日期：2025/06/03

標題：Future Insights Harnessing AI and Social Media for Advanced Event and Epidemic Forecasting

講者：美國維吉尼亞理工學院暨州立大學 呂昌田 教授

關鍵字：Graph Convolutional Network, Graph Neural Networks, Social Media, Neural Network, Support Vector Machine, Graph Convolution, Convolutional Neural Network, Deep Learning, Mean Absolute Error, Multilayer Perceptron, Social Media Data, Social Media Platforms

本次帶來演講的是遠從美國維吉尼亞理工學院暨州立大學而來的呂昌田教授，主要演講內容為運用人工智慧以及社交媒體資訊預測事件，這套系統功能涵蓋流行病預測、大型示威預測以及暴動預測、潛在集團犯罪預測以及自然災害預測。該項系統已經在美國國家單位上做使用。

呂昌田教授的演講深入探討了如何利用人工智慧與社交媒體數據進行事件預測，這套系統的應用範圍之廣令人印象深刻。從流行病監控到大型示威活動的預測，甚至涵蓋了暴動、潛在集團犯罪以及自然災害的預測，顯示了該技術在多領域的潛力。特別是這套系統已在美國國家單位中實際應用，讓我感受到人工智慧技術在公共安全與社會治理中的重要性。

演講中，呂教授介紹了「EMBERS」系統的核心技術與其運作原理。他提到該系統整合了多個數據來源，包括新聞報導、Twitter（現為X）貼文以及Tor網絡數據，並通過Lasso Regression模型進行數據分析。我特別注意到「Ablation testing」部分的展示，該方法通過排除不相關的數據來源（如部落格和通脹數據），顯著提升了模型的預測準確度。這讓我意識到數據篩選在機器學習中的關鍵性，尤其是在處理社交媒體這類噪音數據較多的情境下，如何有效過濾無用訊息是技術成功的關鍵。

另一個令我印象深刻的部分是「Dynamic Query Expansion」技術。呂教授以墨西哥的#MegaMarcha活動為例，展示如何利用種子詞（如「march」和「protest」）動態擴展查詢，從而精準定位事件發生的地點與時間。這種方法解決了社交媒體數據中常見的非正式語言與模糊表達問題，讓系統能夠更高效地捕捉關鍵訊息。作為資訊工程碩士生，我認為這項技術在自然語言處理（NLP）領域有很大的應用潛力，特別是在多語言環境下，如何進一步最佳化查詢擴展以適應不同語言的語法與文化差異，是未來值得研究的課題。

此外，演講中提到的「Data reduction helps!」部分也讓我受益匪淺。呂教授介紹了一種通過空間聚類和時間鏈接來降低Twitter訊息複雜度的方法，這不僅減少了系統的運算負擔，還顯著降低了誤報的數量。這讓我聯想到在機器學習中，數據預處理的重要性。尤其在處理大規模數據時，如何設計有效的降低維度與過濾策略，直接影響模型的性能與應用效果。我認為這項技術對於台灣這樣一個高度數位化的社會來說，具有很大的應用價值，例如可以用來監測颱風等自然災害的即時影響，或是預測大型活動可能引發的交通問題。

在「Fusion of Predictions」部分，呂教授展示了如何利用Kernel Density方法，基於地理接近度和同日預測，提升墨西哥地區的預測質量。這讓我對多模型融合技術有了更深的理解。通過結合多個模型的預測結果，系統能夠更全面地捕捉事件的特徵，這對於複雜事件的預測（如暴動或流行病擴散）尤為重要。我認為這種方法可以進一步應用於台灣的公共衛生領域，例如在登革熱疫情監控中，結合氣象數據與社群媒體訊息，提升預測的準確性與時效性。

演講的最後部分聚焦於「Epidemics Modeling」，呂教授提到利用社交媒體數據進行即時流行病監控與個人健康狀況挖掘。這部分內容讓我特別感興趣，因為台灣近年來在公共衛生領域面臨許多挑戰，例如新冠疫情的應對。呂教授提到，社交媒體上用戶分享的健康相關訊息（如流感症狀）可以作為重要的數據來源，但挑戰在於真實的疾病接觸網絡與擴散過程往往難以觀測。他提出的解決方案是將每個地點視為一個任務，利用多模型結合正則化特徵權重，並通過分類器整合靜態與動態特徵。這套方法讓我對多任務學習（Multi-task Learning）有了新的認識，也讓我思考如何將這項技術應用於台灣的醫療數據分析中。

整體而言，呂昌田教授的演講不僅展示了人工智慧與社交媒體數據在事件預測中的強大潛力，也讓我深刻體會到跨領域技術整合的重要性。作為一名資訊工程碩士生，我希望未來能在相關領域進行更深入的研究，特別是在數據預處理、模型融合與多任務學習方面，為台灣的公共安全與社會治理貢獻一份力量。

參考資料：

L. Zhao, J. Chen, F. Chen, W. Wang, C. -T. Lu and N. Ramakrishnan, "*SimNest: Social Media Nested Epidemic Simulation via Online Semi-Supervised Deep Learning*," 2015 IEEE International Conference on Data Mining, Atlantic City, NJ, USA, 2015, pp. 639-648, doi: 10.1109/ICDM.2015.39.

L. Alkulaib, L. Zhang, Y. Sun and C. -T. Lu, "*Twitter Bot Identification: An Anomaly Detection Approach*," 2022 IEEE International Conference on Big Data (Big Data), Osaka, Japan, 2022, pp. 3577-3585, doi: 10.1109/BigData55660.2022.

S. Mitra et al., "*Towards Inclusive Cybersecurity: Protecting the Vulnerable with Social Cyber Vulnerability Metrics*," 2024 IEEE 6th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA), Washington, DC, USA, 2024, pp. 442-445, doi: 10.1109/TPS-ISA62245.2024.00059.