWARE REQUIREMENTS:

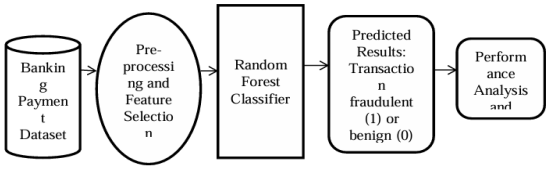
Operating system	:	Windows XP/7.
Coding Language	:	Python
Web Framework	:	Flask

1.3 MODULES:

- Tensor flow
- NumPy

- pandas
- Matplotlib
- Scikit -learn

IV. SYSTEM ARCHITECTURE

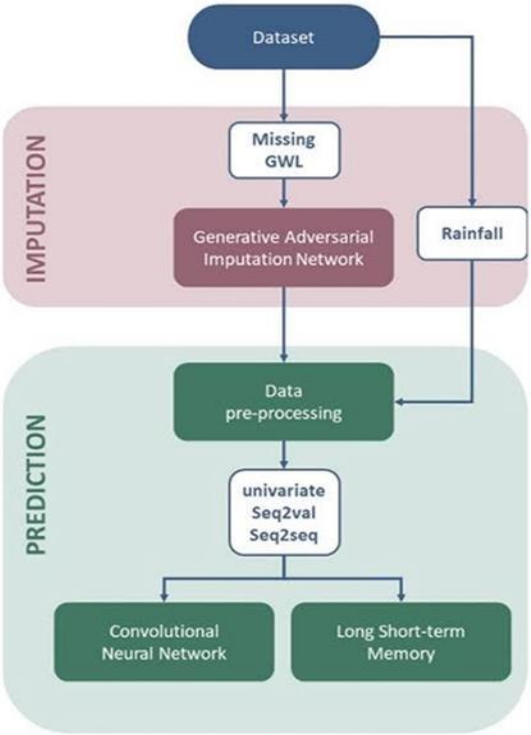


V. SYSTEM DESIGN:DATA FLOW DIAGRAM

5.1 DATA FLOW DIAGRAM:

DATA FLOW DIAGRAM:

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.



5.2 CLASS DIAGRAM:

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. Active category is employed in a very category diagram to represent the concurrency of the system.

VI. SOFTWARE ENVIRONMENT

What is Python :- Below are some facts about Python. Python is currently the most widely used multi-purpose, high-level programming language. Python allows programming in Object-Oriented and Procedural paradigms. Python programs generally are smaller than other programming languages like Java. Programmers have to type relatively less and indentation requirement of the language, makes them readable all the time. Python language is being used by almost all tech-giant companies like – Google, Amazon, Facebook, Instagram, Dropbox, Uber… etc.

Advantages of Python: -

Let us see how Python dominates over other languages.

1. Extensive Libraries

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

Complimentary to extensibility, Python is embeddable as well. You can put your Python code in your source

code of a different language, like C++. This lets us add scripting capabilities to our code in the other language.

VII. SYSTEM IMPLEMENTATION

Sample code:

```
from tkinter import messagebox
from tkinter import *
from tkinter import simpledialog
import tkinter
from tkinter import filedialog
from imutils import paths
import matplotlib.pyplot as plt
import numpy as np
from tkinter.filedialog import askopenfilename
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn import linear_model
from sklearn.metrics import accuracy_score
from sklearn.model_selection import train_test_split
from keras.models import Sequential
from keras.layers import Dense, Dropout, Activation
from keras import optimizers
from genetic_selection import GeneticSelectionCV
import webbrowser
from sklearn.metrics import mean_squared_error
main = tkinter.Tk()
main.title("Groundwater Level Prediction Using
Hybrid Artificial Neural Network with
Genetic Algorithm")
main.geometry("1300x1200")
global filename, dataset
global X, Y, X_train, X_test, y_train, y_test, Y1
global mse, text, pathlabel
```

VIII SYSTEM TESTING

SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way 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 in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement.

TYPES OF TESTS

Unit testing

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. This is a structural testing, that relies on knowledge of 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 defined inputs and expected results. Integration testing Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. 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 testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

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:

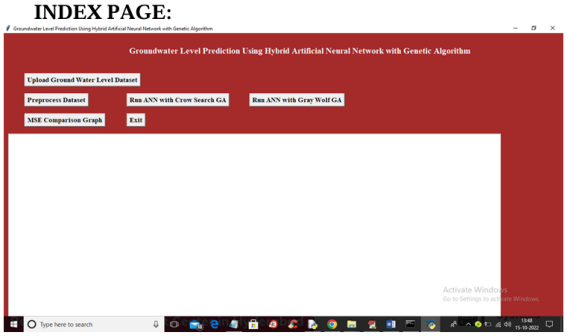


Fig 9.1: Index Page

Uploading Data Sets:

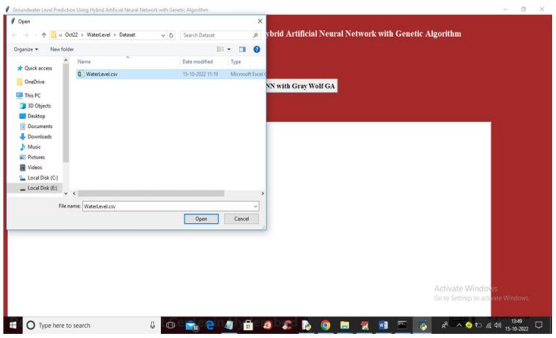


Fig 9.2: Uploading Data Sets

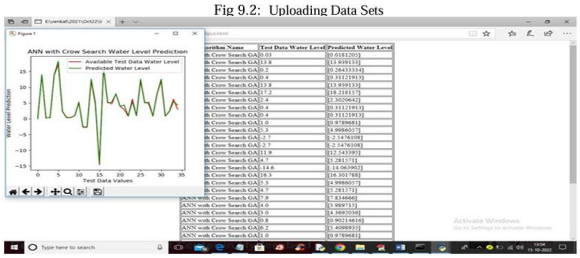


CHART:

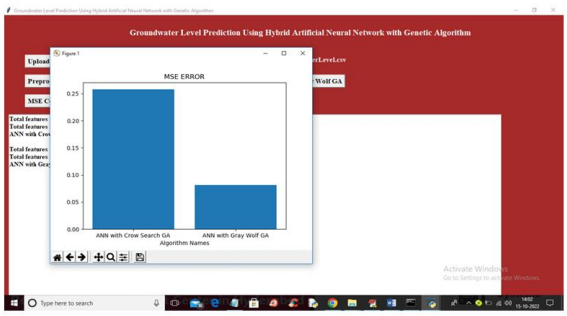


Fig 9.8: Chart

X.CONCLUSION

In this paper, two approaches of soft computing have been developed for predicting groundwater level in an observation well identified in Udupi district. Initially ANN modelling was carried out using feed forward neural network architecture to predict groundwater 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 ANN model. Thus, ANN-GA hybrid algorithm can be used for predicting ground water levels over the study 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 region using artificial neural network., Journal of Environmetal Geology, 58(6),pp 1239-1246, 2009

[4] S. M.Nasseri, K.Asghari and M.J.Abedini, ,Optimized scenario for rainfall forecasting using GA coupled with artificial neural network, Science direct, Elsevier, Vol 35, pp. 1415-1421,2008.

[5] Coulibaly P, Anctil F, Aravena R, Bobee B., Artificial neural network modeling of water table depth fluctuations, Water Resources Research, 37(4): 885-896,2001.

[6] Lallahem, S., and Mania, J., Evaluation and forecasting of daily groundwater inflow in a small 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., and Tsanis, I. K., Groundwater level forecasting using artificial neural networks., Journal of Hydrology., 309, pp.229-240,2005.