

UNIVERSIDAD PERUANA UNIÓN

ESCUELA DE POSGRADO

Unidad de Posgrado de ingeniería y arquitectura



**Enfoque predictivo para la concentración de contaminante del
aire basado en un modelo de red neuronal artificial**

Tesis para obtener el Título de Segunda Especialidad Profesional de
Ingeniería: Estadística Aplicada para Investigación

Autor:

Carlos Andres Guerra Bendezu

Vivian Isabel Romani Franco

Asesor:

Dr. Javier Linkolk Lopez Gonzales

Lima, Abril de 2024

DECLARACIÓN JURADA DE ORIGINALIDAD DEL TRABAJO ACADÉMICO

Yo Javier Linkolk Lopez Gonzales, docente de la Unidad de Posgrado de Ingeniería y Arquitectura de la Escuela de Posgrado de la Universidad Peruana Unión.

DECLARO:

Que la presente investigación titulada: **“Enfoque predictivo para la concentración de contaminante del aire basado en un modelo de red neuronal artificial”** de los autores Vivian Isabel Romani Franco y Carlos Andres Guerra Bendezu tiene un índice de similitud de 15% verificable en el informe del programa Turnitin, y fue realizada en la Universidad Peruana Unión bajo mi dirección.

En tal sentido asumo la responsabilidad que corresponde ante cualquier falsedad u omisión de los documentos como de la información aportada, firmo la presente declaración en la ciudad de Lima, a los 15 días del mes de Abril del año 2024



Javier Linkolk López Gonzales

ACTA DE SUSTENTACIÓN DE TESIS

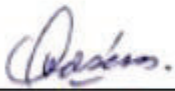
En Lima, Naña, Villa unión a 15 días del mes de abril del año 2024, siendo las 10:30 horas, se reunieron de forma online sincrónica, bajo la dirección del presidente del jurado Dra. Dámaris Susana Quinteros Zuñiga, el secretario Mg. Carlos Daniel Abanto Ramirez; los demás miembros: Mg. Lizeth Huanca López, Mg. Esteban Tocto Cano y el asesor PhD. Javier Linkolk López Gonzales, con el propósito de administrar el acto académico de sustentación de Tesis de la Segunda Especialidad titulada "Enfoque predictivo para la concentración de contaminante del aire basado en un modelo de red neuronal artificial", conducente a la obtención del Título de Segunda Especialidad Profesional de ingeniería: Estadística Aplicada para Investigación.

El presidente inició el acto académico de sustentación invitando a los candidatos a hacer uso del tiempo determinado para su exposición. Concluida la exposición, el presidente invitó a los demás miembros del jurado a efectuar las preguntas, cuestionamientos y aclaraciones pertinentes, los cuales fueron absueltos por los candidatos. Luego, se produjo un receso para las deliberaciones y la emisión del dictamen del jurado. Posteriormente, el jurado procedió a dejar constancia escrita sobre la evaluación en la presente acta, con el dictamen siguiente:

Candidatos: Vivian Isabel Romani Franco y Carlos Andres Guerra Bendezu

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	20	B	Bueno	Excelente

Finalmente, el presidente del jurado invitó a los candidatos a ponerse de pie, para recibir la evaluación final. Además, el presidente concluyó el acto académico de sustentación, procediéndose a registrar a registrar las firmas respectivas.



Presidente



Secretario



Asesor



Miembro



Miembro



Candidato



Candidato

Enfoque predictivo para la concentración de contaminante del aire basado en un modelo de red neuronal artificial

Carlos Andres Guerra Bendezu ¹, Vivian Isabel Romaní Franco ¹, Ana María Gómez Lamus ², Romina Torres ³, Paulo Canas Rodrigues ⁴, Rodrigo Salas ^{5,6} and Javier Linkolk López-Gonzales ¹

¹ Escuela de Posgrado, Universidad Peruana Unión, Lima 15468, Peru; javierlinkolk@gmail.com (J.L.L.-G.)

² Statistical Engineering, Escuela Colombiana de Ingeniería Julio Garavito, Bogotá, Colombia; ana.gomez@escuelaing.edu.co (A.M.G.L.)

³ Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Viña del Mar, Chile; romina.torres.f@uai.cl(R.T.)

⁴ Department of Statistics, Federal University of Bahia, Salvador 40170-110, Brazil; paulocanas@gmail.com (P.C.R.)

⁵ Escuela de Ingeniería C. Biomédica, Universidad de Valparaíso, Valparaíso, Chile; rodrigo.salas@uv.cl (R.S.)

⁶ Millennium Institute for Intelligent Healthcare Engineering (iHealth), Santiago, Chile

* Correspondence: javierlinkolk@gmail.com

Abstract: In this study, we propose a new hybrid method based on artificial neural networks to forecast daily extreme events of PM_{2.5} pollution concentration. The hybrid method combines self-organizing maps to identify temporal patterns of excessive daily pollution found at different monitoring stations, with a set of multilayer perceptron to forecast extreme values of PM_{2.5} for each cluster. The proposed model was applied to analyze five-year pollution data obtained from nine weather stations in the metropolitan area of Santiago, Chile. Simulation results show that the hybrid method improves the performance metrics when forecasting daily extreme values of PM_{2.5}.

Keywords: Air pollution; Hybrid methodology; Artificial Neural Networks; Time series Forecasting

1. Introduction

Air quality monitoring is important for the sustainable growth of cities, mitigating the risks that may affect human health. According to the American Environmental Protection Agency [1], particle pollution includes inhalable particles of 2.5 (PM_{2.5}) and 10 (PM₁₀) microns or smaller, whose source could be construction sites, unpaved roads, and forest fires, to name a few. When inhaled, these particles may cause serious health problems [2–4]. There is a correlation between air pollution exposure and severe health concerns such as the incidence of respiratory and cardio-respiratory diseases [5] and even deaths. Thus, PM_{2.5} concentrations are measured constantly in the main cities to determine if they are under the national permissible limit [6–9]. To be prepared for the next extreme event, it is essential to forecast the intensity of the pollution levels, being the air pollutant concentrations influenced by emission levels, meteorological conditions, and geography.

These methods successfully model PM_{2.5} and other pollution levels, generating accurate forecasts given their non-linear mapping and learning capabilities. Deep learning methods have been successfully applied to predict PM_{2.5} concentrations [10]. For instance, attention-based LSTM [11] and convolutional neural networks (CNN)–LSTM models [12,13] have been used for improving the performance of PM_{2.5} forecasting using data from multiple monitoring locations. Even though predicting the average PM_{2.5} concentrations is a problem reasonably addressed in the literature with various techniques, approaches that can also address the forecasting of pollution peaks are still challenging.

In this work, we propose a new hybrid method based on machine learning to predict extreme daily events of PM_{2.5} pollution concentration. Without loss of generality, in this study we consider extreme events to be the highest daily levels, particularly in the 75th and
