



# Reduction of the Size of Datasets by using Evolutionary Feature Selection: the Case of Noise in a Modern City\*

\***Note:** The full contents of this paper have been published in the volume *Lecture Notes in Artificial Intelligence 11160* (LNAI 11160)

Javier Luque, Jamal Toutouh, Enrique Alba  
*Departamento de Lenguajes y Ciencias de la Computación*  
*Universidad de Málaga*  
Málaga, Spain  
javierluque@uma.es, {jamal, eat}@lcc.uma.es

**Abstract**—Smart city initiatives have emerged to mitigate the negative effects of a very fast growth of urban areas. Most of the population in our cities are exposed to high levels of noise that generate discomfort and different health problems. These issues may be mitigated by applying different smart cities solutions, some of them require high accurate noise information to provide the best quality of serve possible. In this study, we have designed a machine learning approach based on genetic algorithms to analyze noise data captured in the university campus. This method reduces the amount of data required to classify the noise by addressing a feature selection optimization problem. The experimental results have shown that our approach improved the accuracy in 20% (achieving an accuracy of 87% with a reduction of up to 85% on the original dataset).

**Index Terms**—Smart city, Genetic algorithm, Feature selection, Noise