S8: Appendix H - Methodology's Algorithmic Processes

Figure S13 depicts the methodology's algorithmic processes.

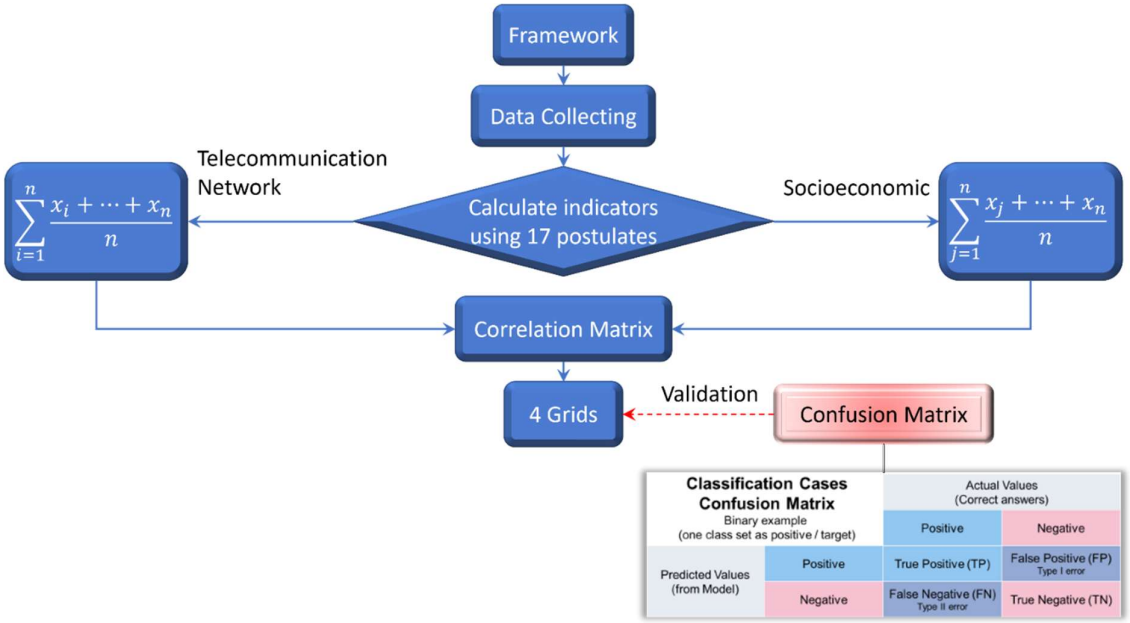

Figure S13. Algorithm Flowchart.

The result is validated by comparable data (between actual and predicted values) and confusion matrix. A confusion matrix is a tabular breakdown of the number of accurate and inaccurate predictions generated. Accuracy is percentage of test data that are correctly classified, where "Accuracy = $\frac{TP+TN}{TP+TN+FP+F}$ " and error rate is one minus accuracy or "Error Rate = $\frac{FP+FN}{TP+TN+FP+FN}$". In addition, we employ two validation strategies:

(i) Test the results in the most developed metropolitan area in Indonesia, namely the province and capital of Indonesia, "DKI Jakarta". Testing was conducted in 44 districts. The actual value is derived from numerous media references pertaining to the situation of the capital, which is still behind in terms of development (for example: slums, absence of communication signals).

(ii) Conduct a national-scale test of results by taking a random sampling of 1%. There were 63 districts [56] tested. The actual value is taken from the Presidential Regulation of the Republic of Indonesia Number 63/2020 concerning the Determination of Disadvantaged Areas in 2020-2024.

REFERENCE

[56] Presidential Regulation of the Republic of Indonesia concerning the Determination of Disadvantaged Regions in 2020-2024 (in the original language (Bahasa): "Peraturan Presiden Republik Indonesia

tentang Penetapan Daerah Tertinggal Tahun 2020-2024", Ministry of Law and Human Rights 63, 2020. Available online: https://peraturan.bpk.go.id/Home/Details/136563/perpres-no-63-tahun-2020 (accessed on 14 April 2022).