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

美國維吉尼亞理工大學資工系 呂昌田教授

報告日期：2025年5月13日

## 關鍵字

數據預測、社會事件、多語言分析、動態模型、事件地點、歷史數據

## 心得報告

這場演講介紹了一個預測全球抗議活動與疾病爆發的研究項目，並著重說明了他們所面對的挑戰、所採用的解決方法以及背後使用的技術。這項研究的核心目標是結合 AI 模型與社群媒體（如 Twitter）資料，自動預測未來特定地點可能發生的社會事件，例如抗議或疫情，並以無需人為介入的方式將預測結果自動寄出，供獨立機構如 MITRE 評估其準確性與時效性。

該項目中最主要的挑戰之一在於社群媒體本身的雜訊性。以 Twitter 為例，推文內容常常短小、非正式、用語多樣化，加上資訊流變動快速，使得有效地擷取與預測目標事件變得困難。此外，不同地區之間的語言、文化差異也會影響事件的表現形式與用詞，導致預測模型無法輕易泛化。針對這些問題，團隊發展了「動態查詢擴展」技術來進行內容過濾與訊息分類，並引入多任務學習（Multitask Learning）來同時處理不同地點與事件類型的預測。

為了進一步提升預測效能，研究團隊也利用資料降維與推文聚類（clustering）方式來減少模型的複雜度與計算成本。尤其是在資料稀疏的小城市地區，他們採用動態模型產生（dynamic model generation）方法來根據特定地點調整模型架構，讓系統更能因地制宜地進行預測。此外，預測評估使用了包括預測品質、提前時間（lead time）、召回率（recall）、準確率（precision）與機率等五項指標，進一步量化模型效能。

這場演講不僅展現了 AI 在社會事件預測上的實務應用，也讓我對如何處理高雜訊、動態變動的大數據有了更深刻的理解。從技術選擇到實作策略，這個案例說明了結合社群媒體資料與機器學習模型，確實可以為公共安全、國際政策與疾病控制等領域提供早期預警與決策參考。未來若能將這類系統部署在更多語言與地理區域，相信能對全球應變能力大有助益。

## 參考文獻

1. Borge-Holthoefer, J., et al. (2016). "Content and Network Dynamics Behind Egyptian Social Media Activism." Scientific Reports, 6, 31956.

2. Conneau, A., et al. (2020). "Unsupervised Cross-lingual Representation Learning at Scale." Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 8440–8451.

3. Batty, M. (2013). "Big Data, Smart Cities and City Planning." Environment and Planning B: Urban Analytics and City Science, 40(4), 191–193.

4. Ramakrishnan, N., et al. (2014). "‘Beating the News’ with EMBERS: Forecasting Civil Unrest Using Open Source Indicators." Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1799–1808.