**Air Quality Index Forecasting via Genetic Algorithm-Based Improved Extreme Learning Machine**

**Abstract**

Air quality has always been one of the most important environmental concerns for the general public and society. Using machine learning algorithms for Air Quality Index (AQI) prediction is helpful for the analysis of future air quality trends from a macro perspective. When conventionally using a single machine learning model to predict air quality, it is challenging to achieve a good prediction outcome under various AQI fluctuation trends. In order to effectively address this problem, a genetic algorithm-based improved extreme learning machine (GA-KELM) prediction method is enhanced. First, a kernel method is introduced to produce the kernel matrix which replaces the output matrix of the hidden layer. To address the issue of the conventional limit learning machine where the number of hidden nodes and the random generation of thresholds and weights lead to the degradation of the network learning ability, a genetic algorithm is then used to optimize the number of hidden nodes and layers of the kernel limit learning machine. The thresholds, the weights, and the root mean square error are used to define the fitness function. Finally, the least squares method is applied to compute the output weights of the model. Genetic algorithms are able to find the optimal solution in the search space and gradually improve the performance of the model through an iterative optimization process. In order to verify the predictive ability of GA-KELM, based on the collected basic data of long-term air quality forecast at a monitoring point in a city in China, the optimized kernel extreme learning machine is applied to predict air quality ( $SO\_{2}$ , $NO\_{2}$ , $PM\_{10}$ , $CO$ , $O\_{3}$ , $PM\_{2.5}$ concentration and AQI), with comparative experiments based CMAQ (Community Multiscale Air Quality), SVM (Support Vector Machines) and DBN-BP (Deep Belief Networks with Back-Propagation). The results show that the proposed model trains faster and makes more accurate predictions.

**Existing System**

In the Existing system, many deep and machine learning algorithms are introduced but their performance is not accurate as training weights of those algorithms are not accurate enough to predict Air Quality with high accuracy and less error rate like MSE (mean square error) and RMSE (root mean square error). Both MSE and RMSE refers to difference between original and predicted values so the lower the MSE the better is the mode.

**Disadvantages**

1.Less accuracy.

**PROPOSED SYSTEM**

In this proposed system, to update weights accurately author of this paper enhancing Extreme Learning Machine with Genetic Algorithm (GA-KLEM). To solve the issue of the conventional limit learning machine where the number of hidden nodes and the random generation of thresholds and weights lead to the degradation of the network learning ability, a genetic algorithm is then used to optimize the number of hidden nodes and layers of the kernel limit learning machine. The thresholds, the weights, and the root mean square error are used to define the fitness function. Finally, the least squares method is applied to compute the output weights of the model. Genetic algorithms are able to find the optimal solution in the search space and gradually improve the performance of the model through an iterative optimization process.

**Advantages**

1. High Accuracy

**SYSTEM REQUIREMENTS**

**HARDWARE REQUIREMENTS**:

Processor - Intel i3 or higher

Speed - 1.1 GHz

RAM - 4 GB (min)

Hard Disk - 500 GB (min)

**SOFTWARE REQUIREMENTS:**

Operating System - Windows 10 or above

Programming Language - Python with Jupiter notebook