

機器學習應用於臺灣梅雨季天氣類型分析研究

Weather type clustering and classification for Taiwan's Meiyu season using machine learning methods

徐理寰 (Hsu L.-H.) 林冠伶 (Lin K.-L.) 江宙君 (Chiang C.-C.) 吳宜昭 (Wu Y.-c.)
朱容練 (Chu J.-L.) 于宜強 (Yu Y.-C.)

國家災害防救科技中心
National Science and Technology Center for Disaster Reduction

摘要

本研究利用機器學習技術進行臺灣梅雨季天氣類型的聚類與分類研究。資料使用包含1979至2023年東亞地區ERA5大氣環流再分析資料、2001年至2022年GPM IMERG降水觀測與TCCIP台灣1公里高解析度降雨資料。本研究首先利用卷積神經網路自編碼器深度學習演算法(CNN-Autoencoder)建立了萃取資料潛在特徵的模組，進而再對這些潛在特徵進行無監督聚類與監督式分類研究。聚類分析結果可辨識出八種與台灣降雨變化相關的綜觀尺度天氣型態：強鋒面(IF)、中等鋒面(MF)、西南氣流(SW)、南風(S)、南海氣旋環流(CC-SCS)、偏南西南氣流(SS-SW)、東海鋒面(F-ECS)以及副熱帶高壓主導型(SH)。在監督式分類研究方面，使用CNN-Autoencoder萃取出之潛在特徵與聚類得到的天氣類型標籤，共建構出五種基本機器學習分類器，以及一個集成投票分類器(EVC)。該EVC透過投票機制整合五個獨立訓練之分類器的預測結果，來獲得最終的天氣類型分類判斷。校驗結果顯示，採用軟投票機制的EVC表現最佳，在所有天氣類型中平均準確率與F1分數皆可超過0.9。藉由結合高解析度的在地降雨資料與綜觀尺度大氣環流，本研究架構可精確辨識與台灣降水相關的天氣類型，展現出量化天氣類型歷史年際變異性與評估長期模式天氣類型預報效能上的高度潛力。

關鍵字：臺灣梅雨、機器學習聚類分析、天氣類型分類

Abstract

This study investigates Taiwan's Meiyu season using machine learning techniques, focusing on weather type clustering and classification. We analyze the ERA5 circulation data from 1979 to 2023, and GPM IMERG precipitation and TCCIP Taiwan rainfall from 2001 to 2022. A CNN-Autoencoder is employed to extract latent features, followed by unsupervised clustering and supervised classification. The results identify eight synoptic-scale weather patterns linked to varying rainfall over Taiwan: intense fronts (IF), moderate fronts (MF), southwesterly (SW), southerly (S), cyclonic circulation over the South China Sea (CC-SCS), southward-shifted southwesterly (SS-SW), fronts over the East China Sea (F-ECS), and subtropical high dominant (SH). Five machine learning classifiers and one ensemble voting classifier (EVC) are developed using the extracted features and weather type labels. The EVC integrates predictions from the five classifiers via a voting mechanism to determine the final classification. The soft voting EVC achieves the highest average accuracy and F1 score, both exceeding 0.9 across all types. By integrating high-resolution local rainfall with synoptic-scale circulation, the framework enables precise classification of Taiwan's precipitation-related weather types. This study shows strong potential for analyzing historical interannual variability and evaluating long-term model performance across synoptic regimes.

Key words: Meiyu season, Machine learning clustering, Weather type classification