Application of historical weather data mining in forecasting storm and flood phenomena

Dao Duc Chien

Abstract: This study applies data mining techniques on historical weather datasets to build a disaster prediction model. The author collects data from Open-Meteo and meteorological stations, then applies algorithms such as Random Forest and LightGBM to analyze the correlation between pressure, temperature, and precipitation. The model provides early warnings about the risk of storms and floods in Vietnam, supporting disaster prevention efforts and minimizing human and property damage based on current climate change scenarios.

Keywords: Data Mining, Storm and Flood Forecasting, Historical Weather

References

[1] L. Breiman, "Random Forests," Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

[2] G. Ke et al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in Proc. NeurIPS, 2017.

[3] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, Applied Logistic Regression, 3rd ed. Wiley, 2013.

[4] D. C. Montgomery et al., Introduction to Linear Regression Analysis, 5th ed. Wiley, 2012.

[5] Open-Meteo, "Historical Weather API Documentation," 2024. [Online]. Available:

[6] Nguyễn Đức Ngữ, Bão và áp thấp nhiệt đới ở Việt Nam. Hà Nội: NXB KH&KT;, 2008.