深度學習與系集分類應用於梅雨季及相關天氣類型分析

Exploring Taiwan's Meiyu Season and Associated Weather Types Through Deep Learning and Ensemble Classification

江宙君 (Chiang C.-C.) 林冠伶 (Lin K.-L.) 徐理寰 (Hsu L.-H.) 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

本研究提出一個深度學習框架，客觀辨識台灣梅雨季的期間、並對此季節內的綜觀天氣型態進行分群及分類分析。本研究針對不同氣象變數訓練了獨立的卷積神經網路自編碼器(CNN-Autoencoders)，用來提取變數的潛在特徵。基於ERA5 變數的模型使用了1979至2023年的資料進行訓練，而基於GPM IMERG與TCCIP降水資料的模型則是使用了2001至2023年的資料進行訓練。為了界定梅雨季的起訖時間，本研究透過餘弦相似度評估5天平均ERA5環流特徵(1979–2023)，為次季節的辨識提供了一種新穎的多變數判定方法。進一步針對2001–2023期間，以非監督式k-means演算法對整合後的每日特徵(ERA5、GPM IMERG及TCCIP)進行分群分析，共辨識出8種綜觀天氣類型。其中4種類型與鋒面系統及西南氣流驅動的顯著降水有關，而另外4種則對應與副熱帶高壓相關較乾燥的天氣條件。最後，這些特徵與分群標籤被用來訓練6種機器學習模型，其中包含一個系集投票分類器(Ensemble Voting Classifier, EVC)。透過5次交叉驗證法與單年獨立測試的評估，EVC的準確率、F1分數及精確率均可超過0.94，有效減少了使用單一模型時可能造成的變異性，從而提供較一致的分類結果。這些研究結果顯示，以深度學習提取的變數特徵與系集分類器相結合，對於天氣系統分類研究提供了一套客觀穩健的方法。

關鍵字：深度學習、梅雨、系集投票分類器

Abstract

This study presents a deep learning framework to objectively identify Taiwan's Meiyu season duration, cluster synoptic weather patterns, and classify daily weather types. Independent Convolutional Neural Network (CNN)-Autoencoders are trained to extract latent features. The ERA5-based models are trained on 1979–2023 data, while the rainfall-based models use GPM IMERG and TCCIP data from 2001–2023. To define the Meiyu season boundaries, five-day moving mean ERA5 features (1979–2023) are evaluated via cosine similarity, providing a novel multi-variable approach for sub-season identification. For 2001–2023 period, unsupervised k-means clustering of the integrated daily features (ERA5, GPM IMERG, and TCCIP) identifies eight synoptic weather types. Four reflect significant rainfall driven by frontal systems and southwesterly flows, while four represent drier conditions associated with the subtropical high. Finally, these features and cluster labels train six machine learning models, including an Ensemble Voting Classifier (EVC). Evaluated via 5-fold cross-validation and single-year testing, the EVC exceeds 0.94 in accuracy, F1-score, and precision, effectively mitigating single-model variability to deliver consistent categorizations. These results demonstrate that integrating variable-specific deep-learning feature extraction with an ensemble classifier provides a robust method for objectively characterizing weather systems.

Key words : Deep learning, Meiyu season, Ensemble Voting Classifier