Machine learning algorithms to forecast air quality: a survey

Manuel Méndez[1] · Mercedes G. Merayo[1] · Manuel Núñez[1]

Abstract

Air pollution is a risk factor for many diseases that can lead to death. Therefore, it is important to develop forecasting mechanisms that can be used by the authorities, so that they can anticipate measures when high concentrations of certain pollutants are expected in the near future. Machine Learning models, in particular, Deep Learning models, have been widely used to forecast air quality. In this paper we present a comprehensive review of the main contributions in the field during the period 2011–2021. We have searched the main scientific publications databases and, after a careful selection, we have considered a total of 155 papers. The papers are classified in terms of geographical distribution, predicted values, predictor variables, evaluation metrics and Machine Learning model.

Keywords Machine learning · Deep learning · Regression algorithms · Air quality

1 Introduction

The economical and urban development of cities has brought that the interest in environmental pollution has risen during the last years. Among other problems, air pollution has a big impact in human health, being a risk factor for many diseases that can lead to death. According to the World Health Organisation (WHO), air pollution is a "silent killer" that produces the premature death of almost seven million people each year,[1] including 600.000 children.

---

[1] https://www.scmp.com/news/world/europe/article/3149735/air-pollution-kills-7-million-year-says-who-it-tightens, last accessed on 2022-02-15.

Mercedes G. Merayo and Manuel Núñez have contributed equally to this work.

✉ Manuel Méndez
  manumend@ucm.es

  Mercedes G. Merayo
  mgmerayo@fdi.ucm.es

  Manuel Núñez
  mn@sip.ucm.es

[1] Design and Testing of Reliable Systems Research Group, Universidad Complutense de Madrid, C/ Profesor José García Santesmases, 9, 28040 Madrid, Madrid, Spain