link: https://jurnalteknik.unisla.ac.id/index.php/informatika/article/view/1674

Climate-Driven Disease Outbreak Anomaly Detection Using PCA, DBSCAN, and Isolation Forest

Authors

- **Titin Wahyuni**Universaitas Muhammadiyah Makassar
- **Muhammad Faisal**Universitas Muhammadiyah Makassar
- **Muhammad Syafa'at S. Kuba**Universitas Muhammadiyah Makassar
- **Eny Syatriana**Universitas Muhammadiyah Makassar
- **Muh Al Haikal**Universitas Muhammadiyah Makassar

DOI:

https://doi.org/10.30736/jti.v11i02.1674

Keywords:

Climate-driven disease, surveillance, Dangue, Anomaly Detection, Unsupervised learning, PCA, DBSCAN, Isolation Forest, Early Warning Systems

ABSTRACT

Climate variability strongly influences dengue transmission, but prediction-focused surveillance often overlooks unusual climate-dengue configurations that emerge before or outside epidemic peaks. This study proposes a fully unsupervised anomaly-oriented framework for national-level dengue surveillance in Brazil. Weekly dengue case data were integrated with NASA POWER climate variables aggregated to the same epidemiological week scale. Seven standardized climate indicators were reduced using Principal Component Analysis (PCA); the first three components retained 93.34% of the total variance. DBSCAN was then used to identify structurally isolated climate regimes, and Isolation Forest was applied to assign anomaly severity scores and examine complementary transitional deviations. The results show that DBSCAN detected 19 epidemiological weeks (5.19%) as structural anomaly candidates, while Isolation Forest highlighted severity-ranked weeks whose case totals were not always the highest in the study period. To support interpretation, detected anomalies are evaluated using temporal identifiers, dengue case counts, climate-variable deviations, and cross-method robustness rather than treated as automatic outbreak declarations. The findings indicate that climate-driven dengue anomalies are episodic and multivariate and may occur outside peak incidence periods. The proposed framework therefore offers an explainable and reproducible screening layer for climate-informed early warning and risk prioritization in public health surveillance