CONTENTS

ABSTRACT

ABSTRACT

One major basic right is clean air that is integral to the concept of citizenship and it's while not a doubt, the responsibility of every subject to try to do his/her half to stay the air clean. Air quality prognostication has been looked into because the key answer of early warning and management of pollution. During this paper, we tend to propose an Associate in nursing air quality prediction system supported by a machine learning framework known as the sunshine GBM model, to predict air quality. This model, trained victimization lightweight GBM classifier, take meteorology knowledge jointly of sources for predicting the air quality thereby increasing the prediction accuracy by creating full use of obtainable abstraction data. the prevailing air quality observance stations and satellite meteorologic knowledge offer period air quality observance info that is employed to predict the trend of air pollutants within the future. The projected system was found to administer Associate in nursing accuracy of ninety-two

LITERATURE SURVEY

1) Air pollutant severity reduction using Bi-directional LSTM Network:

Author: Verma, Ishan, Rahul Ahuja, Hardik Meisheri, and Lipika Dey

Air pollution has emerged as a universal concern across the globe affecting human health. This increasing danger motivates the study of systems for predicting air pollutant severities ahead of time. In this paper, we have proposed the use of a bi-directional LSTM model to predict air pollutant severity levels ahead of time. We have shown that the predictions can be significantly improved using an ensemble of three Bi-Directional LSTMs (BiLSTM) that model the long-term, short-term and immediate effects of PM2.5 (the key air pollutant) severity levels. Further, weather information data has been taken into account while modelling, since they are found to boost prediction accuracies. Experimental results for multiple locations in New Delhi, India are presented to demonstrate model superiority over earlier techniques.

2) Urban air quality based on Bayesian network

Author: Yang, Ruijun, Feng Yan, and Nan Zhao.

In the recent years, the rapid increase of industrial pollution, scientific evaluation of air quality is an important research topic. This is the Bias network applied to the city air quality evaluation of air quality characteristics, establish a directed acyclic graph (DAG), and the air quality data of Shanghai for training and validation, the experimental results coincide with the actual situation, the accuracy rate reached 99.3169%, that Bias network used in the feasibility and effectiveness of the air quality evaluation.

3)" Evolving Differential evolution method with random forest for prediction of Air Pollution"

Author: Kumar, Dinesh

The aim of this paper is to use a heterogeneous ensemble of differential evolution with random forest method for air pollution prediction. This is different from existing work (independent classifier of Bayesian network and multi-label classifier used for the estimation of air pollutants) as a method is proposing to combine state-of-the-art differential evolution strategies with random forest method instead of focusing on existing single technique. When the existing approach i.e. independent and multi-label classifiers are compared with proposed approach, it shows proposed approach leads to the performance gains. Continuous ambient air quality data of two cities Delhi and Patna from Central Pollution Control Board were publicly made available, from where seven pollutants (C6H6, NO2, O3, SO2, CO, PM2.5 and PM10) dataset are collected with daily average concentration.

4) "Exploring the utility of the random forest method for forecasting ozone pollution in SYDNEY

Author: Jiang, Ningbo, and Matthew L. Riley

This paper explores the utility of an ensemble decision-tree method called random forest, in comparison with the classic classification and regression trees (CART) algorithm, for forecasting ground-level ozone pollution in the Sydney metropolitan region. Statistical forecasting models are developed to provide daily ozone forecasts in November-March for three subregions, i.e., Sydney east, Sydney south-west and Sydney north-west. The random forest models are evaluated in reference to the single decision-tree models developed from the classic CART algorithm. The results show that the random forest models outperform the CART models for forecasting high ozone pollution in Sydney south-west and Sydney north-west, the areas where the highest ozone pollution are observed. The random forest models also show a lift in forecasting skills in Sydney south-west if compared to the existing forecasting practice for the basin as a whole. These results suggest that random forest is a promising method for air quality forecasting in Sydney. This study promotes the application of a statistical ensemble approach to air quality forecasting.

5) Random forest: a classification and regression tool for compound classification and QSAR modelling

Author: Svetnik, Vladimir, et al.

A new classification and regression tool, Random Forest, is introduced and investigated for predicting a compound's quantitative or categorical biological activity based on a quantitative description of the compound's molecular structure. Random Forest is an ensemble of unpruned classification or regression trees created by using bootstrap samples of the training data and random feature selection in tree induction. Prediction is made by aggregating (majority vote or averaging) the predictions of the ensemble. We built predictive models for six cheminformatics data sets. Our analysis demonstrates that Random Forest is a powerful tool capable of delivering performance that is among the most accurate methods to date. We also present three additional features of Random Forest: built-in performance assessment, a measure of relative importance of descriptors, and a measure of compound similarity that is weighted by the relative importance of descriptors. It is the combination of relatively high prediction accuracy and its collection of desired features that makes Random Forest uniquely suited for modelling in cheminformatics.

CHAPTER-1
INTRODUCTION

CHAPTER-2
SYSTEM ANALYSIS

CHAPTER-3
SYSTEM STUDY

CHAPTER-4
SOFTWARE ENVIROMENT

CHAPTER-5
SYSTEM DESING

CHAPTER-6
INPUT AND OUTPUT DESINGS

CHAPTER-7
SYSTEM REQURIMENTS

CHAPTER-8
SAMPLE SOURCE CODE

CHAPTER-9
MODULES

CHAPTER-10
SYSTEM TESTING

CHAPTER-11
SAMPLE TEST CASES

CHAPTER-12
SCREEN SHOTS

CHAPTER 13

FUTURE ENHANCEMENT

CHAPTER 14
CONCLUSION

CHAPTER 15
REFERENCES