World Academy of Science, Engineering and Technology
International Journal of Environmental and Ecological Engineering
Vol:14, No:06, 2020

An Artificial Intelligence Framework to Forecast Air Quality

Authors : Richard Ren

Abstract : Air pollution is a serious danger to international well-being and economies - it will kill an estimated 7 million people every year, costing world economies $2.6 trillion by 2060 due to sick days, healthcare costs, and reduced productivity. In the United States alone, 60,000 premature deaths are caused by poor air quality. For this reason, there is a crucial need to develop effective methods to forecast air quality, which can mitigate air pollution's detrimental public health effects and associated costs by helping people plan ahead and avoid exposure. The goal of this study is to propose an artificial intelligence framework for predicting future air quality based on timing variables (i.e. season, weekday/weekend), future weather forecasts, as well as past pollutant and air quality measurements. The proposed framework utilizes multiple machine learning algorithms (logistic regression, random forest, neural network) with different specifications and averages the results of the three top-performing models to eliminate inaccuracies, weaknesses, and biases from any one individual model. Over time, the proposed framework uses new data to self-adjust model parameters and increase prediction accuracy. To demonstrate its applicability, a prototype of this framework was created to forecast air quality in Los Angeles, California using datasets from the RP4 weather data repository and EPA pollutant measurement data. The results showed good agreement between the framework's predictions and real-life observations, with an overall 92% model accuracy. The combined model is able to predict more accurately than any of the individual models, and it is able to reliably forecast season-based variations in air quality levels. Top air quality predictor variables were identified through the measurement of mean decrease in accuracy. This study proposed and demonstrated the efficacy of a comprehensive air quality prediction framework leveraging multiple machine learning algorithms to overcome individual algorithm shortcomings. Future enhancements should focus on expanding and testing a greater variety of modeling techniques within the proposed framework, testing the framework in different locations, and developing a platform to automatically publish future predictions in the form of a web or mobile application. Accurate predictions from this artificial intelligence framework can in turn be used to save and improve lives by allowing individuals to protect their health and allowing governments to implement effective pollution control measures.Air pollution is a serious danger to international wellbeing and economies - it will kill an estimated 7 million people every year, costing world economies $2.6 trillion by 2060 due to sick days, healthcare costs, and reduced productivity. In the United States alone, 60,000 premature deaths are caused by poor air quality. For this reason, there is a crucial need to develop effective methods to forecast air quality, which can mitigate air pollution's detrimental public health effects and associated costs by helping people plan ahead and avoid exposure. The goal of this study is to propose an artificial intelligence framework for predicting future air quality based on timing variables (i.e. season, weekday/weekend), future weather forecasts, as well as past pollutant and air quality measurements. The proposed framework utilizes multiple machine learning algorithms (logistic regression, random forest, neural network) with different specifications and averages the results of the three top-performing models to eliminate inaccuracies, weaknesses, and biases from any one individual model. Over time, the proposed framework uses new data to self-adjust model parameters and increase prediction accuracy. To demonstrate its applicability, a prototype of this framework was created to forecast air quality in Los Angeles, California using datasets from the RP4 weather data repository and EPA pollutant measurement data. The results showed good agreement between the framework's predictions and real-life observations, with an overall 92% model accuracy. The combined model is able to predict more accurately than any of the individual models, and it is able to reliably forecast season-based variations in air quality levels. Top air quality predictor variables were identified through the measurement of mean decrease in accuracy. This study proposed and demonstrated the efficacy of a comprehensive air quality prediction framework leveraging multiple machine learning algorithms to overcome individual algorithm shortcomings. Future enhancements should focus on expanding and testing a greater variety of modeling techniques within the proposed framework, testing the framework in different locations, and developing a platform to automatically publish future predictions in the form of a web or mobile application. Accurate predictions from this artificial intelligence framework can in turn be used to save and improve lives by allowing individuals to protect their health and allowing governments to implement effective pollution control measures.Air pollution is a serious danger to international wellbeing and economies - it will kill an estimated 7 million people every year, costing world economies $2.6 trillion by 2060 due to sick days, healthcare costs, and reduced productivity. In the United States alone, 60,000 premature deaths are caused by poor air quality. For this reason, there is a crucial need to develop effective methods to forecast air quality, which can mitigate air pollution's detrimental public health effects and associated costs by helping people plan ahead and avoid exposure. The goal of this study is to propose an artificial intelligence framework for predicting future air quality based on timing variables (i.e. season, weekday/weekend), future weather forecasts, as well as past pollutant and air quality measurements. The proposed framework utilizes multiple machine learning algorithms (logistic regression, random forest, neural network) with different specifications and averages the results of the three top-performing models to eliminate inaccuracies, weaknesses, and biases from any one individual model. Over time, the proposed framework uses new data to self-adjust model parameters and increase prediction accuracy. To demonstrate its applicability, a prototype of this framework was created to forecast air quality in Los Angeles, California using datasets from the RP4 weather data repository and EPA pollutant measurement data. The results showed good agreement between the framework's predictions and real-life observations, with an overall 92% model accuracy. The combined model is able to predict more accurately than any of the individual models, and it is able to reliably forecast season-based variations in air quality levels. Top air quality predictor

variables were identified through the measurement of mean decrease in accuracy. This study proposed and demonstrated the efficacy of a comprehensive air quality prediction framework leveraging multiple machine learning algorithms to overcome individual algorithm shortcomings. Future enhancements should focus on expanding and testing a greater variety of modeling techniques within the proposed framework, testing the framework in different locations, and developing a platform to automatically publish future predictions in the form of a web or mobile application. Accurate predictions from this artificial intelligence framework can in turn be used to save and improve lives by allowing individuals to protect their health and allowing governments to implement effective pollution control measures.
Keywords : air quality prediction, air pollution, artificial intelligence, machine learning algorithms
Conference Title : ICEEWMI 2020 : International Conference on Environmental Engineering, Waste Management and Incineration
Conference Location : New York, United States
Conference Dates : June 04-05, 2020