本文章已註冊DOI數位物件識別碼

airiti Library

運用機器學習理論建構崩塌地潛勢－以臺東縣太麻里溪流域為例

Establish a Landslide Potential Prediction Model with Machine Learning Method - A Case Study of Taitung Taimali Watershed

doi:10.6857/CDMT.202112.0017

臺灣災害管理研討會, 2021, 2021

Conference for Disaster Management in Taiwan, 2021, 2021

作者/Author： 余軒誼(Yu,Hsuan-Yi);許勝峯(Hsu,Sheng-Feng);蔡元芳(Tsai,Yuan-Fang)

頁數/Page： 215-222

出版日期/Publication Date：2021/12

引用本篇文獻時，請提供DOI資訊，並透過DOI永久網址取得最正確的書目資訊。

To cite this Article, please include the DOI name in your reference data.

請使用本篇文獻DOI永久網址進行連結:

To link to this Article:

http://dx.doi.org/10.6857/CDMT.202112.0017

DOI Enhanced

DOI是數位物件識別碼（Digital Object Identifier, DOI）的簡稱，
是這篇文章在網路上的唯一識別碼，
用於永久連結及引用該篇文章。

若想得知更多DOI使用資訊，

請參考 http://doi.airiti.com

For more information,

Please see: http://doi.airiti.com

請往下捲動至下一頁，開始閱讀本篇文獻

PLEASE SCROLL DOWN FOR ARTICLE

最完整的學術整合查詢平台—華藝線上圖書館 ● http://www.airitilibrary.com ● Find more academic articles on Airiti Library

airiti

運用機器學習理論建構崩塌地潛勢-以臺東縣太麻里溪流域為例

Establish a Landslide Potential Prediction Model with Machine Learning Method – A Case Study of Taitung Taimali Watershed

余軒誼 [1] Yu,Hsuan-Yi
瑞竣科技股份有限公司 專案專員

許勝峯 [2]Hsu,Sheng-Feng
國立臺北教育大學社會與區域發展學系 碩士生

蔡元芳 [3] Tsai,Yuan-Fang
國立臺北教育大學社會與區域發展學系 教授

摘要

近年隨氣候變遷影響下極端降雨事件頻傳，臺灣每年春夏季面對梅雨與颱風所帶來的集中性豪大雨侵襲，常常造成山區發生嚴重的坡地災害。2009 年的莫拉克颱風過後全臺總共新增 39,492 公頃的崩塌地，其中太麻里溪流域範圍為當時影響較為嚴重的流域，流域內山坡地多處發生自然崩塌，而多起的土石崩塌造成了許多公共建設的毀損及民眾生命財產的損失。崩塌地發生因素眾多且複雜，為有效執行災害預防，本研究運用機器學習理論建構山崩潛勢圖，並選定七個因子包含高程、坡度、坡向、曲率、地質、距河流遠近及颱風事件累積雨量。本研究以流域近年最大三場颱風事件，並採用支持向量機（SVM)、類神經網路（ANN)與極限梯度提升（XGBoost)三種演算法進行訓練，運用混淆矩陣計算效能指標。在 ROC 曲線面積中三種演算法的值分別為 0.80、0.77 及 0.88 皆具有良好的分類效能，又以 XGBoost 的值最高。最後本研究以天兔颱風事件建構流域崩塌地潛勢，而各演算法模型總體準確率分別為支持向量機 71.9%、類神經網路 67.1%及極限梯度提升 70.4%。其中極限梯度提升模式建構的崩塌地潛勢圖中，高潛勢地區與實際崩塌區位趨勢較為符合。未來若可利用颱風預報雨量圖層，即可根據此三種模型建構崩塌地潛勢圖，提早掌握太麻里溪流域可能崩塌區域，以提早做災前預備，減少災損。

關鍵詞：崩塌潛勢圖、機器學習、監督式學習

一、 前言

隨近年氣候變遷及極端氣候影響，我國春夏季受梅雨季及颱風侵襲所帶來之豪大雨衝擊致使山區多處發生坡地災害。2009 年 8 月莫拉克颱風挾帶強風及豪雨肆虐，造成我國人員、山區重大衝擊。依據行政院農委會水保局推估，莫拉克颱風過後共新增 39,492

公頃的崩塌地，也造成河川中上游產生 16 處堰塞湖。另外莫拉克颱風更造成臺東縣太麻里溪流域發生多起土石流及崩塌地災害，導致多處流域周邊重要設施損毀。

此外，根據國家災害防救科技中心出版的臺灣氣候變遷科學報告指出，臺灣在未來降雨型態將呈現短延時強降雨型態發展，且整體氣候呈現極端的極度乾旱與潮濕狀態。而我國山區因坡陡流急，且坡面長時間受颱風及季風梅雨影響，致使其容易風化與侵蝕而產生大量崩塌地，且崩塌地之土砂也成為土石流及山崩等致災因素之一。

故為能使我國有效改善流域崩塌地所造成之災損，本研究將透過機器學習之分析預測來應用，並透過支持向量機（SVM）、類神經網路（ANN）及極限梯度提升（XGBoost）三種演算法來建立最適合預測流域崩塌地潛勢之模型，並藉此預測流域內容易發生之崩塌地的地點及機率，以此提供災前預防準備。

二、 研究方法

1、 研究流程與研究範圍

本研究選定臺東縣太麻里溪流域為研究範圍（詳見圖 1），因太麻里溪流域於莫拉克颱風時受重創，造成流域內山坡地多起大型崩塌，另因崩塌地發生範圍多位於森林用地，其受人為因子干擾較少，較適於崩塌地潛勢因子分析。

本研究流程將先行蒐集研究範圍內之環境資料及歷年新增崩踏地位置，並參考文獻選定崩塌因子分為地文、地質、區位及水文因子，並過地理資訊系統（GIS）萃取流域空間數值資料，再進行機器學習模型建置，詳可見圖 1。

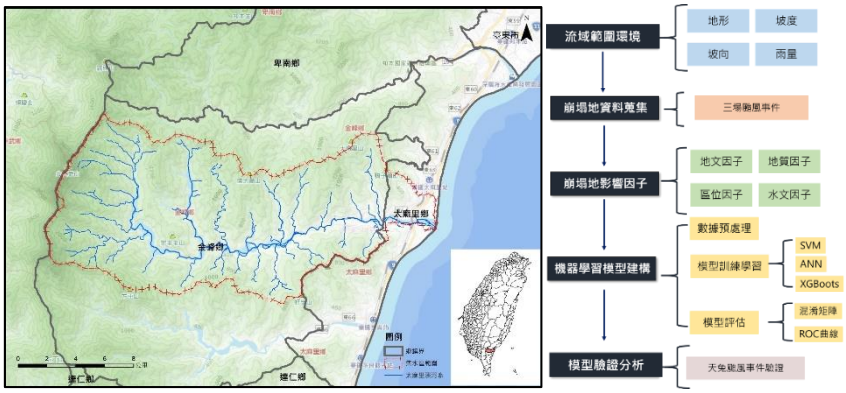

圖 1、研究範圍與研究流程圖

2、 崩塌地資料及影響因子蒐集

地文相關因子透過 GIS 計算太麻里溪流域之高程、坡度與曲率，地質則由中央地質調查所提供。水系相關因子以太麻里河系為主，透過 GIS 環域分析依距離分四級，共分為距離河流 100 公尺內、200 公尺內、300 公尺內與 300 公尺外。以此為距河遠近因子設定。降雨相關因子透過鄰近太麻里溪流域周邊雨量測站資料來進行克

利金法分析，並假設以當年颱風事件內累積雨量作為崩塌地發生之雨量資料。

3、 克利金（Kriging）法雨量推估

降雨為造成崩塌發生的重要誘發因子，但受限研究範圍流域較大，無法透過單一雨量站資料進行克利金降雨量推估，故將依據各颱風事件蒐集離太麻里溪流域最近的雨量測觀測站資料來進行克利金分析，詳見圖 2。為配合不同時期崩塌事件資料，本研究選用三場颱風事件累積雨量，分別為莫拉克颱風（2009）909.5mm、南瑪都颱風（2011）515mm、天兔颱風（2013）514mm，來進行分析。

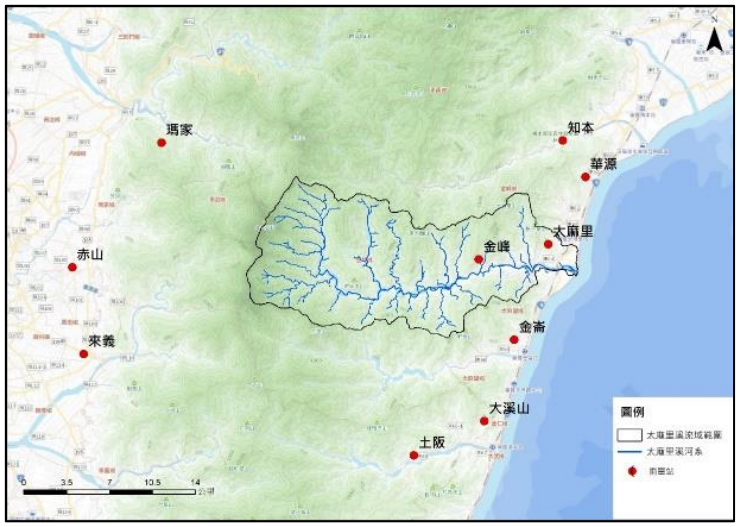

圖 2、太麻里溪流域周為雨量測站分布位置

4、 機器學習模型建構

本研究使用 python 程式語言進行機器學習模型建構，此為目前數據科學中較新也是最受歡迎的程式語言之一，本身含有大量附加函示庫可供使用，研究中主要使用的為 Scikit-learn 與 TennsorFlow 數值計算函示庫。而機器學習模型建構具有主要四個部分，其步驟分別為資料預處理、模型訓練、模型評估與結果預測。

三、 研究結果

1、流域分析單元劃設

本研究將太麻里溪流域製成 20 公尺×20 公尺網格圖層（詳見圖 3），共分割為 538806 個網格，再將因子以 Spatial join 讓每個網格擁有各因子數值資料。

而再進行網格分類判別其崩塌與非崩塌之辦法為，該網格崩塌面積達一半以上，即 200 平方公尺即判定為發生崩塌網格，詳見圖 4。

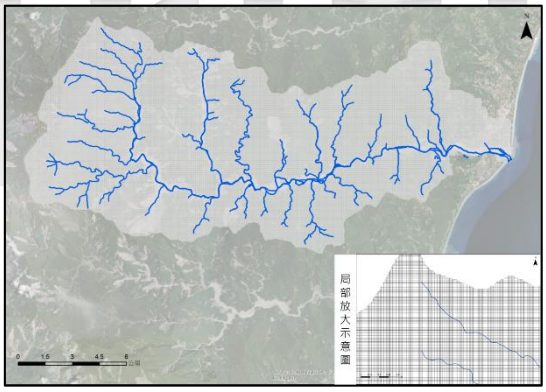

圖 3、太麻里溪流域網格示意圖

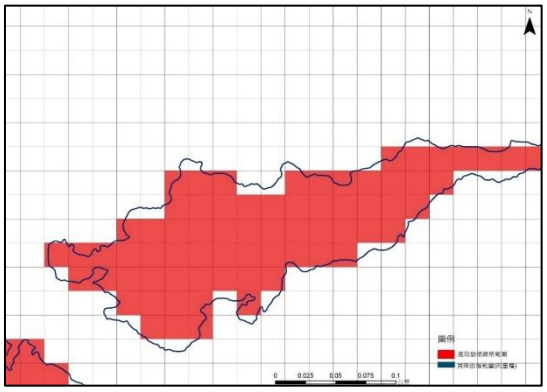

圖 4、發生崩塌網格選取示意圖

2、崩塌地潛勢因子選取

本研究崩塌地潛勢因子依先前相關文獻回顧選定為高程、坡度、坡向、地質、曲率、距河遠近及以颱風事件累積雨量作為雨量因子進行分析。

3、研究流域資料預處理

研究流域範圍網格資料共計 538806 個網格，其中三次颱風事件發生新增崩塌網格數共計 94805 個，未發生崩塌網格共 1521613 個，詳可見表 1。但因新增崩塌網格樣本僅佔未崩塌網格樣本 6.23%，有嚴重不平衡狀況，且因流域網格資料量過大，未必免上述因素造成分析結果誤差即訓練時間過長，本研究將透過下採樣方式進行處理，將未崩塌網格樣本以亂數隨機不重複選曲與崩塌網格樣本數量比例變為 1:1 平衡狀態，詳可見圖 5。

表 1、流域範圍內新增崩塌與未崩塌網格表

事件	新增崩塌網格	未崩塌網格
莫拉克颱風（2009）	59,326	479,480
南瑪都（2011）	21,524	517,282
天兔颱風（2013）	13,955	524,851

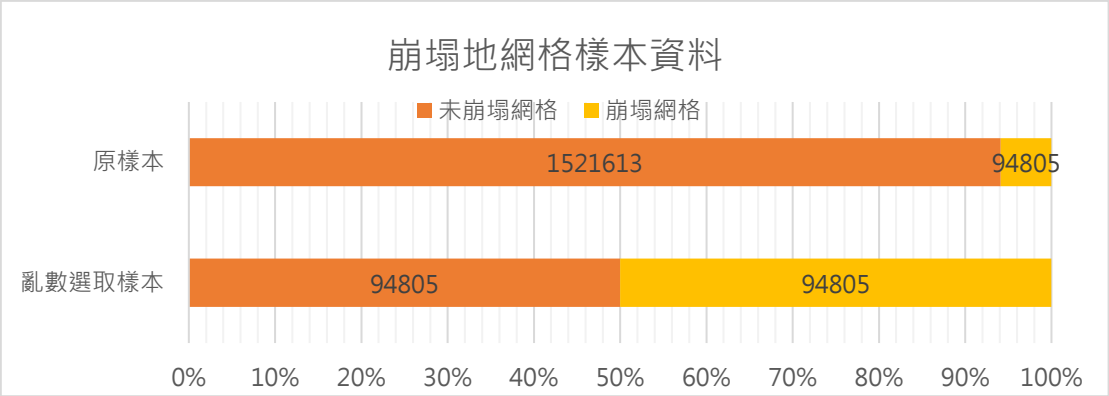

圖 5、流域範圍網格樣本資料預處理

airiti

4、機器學習模型訓練與評估

（1）三種模型訓練時間

本研究進行機器學習模型訓練之時間分別為 SVM（9 時 24 分 39 秒）、ANN（23 分 23 秒）、XGBoost（37 秒）。

（2）混淆矩陣

模型進行演算完後，透過混淆矩陣進行效能評估計算。其中包含真陽性（TP）、假陽性（FP）、假陰性（FN）、真陰性（TN），可用來比較預測結果與真實結果的類別。另可計算總準確率、精確率、召回率及 F1 分數。詳見表 2。

表 2、崩塌地預測混淆矩陣說明表

預測分類				
實際分類		未崩塌網格	崩塌網格	
	未崩塌網格	TN	FP	N:未崩塌網格
	崩塌網格	FN	TP	P:崩塌網格

準確率(Accuracy) = (TP+TN) / (TP+TN+FP+FN)

精確率(Precision) = TP / (TP+FP)

召回率(Recall) = TP / (TP+FN)

$$\text{F1 分數} = 2\frac{\text{PRE}\times\text{REC}}{\text{PRE}\times\text{REC}}$$

圖 6 為各演算法測試集之模型混淆矩陣圖，類別 1 為發生崩塌表格，類別 2 為未發生崩塌網格。以 XGBoost 模型舉例說明，測試集樣本共計 37922 個網格樣本；預測崩塌網格樣本共計 21483 個，其中 16458 為正確預測；預測未崩塌網格樣本為 16439 個，其中 14003 個為正確預測未崩塌網格樣本。XGBoost 精確率為 0.77；召回率為 0.87；模型總體準確率為 0.80；F1 分數則為 0.82，此為運用精確率與找回率計算得知。詳細演算法之混淆矩陣各指標值請見表 3。

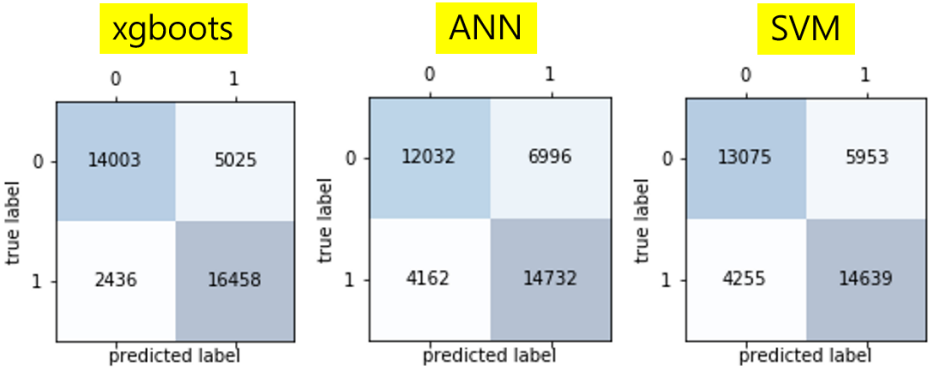

圖 6、各演算法（測試集）模型混淆矩陣圖

表 3、各演算法（測試集）模型指標值

模型演算法	XGBoots	ANN	SVM
準確率	0.80%	0.71%	0.73%
精確率	0.77%	0.68%	0.71%
召回率	0.87%	0.78%	0.78%
F1 分數	0.82%	0.73%	0.74%

（3）接收操作特徵曲線

接受操作特徵曲線（Receiver Operating Characteristic Curve）又稱 ROC 曲線。其可透過 ROC 曲線的曲下面積（Area Ubder the Curve,AUC）來評估預測效能。AUC 值會界於 0~1 之間，且數值越大越好，若數值等於 0.5 則為隨機猜測，沒有預測價值。圖 7 將呈現各演算法模型 ROC 曲線圖，各模型 AUC 值跟別為 ANN（0.77）、XGBoost（0.88）、SVM（0.80），三種皆大於 0.7 具有較佳的判別能力。

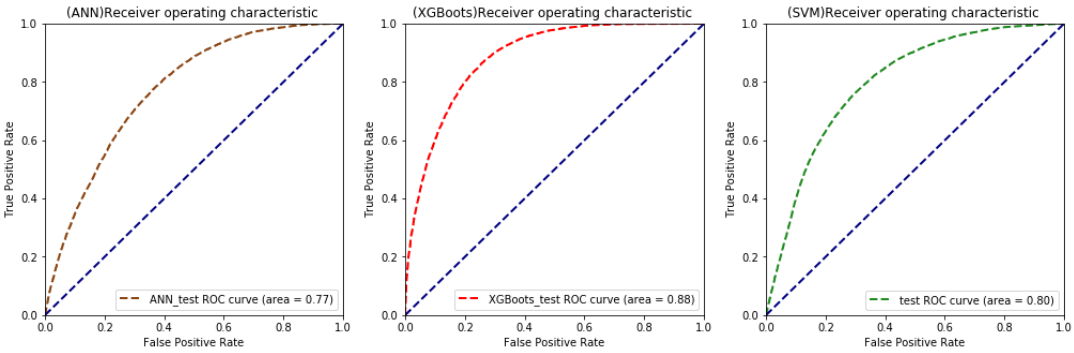

圖 7、各演算法（測試集）模型接受操作特徵曲線

5、模型驗證分析

前述經過模型訓練及評估後，將透過 2013 年天兔颱風事件進行模擬驗證分析。本次預測之選用因子及排序皆與訓練模型一樣，僅累積雨量改為以天兔颱風事件累積雨量資料。而本次驗證將各演算模型預測結果值分為三等級，詳可見表 4。

表 4、崩塌潛勢等級劃分表

崩塌分類	未崩塌網格	崩塌網格	
崩塌潛勢分類	低潛勢	中潛勢	高潛勢
預測結果值	0-0.5%	0.5-0.75%	0.75%以上

本次模型驗證之各模型預測結果可見表 5 及圖 8，由該預測結果可得知 XGBoost 及 SVM 兩種演算法預測模式總體準確率較高，皆高於 0.7 以上，與前述模型訓練時

airiti

的混淆矩陣及 ROC 曲線呈現結果大致相符。

表 5、各演算法預測結果表

模型演算法	ANN	XGBoost	SVM
總體準確率	67.1%	70.4%	71.9%
中、高潛勢網格	8,539	12,602	8,981
佔總崩塌網格率	61.1%	90.3%	64.3%

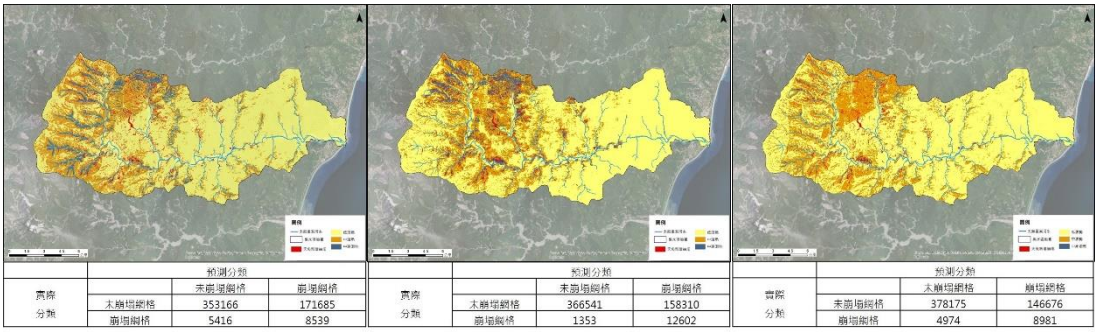

圖 8、各演算法模型崩塌潛勢圖（由左而右依序為 ANN、XGBoost、SVM）

四、 結論與建議

1、結論

（1）建置太麻里溪流域內崩塌地空間資料庫運用 GIS 系統作空間分析，從中萃取崩塌地相關因子，並運用克利金推估三場颱風事件累積雨量，作為機器學習模型建立所需數據資料。

（2）本次選用三種演算法 SVM、ANN、XGBoost 於 ROC 曲線中可見其各 AUC 值分別為 SVM（0.8）、ANN（0.77）、XGBoost（0.88），皆大於 0.7，具有較好的分類效能及預測價值。

（3）SVM 與 XGBoost 的總體準確率皆高於 0.7 以上，但 XGBoost 中高潛勢區與實際崩塌網格趨勢較相符，且 90.3%崩塌五格位於中高潛勢區中，另在時間訓練也明顯較另外兩種演算法快速。

2、建議

（1）崩塌地發生原因複雜，可在未來新增其他因子，如土壤分佈或植生覆蓋情形。

（2）本研究受限無法抓取崩塌地發生之最大降雨強度及發生時間點，而僅能使用颱風事件累積雨量作，若未來能成功抓取上述兩點，對於分析結果會更具可信度。

（3）若能未來能新增更多颱風作為訓練集，可提升整體模型效能。

（4）研究目前為針對太麻里溪流域作為研究區域，未來可在針對其他流域蒐集相關因子後並採用此三種演算法進行崩塌潛勢模型驗證。

參考文獻

Chakraborty, T.,Alam, S.&Islam, D.(2019). ***Usceptibility Mapping Using Xgboost Model In Chittagong District***, Bangladesh. International Conference On Disaster Risk Management,431-434.

Corters, C.,Vapnik, V. (1995).***Support-vector networks***. Machine Learning, 20(3), 273–297.

Kumar, D., Thakur, M., Dubey, C.S.,&Shukla, D.P. (2017). ***Landslide susceptibility mapping & prediction using Support Vector Machine for Mandakini River Basin***, Garhwal Himalaya, India. Geomorphology(295),115-125.

Pourghasemi, H.R.,Jirandeh, A.G.,Pradhan, B.(2013).***Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province***, Iran. J Earth Syst Sci 122, 349–369.

水土保持手冊(2006)。行政院農業委員會水土保持局。

黃雅喬、林國峰、張明瑞、何瑞益(2016)。***支援向量機於降雨引致崩塌潛勢分析之研究***。中國土木水利工程學刊:28 卷 1 期 P57~66。

蔡光榮、謝正倫、周士傑、曾皇銘、陳昆廷(2010)。***類神經網路分析法應用於高雄壽山坡地社區邊坡崩塌潛感區之劃定***。臺灣鑛業，第 62 卷第 3 期，35-42。

謝依霖(2016)。***臺東太麻里集水區崩塌地因子探討***。國立屏東科技大學水土保持學系，碩士論文。