

An Accurate PM2.5 Concentration Prediction of Low Cost Sensor Tool Using Multiple Regression Model with Gradient Descent Algorithm

Nabila Zahra Ekaputri¹, Yuyun Hidayat², Gumgum Darmawan³

^{1, 2, 3} Department of Statistics, Padjadjaran University, Indonesia.

(E-mail: nabila20005@mail.unpad.ac.id, yuyun.hidayat@unpad.ac.id, gumgum@unpad.ac.id)

Keywords: Gradient Descent, Multiple Regression, PM2.5

ABSTRACT

BMKG (Meteorology, Climatology, and Geophysics Agency) is a government agency in Indonesia responsible for weather, climate, and geophysical monitoring [1]. BMKG is tasked with making observations covering climate and air quality as mentioned in Law No. 31 of 2009. Air quality parameters observed according to BMKG Regulation of the Republic of Indonesia No. 2 of 2020 concerning the provision and dissemination of air quality information, one of which is PM2.5. PM2.5 (Particulate Matter 2.5) refers to small particles in the air less than or equal to 2.5 micrometers (μm). PM 2.5 is a concern because it is small enough to go directly into the lungs and even the bloodstream [2].

BMKG has been monitoring PM2.5 since 2020 using BAM 1020 tools spread throughout Indonesia. BMKG as an agency that always adapts to technological developments is trying to develop a PM2.5 concentration measurement tool, namely a low cost sensor. However, the low cost sensor cannot be used commercially due to an overestimation between the PM2.5 monitoring results of the BAM 1020 tool and the low cost sensor tool. In this study, model assessment will be carried out using the gradient descent algorithm as an optimization method in estimating parameters in the PM2.5 concentration model with the aim of producing prediction values that tend to be close to the PM2.5 concentration value of the BAM 1020 tool based on the factors that influence it. Gradient descent is an iteration process that aims to find the parameter values (coefficients) of a function (f) that minimize the cost function [3]. This study is expected to produce a predicted value of PM 2.5 concentration that is close to the true value, where the difference between the predicted value and the actual value is less than 10%.

References:

- [1] BMKG. (2022). Retrieved August, 2023 from Badan Meteorologi, Klimatologi, dan Geofisika: <https://www.bmkg.go.id/>
- [2] U.S. Embassy. (2023). Retrieved August, 2023 from U.S . Embassy & Consulates in Indonesia: <https://id.usembassy.gov/>
- [3] Juliastoro, W., Cholissodin, I., Bachtiar, F. (2021). Sistem Prediksi Hasil Produksi Udang Vaname menggunakan Algoritma Multiple Linear Regression (MLR) Kombinasi Gradient Descent (GD) dengan Apache Spark. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer. Vol. 5, No.11, hlm. 4920-4928