2023-02-20

Extreme Rainfall Events Classification Using Machine Learning for Kikuletwa River Floods

Mdegela, Lawrence

Preprints

Provided with love from The Nelson Mandela African Institution of Science and Technology

Article

Extreme Rainfall Events Classification Using Machine Learning for Kikuletwa River Floods

Lawrence Mdegela [1,4]*, **Esteban Municio** [2], **Yorick De Bock** [1], **Edith Talina Luhanga** [3], **Judith Leo** [4] **and Erik Mannens** [1]

1. University of Antwerp - imec; {lawrencenehemiah.mdegela,Yorick.DeBock, erik.mannens}@uantwerpen.be
2. i2CAT Foundation; esteban.municio@i2cat.net
3. Carnegie Mellon University Africa; eluhanga@andrew.cmu.edu
4. The Nelson Mandela African Institution of Science and Technology; judith.leo@nm-aist.ac.tz
* Correspondence: lawrencenehemiah.mdegela@uantwerpen.be; Tel.:+32494594736

Abstract: Advancements in Machine Learning techniques, availability of more data-sets, and increased computing power have enabled a significant growth in a number research areas. Predicting, detecting and classifying complex events in earth systems which by nature are difficult to model is one of such areas. In this work, we investigate the application of different machine learning techniques for detecting and classifying extreme rainfall events in a sub-catchment within Pangani River Basin, found in Northern Tanzania. Identification and classification of extreme rainfall event is a preliminary crucial task towards success in predicting rainfall-induced river floods. To identify a rain condition in the selected sub-catchment, we use data from five weather stations which have been labeled for the whole sub-catchment. In order to assess which Machine Learning technique suits better for rainfall classification, we apply five different algorithms in a historical dataset for the period of 1979 to 2014. We evaluate the performance of the models in terms of precision and recall, reporting Random Forest and XGBoost as the ones with best overall performance. However, since the class distribution is imbalanced, the generic Multi-layer Perceptron performs best when identifying the heavy rainfall events, which are eventually the main cause of rainfall-induced river floods in the Pangani River Basin.

Keywords: Heavy rainfall; River floods; Machine learning ;

1. Introduction

Rainfall-induced river floods are among Earth's most common and most catastrophic natural hazards [1]. Worldwide, flash floods account for more than 5000 deaths annually with a mortality rate more than 4 times greater than other types of flooding [2], and subsequently, their social, economic, and environmental impacts are significant. According to the Tanzania Meteorological Agency, in the last decade, the northern part of the country has experienced its heaviest rainfall accompanied by strong winds, causing the most severe floods of the last 50 years [3]. It is without a doubt that, with the changing climate, such events are likely to become more frequent, not only in Tanzania as evidenced in several reports [1], [2], [3], but across the globe. The effects of floods are notably severe in developing or low-income countries like Tanzania because of their vulnerability to the occurrence of these phenomena. The vulnerability is partly due to limited human capacity and limited resources invested in dealing with the problem [4].

---

1. Burundi and Tanzania – Floods Leave Homes Destroyed, Hundreds Displaced. https://floodlist.com/africa/burundi-tanzania-floods-late-february-2021
2. Tanzania – Severe Flooding in Mtwara Region After Torrential Rainfall. https://floodlist.com/africa/tanzania-flood-mtwara-january-2021.
3. Tanzania – 12 Killed in Dar Es Salaam Flash Floods. https://floodlist.com/africa/tanzania-daressalaam-floods-october-2020.

Understanding the trends and key patterns in the occurrence of rainfall events, is an important step towards better flood risk management plans that will help in designing more accurate early warning systems. [5].

Machine Learning (ML) presents the ability to identify the hidden patterns and trends in historical climate data [6], and may be used to classify and predict the key rainfall events that are associated with the occurrence of floods. The potential of machine learning techniques to improve the classification and prediction of extreme rainfall events has been demonstrated by several studies in the past. The techniques provide valuable insights into the spatial and temporal patterns of extreme rainfall events and their impacts on flood generation, water resources management, and climate change impact assessment. Authors in a study [7] proposed an event-based flood classification method to study the global river flood generation processes. The approach is based on a combination of unsupervised and supervised machine learning methods that can provide event-based information for better understanding of flood generation processes. Another machine learning-based down-scaling approach is demonstrated in Pham *et al.* [8] where a combination of random forest and least square support vector regression to improve the accuracy of extreme rainfall predictions at a local scale is demonstrated. The inspiration of the method used here is that it provides valuable insights into the extreme rainfall events and their spatial and temporal characteristics, which are useful for water resource management and flood risk assessment.

Similarly, [9] developed a machine learning-based classification method to categorize extreme precipitation events over Northern Italy. The study employed the k-means clustering technique to identify distinct clusters of extreme rainfall events and used decision trees to develop a classification scheme. Despite the fact that most of these studies were conducted in developed counties where there is advancement in both technology and human resource, the results show a big potential to be used as models for similar studies in other developing regions as Tanzania where the impacts of climate change are expected to be more severe [10].

Furthermore, [11] presented a study that used three different machine learning algorithms (XGBoost, LightGBM, and CatBoost) to forecast daily stream-flow in a mountainous catchment. The study compared the performance of the three algorithms and showed that machine learning can provide accurate stream-flow forecasts, which are valuable for water management and flood prediction. An analysis of physical causes of extreme precipitation [12] can also be used to identify key climatic variables that drive extreme precipitation events and machine learning based approaches can be applied to predict the occurrence of extreme precipitation.

We apply five machine learning techniques namely Random Forest, eXtreme Gradient Boost, Support Vector Machine, k-Nearest Neighbors and Multilayer Perceptron to identify and classify rainfall events in the Karanga-Weruweru-Kikafu sub-catchment, located within the Pangani River Basin, Tanzania. Random Forest is preferred for its robustness to large and noisy data sets [13] and its ability to handle imbalanced data sets [14,15]. XGBoost is computationally efficient [16] and it can also perform better on imbalanced data sets [17]. SVM is suitable for high dimensional input space and modeling complex, non-linear relationships between inputs and outputs [18,19]. k-Nearest neighbors although considered one of the simplest machine learning algorithms, has been successful in a number of applications, from recognition of handwritten texts [20] to satellite image scenes [21] and mostly success in classification problems with irregular decision boundary. MLP as a feed forward neural network has also shown success in classification problems, including extreme natural events like droughts [22].

We compare these techniques and discuss the suitability of each in successfully classifying rainfall events. To train these models, a historical labeled data-set from Pangani Water Board (PWB) and Tanzania Meteorological Agency (TMA) collected from five stations located across the Karanga-Weruweru-Kikafu sub-catchment was used. The nature of the dataset gives us an imbalanced multi-class classification problem. There are three categories in the target class (Heavy, Light and No-rain). Of these, heavy rainfall is the

smallest making up just 0.32% of the whole data set. The distribution is highly skewed towards the majority class, in this case, light rain, which make up 83.22% of the whole data set leaving 16.46% to no-rain class. This simply means for every single example of heavy rainfall event, there are 51 examples of no-rain and 260 examples of light rain.

2. Materials and Methods

The study area under consideration is situated in the Northern part of Tanzania in the south of Kilimanjaro region. The Karanga-Weruweru-Kikafu (KWK) sub-catchment (Figure 1) and the villages along the Kikuletwa river are intensely affected by flash river floods from heavy rainfall. The aim of this work was to classify rainfall intensity (heavy, light, none) based on 6 weather parameters namely temperature (Minimum and Maximum), relative humidity, precipitation, solar, and wind speed. Data records for the study covering from 1979 to 2014 was provided by Tanzania Meteorological Agency [4], and the Pangani Basin Water Board(PBWB) [5]. The data contains daily weather data for 35 years, with 7 parameters namely Maximum and Minimum temperatures in centigrade(C), precipitation in millilitres(mm), wind speed in meter per second(m/s), relative humidity expressed as percentage, and solar irradiance in mega-joules per square metre(MJ/m^2) and rain category. It is worthy noting that, the data was consolidated from two main sources, the ground gauges and the satellite estimates. The basis of the consolidation and to compliment the ground gauge data by the satellite estimates was based on several previous studies that were done to check the validity of the satellite estimates over the region. Three studies, [23], [24], and [25], investigated the spatio-temporal characteristics and accuracy of satellite-derived rainfall estimates in Tanzania, in comparison to ground-based measurements. The studies revealed a positive correlation between the two data sources, indicating the potential of satellite-based rainfall estimates as a useful complement to ground-based measurements, especially in areas with complex topography and limited ground-based measurement stations. Nevertheless, the satellite estimates exhibited a tendency to overestimate the ground-based measurements, and their accuracy varied in different locations. The findings suggest that the use of satellite-based rainfall estimates can enhance rainfall monitoring and prediction in regions where traditional measurement methods are sparse or challenging to implement, albeit with the need for continual improvements in their accuracy and uncertainty estimation.
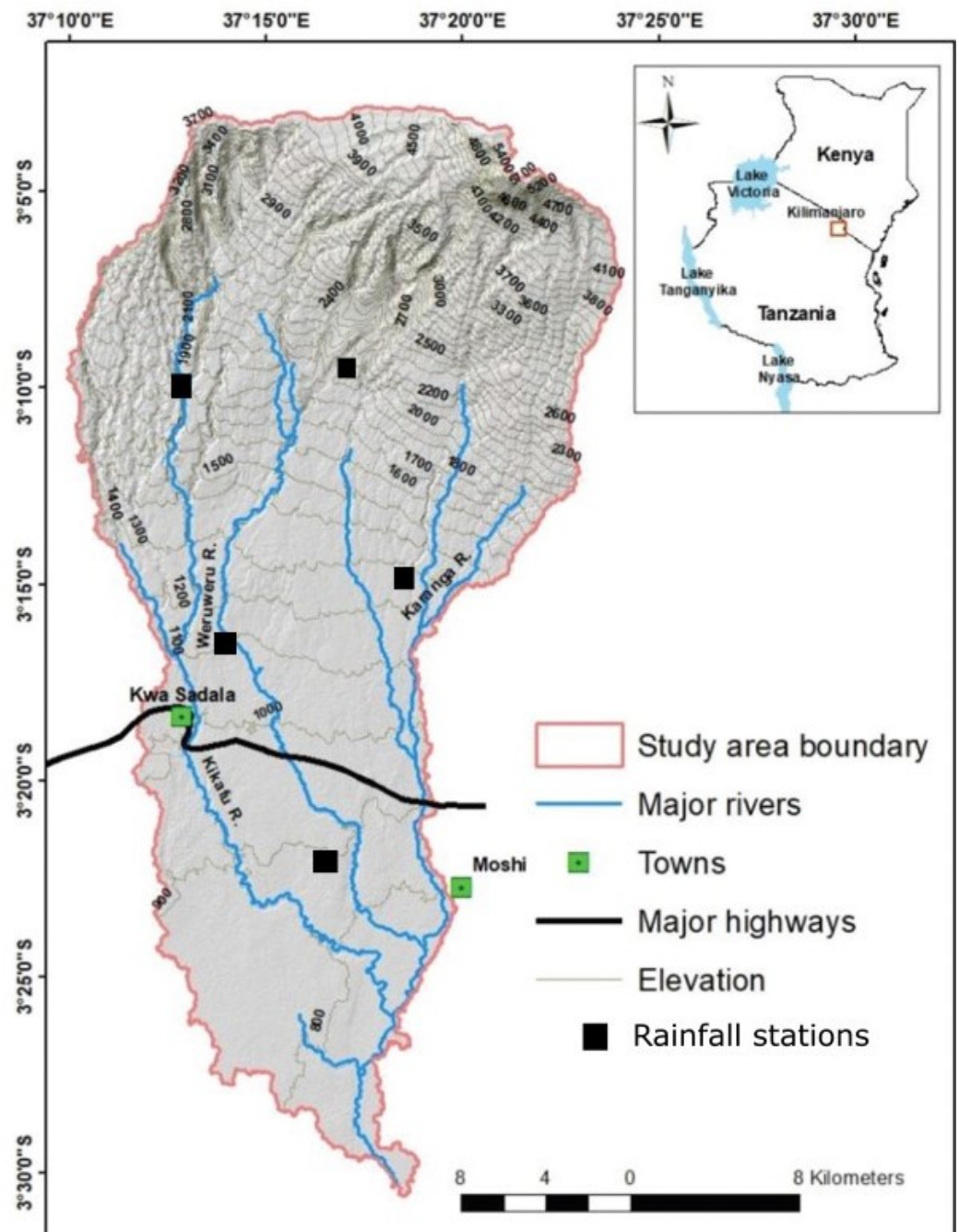

Figure 1. Karanga-Weruweru-Kikafu(KWK) sub-catchment.

2.1. Data preparation

The data from each station were checked for missing values before being merged into a single data set. Simple line plots were used to check whether all stations had similar patterns in the features and to identify any outliers. Simple statistical analysis was done on the numerical features and summarized in Table 1

Data was split into train and test set in the ratio of 80% to 20%, in order to even out the distribution as there is an imbalance in the target class distribution. In total there were 6 features and 1 target class. The features are maximum temperature, minimum

temperature, precipitation, wind, relative humidity and solar radiance, while the target is rainfall category. Further feature engineering was done, where all the object type columns were encoded to numeric type. Pivoting was also done to put the data in a format that is convenient for model training. Training data was normalized using *MinMaxScaler* from sklearn library.

Table 1. Descriptive statistics of weather variables used in training.

Variable	Count	Mean	Std. Dev.	Min	25th Percentile	75th Percentile
Max Temperature	10,389	22.71	3.08	13.26	20.53	24.93
Min Temperature	10,389	12.90	1.96	5.71	11.61	14.41
Precipitation	10,389	3.17	6.34	0.00	0.15	3.66
Wind	10,389	2.49	0.55	0.65	2.14	2.87
Relative Humidity	10,389	0.76	0.10	0.32	0.69	0.84
Solar	10,389	16.92	7.23	0.00	11.18	22.19

2.2. Model Building

Two main things were considered during this stage before jumping into the models. First the target class distribution and second the multi-class classification. Our target class was somehow severely imbalanced, with the distribution being highly skewed towards the majority class, light rain(83.22%), followed by no-rain (16.46%),and heavy rainfall, which is our class of interest, is the minority (0.32%). This simply means for every single example of heavy rainfall, we have 51 examples of no-rain and 260 examples of light rain.

Consideration on the part of multi-class classification was to use a multi-class strategy from scikit-learn [6] library known as One-vs-the-rest(OvR). OvR is a heuristic technique of dealing with multi-class problems by fitting one classifier per class. For each classifier, the class is fitted against all the other classes. One of the implementation of OvR is from the sklearn library, which provides a separate *OneVsRestClassifier* class that allows the one-vs-rest strategy to be used with any classifier. A classifier that is inherently for binary classification is just provided to the *OneVsRestClassifier* as an argument.

Each model was then trained, tested and evaluated. Since our problem fall under multi-class imbalanced classification, selecting a metric for evaluation was the most important step in the project. A wrong metric would mean choosing the wrong algorithm, consequently solving a different problem from the one you wan to solve.

2.3. Model Evaluation

Since we are dealing with a highly skewed data-set we chose Precision and Recall as our performance evaluation metric. Precision(Equation 1) is a ratio of the number of true positives divided by the sum of the true positives and false negatives. In other words, it inform about how good a model is at predicting the positive class.

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} \tag{1}$$

Recall (Equation 2) on the other hand is the ratio of the number of true positives divided by the sum of the true positives and the false negatives.

$$Recall = \frac{TruePositives}{TruePositives + Falsenegatives} \tag{2}$$

One important aspect of precision and recall to take note is that, the calculations do not consider the use of the *truenegatives*. The focus is on the correct prediction of the minority class. Precision-recall curve is a plot of the precision on the y-axis and the recall on the

---

[6] https://scikit-learn.org/stable/modules/generated/sklearn.multiclass OneVsRestClassifier.html

x-axis for different thresholds. They give a more informative picture of an algorithm in skewed data-sets as it has also been evident in a number of studies [26], [27]. In that sense we identified our positive class to be H for heavy rainfall and other collectively as negative classes(No rain and Light rain). Nevertheless, precision and recall, are in a trade-off relationship, at some point you may need to optimize one at the expense of the other [28]. Contextually, at some point you would want classifier that is good at minimizing both the false positives and false negatives, meaning, it would make more sense to have a model that is equally good at identifying cases were a false alarm of a heavy rainfall event goes off and when an alarm is not going-off while there is an event coming. In the view of that, we applied another metric called F1_score. F1score is the harmonic mean (Equation 3) of precision and recall and ranges from 0 to 1.

$$F1\_score = \frac{2 * (precision * recall)}{precision + recall} \tag{3}$$
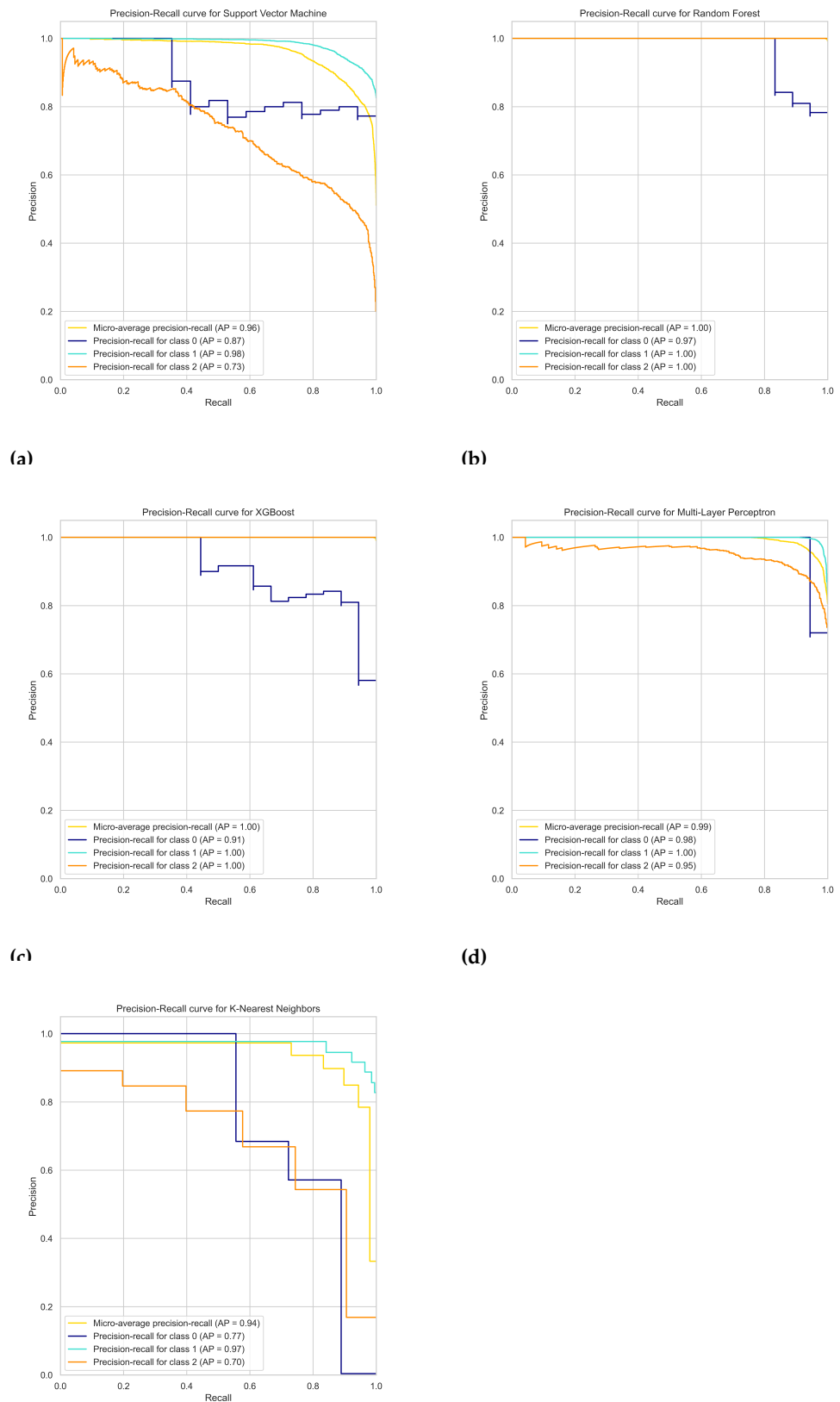
3. Results

In our experiments, five different machine learning algorithms were used to identify and classify extreme rainfall events among three rainfall categories. As stated in the introduction, the ability to identify extreme rainfall events is crucial for predicting rainfall-induced river floods. Results from evaluation show that, overall Random Forest and XGBoost performed better than the rest as we can see from the F-scores summarized in Table 2.

Table 2. Summary of F1_score measures for the models.

Random Forest	XGBoost	Support Vector Machine	KNN	Multi-layer Perceptron
0.998	0.998	0.878	0.898	0.950

Ideally, the scores from F1_score means both XGBoost and Random forest have perfect precision and recall when you give equal importance to both false negatives and false positives. However, that is not the case in our problem. In classifying the heavy rains, which in our case is the minority class, false negatives were the most important. Intuitively, in our context, it is not helpful if we are successful in predicting all data points as negative, that is no heavy rainfall event, instead, we focused on identifying the positive cases, the occurrence of a heavy rainfall event. Coming back to the metrics language this simply means we maximized the *recall*, the ability of our model to find all the relevant cases within a given data-set. This notion is supported by a number of past studies [29], [30], [31], [32]. In the view of that, F-score is not the determinant for the appropriate model to use in this scenario, as we said it being the harmonic mean, it takes into account both the precision and recall. Our main goal is favor the minimization of the false negatives and not to cast equal importance to both the false negatives and false positives. We focused on having a model with high recall which is able to identify most of the heavy rainfall events(true positives), that way saving lives and properties from the consequences that come along with such events. On the other hand of course, that is at the expense of issuing false alarms of heavy rainfall events as though they will happen (false positive) but they won't. Potentially, the associated costs of false positives will be the unnecessary anxiety to the people and at the worst, the costs associated with taking unnecessary precautions. In most cases, the false positives will not be fatal. Therefore, since false negatives will results into fatalities and destruction, we want to have our classification threshold to favor the optimization of recall over precision. This is the point where we turn our attention to the precision-recall curves for more insight. Despite the fact that Random Forest and XGBoost were the general best performers when we put equal weight on both the false negatives and the false positive, it was the generic Multilayer Perceptron that came on top when we focused on the minority

positive class. Multilayer Perceptron (MLP) was 98% in identifying our class of interest as can be seen in the precision-recall curves of Figure 2(d).

(a)

(b)

(c)

(d)

(e)

Figure 2. Precision-Recall Curves for the models.

4. Discussion

To get deeper into the results highlighted above, we see that we based the evaluation on two sets of metrics, the precision-recall (PR) curve and the F1 score, to assess the models' ability to classify the minority class in a data set. The evaluation of algorithms in classifying the minority class in imbalanced data sets is a topic of ongoing research in the field of machine learning. For example, in a study by Batista et al. [33], the authors found that precision-recall curves were more reliable for evaluating imbalanced datasets than other metrics such as ROC curves.

4.1. Precision-Recall curve results analysis

The precision-recall curve is a valuable metric for evaluating algorithms in imbalanced data sets, particularly when the positive class is rare. The PR curve provides a graphical representation of the trade-off between precision and recall, where precision measures the proportion of correct positive predictions among all positive predictions, and recall measures the proportion of correct positive predictions among all actual positive samples.

The micro-average PR curve of each model (Figure 2) summarizes the overall performance of the model in all classes. The micro-average PR curve is computed by treating all the classes as a single binary classification problem. The micro-average PR curve A.P (average precision) score for the Support Vector Machine (SVM), Random Forest, XGBoost, Multi-layer Perceptron (MLP), and K-Nearest Neighbors (KNN) models were 0.96, 1, 1, 0.99, and 0.94, respectively.

Looking at the individual class PR curves, the models achieved high precision-recall performance for classes 1 and 2, indicating a high ability to predict the absence of heavy rain (class 2) and light rain (class 1). However, the models showed lower performance in predicting the occurrence of heavy rain (class 0), which is the minority class in the imbalanced data set.

Among the models, SVM achieved the lowest A.P score of 0.96, and its PR curve for class 0 had the lowest A.P score of 0.87, indicating that the SVM model has the lowest ability to predict the minority class. On the other hand, the Random Forest, XGBoost, and MLP models showed high performance in predicting the minority class, with A.P scores of 1, 1, and 0.99, respectively. The PR curve for class 0 for these models also achieved high A.P scores of 0.97, 0.91, and 0.98, respectively. KNN achieved a moderate A.P score of 0.94, and its PR curve for class 0 had an A.P score of 0.77.

4.2. F1 Score results analysis

The F1 score is a single number that summarizes the harmonic mean of precision and recall. It is another useful metric for evaluating model performance in imbalanced data sets. The F1 scores of the models (Table 2) were as follows: Random Forest (0.998), XGBoost (0.998), SVM (0.878), KNN (0.898), and MLP (0.95).

Comparing the results of the two sets of metrics, the Random Forest and XGBoost models achieved the highest F1 scores of 0.998, indicating their superior overall performance in predicting the occurrence of heavy rain. These models also showed high PR curve performance, particularly for class 0. The MLP model achieved the second-highest F1 score of 0.95, indicating high performance in predicting the minority class. The SVM model had the lowest F1 score of 0.878, consistent with its lower PR curve performance, particularly for class 0.

Overall, the Random Forest and XGBoost models showed the highest performance in predicting heavy rain events, while the SVM model had the lowest performance. The MLP model also showed good performance in predicting the minority class. It is important to note that the imbalanced nature of the data set presented a challenge to all models in predicting the minority class, and thus, the models' performance in this aspect should be carefully considered.

The PR curve analysis revealed that the Random Forest, XGBoost, and MLP models had high precision and recall performance, particularly for the minority class, while the

SVM and KNN models had lower precision and recall performance, especially for the minority class. These results suggest that the Random Forest, XGBoost, and MLP models are more suitable for the prediction of heavy rain events in an imbalanced data set.

On the other hand, the F1 score results revealed that the Random Forest and XGBoost models had the highest F1 scores, indicating their superior overall performance in predicting heavy rain events. The MLP model also showed high F1 score, but the SVM and KNN models had lower F1 scores. These results suggest that the Random Forest, XGBoost, and MLP models are more suitable for predicting heavy rain events in an imbalanced data set based on the F1 score.

It is important to consider that the models' performance might be affected by the choice of evaluation metrics, and it is recommended to use multiple evaluation metrics to assess model performance. In this study, the PR curve and F1 score were used to provide a comprehensive evaluation of the models' performance.

In this view, the Random Forest and XGBoost models showed high performance in predicting heavy rain events in an imbalanced data set, as indicated by both the PR curve and F1 score results. The MLP model also showed good performance in predicting the minority class. The SVM and KNN models had lower performance in predicting heavy rain events, especially for the minority class. These results suggest that the Random Forest, XGBoost, and MLP models are more suitable for predicting heavy rain events in an imbalanced data set.

On the other hand, it's important to note that these results were obtained using default hyper-parameters, and there may be additional improvements that can be made by fine-tuning the models. However, these results still provide valuable insights into the relative performance of different algorithms for predicting rainfall classes in imbalanced data sets. Future studies could also investigate the impact of hyper-parameter tuning on the models' performance in identifying extreme rainfall events.

5. Conclusion

The objective of this study was to evaluate different machine learning techniques for detecting and distinguishing heavy rainfall events in a sub-region of the Pangani River Basin in Northern Tanzania. The study employed five different algorithms to identify heavy rainfall occurrences between 1979 and 2014. The models' performance was assessed using precision-recall metrics and F-score to determine the most suitable method for the task. Based on the evaluation results, Random Forest and XGBoost demonstrated superior overall performance. However, it was observed that the Multi-layer Perceptron (MLP) performed better in identifying heavy rainfall events, which are the leading cause of floods in the Pangani River Basin.

The study's results suggest that MLP, despite being outperformed by other algorithms in overall performance, was the most effective technique for identifying heavy rainfall events, highlighting the significance of precision and recall in detecting the minority class in imbalanced classification. The highly imbalanced class distribution in the dataset makes it challenging to identify heavy rainfall events, making the use of MLP a vital approach in the process.

The findings of this study align with previous research that has emphasized the importance of selecting appropriate performance metrics to evaluate algorithms' effectiveness in detecting rare events. Moreover, the study contributes to the literature by demonstrating that the MLP approach is well-suited for recognizing heavy rainfall events in the Pangani River Basin. The research provides valuable insights into the potential of machine learning algorithms in identifying heavy rainfall events, enabling policymakers to take proactive measures in flood management and control. To the government of Tanzania, the study recommends that the ministry responsible for monitoring flood and water levels in rivers and other water bodies should collect water level data with respect to weather parameters. This will enable the replication of the developed model in other rivers and assist in future studies in similar areas.

Overall, these models have potential applications in various fields that require accurate predictions of rare events, such as climate prediction, disaster management, and risk assessment. However, further research is needed to explore the models' performance in different settings and under different conditions, such as changes in climate patterns and data sources.

Future research could also focus on developing more robust models that can handle highly imbalanced data sets and improving feature engineering techniques to enhance model performance. Additionally, ensemble techniques and meta-learning approaches could be explored to improve the models' generalization and transfer learning abilities. Overall, the study provides insights into the potential of machine learning algorithms in predicting rare events and highlights the need for further research to develop more accurate and robust models.

Author Contributions: Conceptualization, L.M and E.M.; methodology, L.M.; software, L.M.; validation, L.M., E.M. and Y.B.; formal analysis, L.M.; investigation, L.M.; resources, L.M and E.M.; data curation, L.M.; writing—original draft preparation, L.M.; writing—review and editing, E.M, E.L, J.L and Y.B.; visualization, L.M.; supervision, E.L, J.L and E.M.; project administration, n/a.; funding acquisition, n/a. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: At the moment, data supporting reported results can be found on request, it will be publicly available at the end of the main project

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

XGBoost	eXtreme Gradient Boost
KNN	k-Nearest Neighbors
ML	Machine Learning
SVM	Support Vector Machine
MLP	Multi-Layer Perceptron
PBWB	Pangani Basin Water Board
TMA	Tanzania Meteorological Agency
KWK	Karanga-Weruweru-Kikavu
OvR	One-vs-the-Rest
ROC	Receiver Operating Characteristic

References

1. World Health Organization. Floods. Retrieved February 17, 2023, from https://www.who.int/health-topics/floods, 2021.
2. Jonkman, S.N. Global perspectives on loss of human life caused by floods. *Natural hazards* **2005**, *34*, 151–175.
3. Tanzania Meteorological Agency. Annual Technical Report on Meteorology, Hydrology and Climate Services 2020-2021 Update. Retrieved on February 17, 2023, from https://www.meteo.go.tz/uploads/publications/sw1628770614-TMA%20BOOK%202020%20-2021%20UPDATE.pdf, 2021.
4. Kimambo, O.N.; Chikoore, H.; Gumbo, J.R. Understanding the Effects of Changing Weather: A Case of Flash Flood in Morogoro on January 11, 2018. *Advances in Meteorology* **2019**, *2019*, Article ID 8505903. https://doi.org/10.1155/2019/8505903.
5. Nayak, M.A.; Ghosh, S. Prediction of extreme rainfall event using weather pattern recognition and support vector machine classifier. *Theoretical and applied climatology* **2013**, *114*, 583–603.

6.   Parmar, A.; Mistree, K.; Sompura, M. Machine learning techniques for rainfall prediction: A review. In Proceedings of the International Conference on Innovations in information Embedded and Communication Systems, 2017, Vol. 3.

7.   Stein, L.; Pianosi, F.; Woods, R. Event-based classification for global study of river flood generating processes. *Hydrological Processes* **2020**, *34*, 1514–1529.

8.   Pham, Q.B.; Yang, T.C.; Kuo, C.M.; Tseng, H.W.; Yu, P.S. Combing random forest and least square support vector regression for improving extreme rainfall downscaling. *Water* **2019**, *11*, 451.

9.   Grazzini, F.; Craig, G.C.; Keil, C.; Antolini, G.; Pavan, V. Extreme precipitation events over northern Italy. Part I: A systematic classification with machine-learning techniques. *Quarterly Journal of the Royal Meteorological Society* **2020**, *146*, 69–85.

10. Bank, W. Climate Change Knowledge Portal: Tanzania, 2019.

11. Szczepanek, R. Daily Streamflow Forecasting in Mountainous Catchment Using XGBoost, LightGBM and CatBoost. *Hydrology* **2022**, *9*, 226.

12. Davenport, F.V.; Diffenbaugh, N.S. Using machine learning to analyze physical causes of climate change: A case study of US Midwest extreme precipitation. *Geophysical Research Letters* **2021**, *48*, e2021GL093787.

13. Fernández-Delgado, M.; Cernadas, E.; Barro, S.; Amorim, D. Do We Need Hundreds of Classifiers to Solve Real World Classification Problems? *Journal of Machine Learning Research* **2014**, *15*, 3133–3181.

14. Khoshgoftaar, T.M.; Golawala, M.; Hulse, J.V. An Empirical Study of Learning from Imbalanced Data Using Random Forest. In Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence(ICTAI 2007), 2007, Vol. 2, pp. 310–317. https://doi.org/10.1109/ICTAI.2007.46.

15. Liu, X.Y.; Wu, J.; Zhou, Z.H. Exploratory Undersampling for Class-Imbalance Learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **2009**, *39*, 539–550. https://doi.org/10.1109/TSMCB.2008.2007853.

16. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Association for Computing Machinery: New York, NY, USA, 2016; KDD '16, p. 785–794. https://doi.org/10.1145/2939672.2939785.

17. Wang, C.; Deng, C.; Wang, S. Imbalance-XGBoost: leveraging weighted and focal losses for binary label-imbalanced classification with XGBoost. *Pattern Recognition Letters* **2020**, *136*, 190–197. https://doi.org/https://doi.org/10.1016/j.patrec.2020.05.035.

18. Pilario, K.E.; Shafiee, M.; Cao, Y.; Lao, L.; Yang, S.H. A Review of Kernel Methods for Feature Extraction in Nonlinear Process Monitoring. *Processes* **2020**, *8*. https://doi.org/10.3390/pr8010024.

19. He, Y.; Ma, J.; Ye, X. A support vector machine classifier for the prediction of osteosarcoma metastasis with high accuracy. *Int J Mol Med* **2017**, *40*, 1357–1364. https://doi.org/10.3892/ijmm.2017.3126.

20. Chychkarov, Y.; Serhiienko, A.; Syrmamiikh, I.; Kargin, A. Handwritten Digits Recognition Using SVM, KNN, RF and Deep Learning Neural Networks. In Proceedings of the CMIS, 2021.

21. Mcroberts, R. A two-step nearest neighbors algorithm using satellite imagery for predicting forest structure within species composition classes. *Remote Sensing of Environment* **2009**, *113*, 532–545. https://doi.org/10.1016/j.rse.2008.10.001.

22. Azorin-Molina, C.; Ali, Z.; Hussain, I.; Faisal, M.; Nazir, H.M.; Hussain, T.; Shad, M.Y.; Mohamd Shoukry, A.; Hussain Gani, S. Forecasting Drought Using Multilayer Perceptron Artificial Neural Network Model. *Advances in Meteorology* **2017**, *2017*, 5681308. https://doi.org/10.1155/2017/5681308.

23. Dinku, T.; Ceccato, P.; Grover-Kopec, E.; Lemma, M.; Connor, S.J.; Ropelewski, C.F. Validation of satellite rainfall products over East Africa's complex topography. *International Journal of Remote Sensing* **2007**, *28*, 1503–1526. https://doi.org/10.1080/01431160600954688.

24. Hamis, M.M. Validation of Satellite Rainfall Estimates Using Gauge Rainfall Over Tanzania, 2013.

25. Lu, S.; ten Veldhuis, M.C.; van de Giesen, N. Evaluation of four satellite precipitation products over Tanzania. In Proceedings of the EGU General Assembly Conference Abstracts, 2018, EGU General Assembly Conference Abstracts, p. 1403.

26. Cook, J.; Ramadas, V. When to consult precision-recall curves. *The Stata Journal* **2020**, *20*, 131–148, [https://doi.org/10.1177/1536867X20909693]. https://doi.org/10.1177/1536867X20909693.

27. Li, W.; Guo, Q. Plotting receiver operating characteristic and precision–recall curves from presence and background data. *Ecology and Evolution* **2021**, *11*, 10192–10206, [https://onlinelibrary.wiley.com/doi/pdf/10.1002/ece3.7826]. https://doi.org/https://doi.org/10.1002/ece3.7826.

28. Erenel, Z.; Altincay, H. Improving the precision-recall trade-off in undersampling-based binary text categorization using unanimity rule. *Neural Computing and Applications* **2012**, *22*. https://doi.org/10.1007/s00521-012-1056-5.

29. Brabec, J., K.; T., F.; V., M. On Model Evaluation Under Non-constant Class Imbalance. In Proceedings of the Computational Science – ICCS 2020; Springer International Publishing: Cham, 2020; pp. 74–87.

30. He, H.; Garcia, E.A. Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering* **2009**, *21*, 1263–1284. https://doi.org/10.1109/TKDE.2008.239.

31. Kotsiantis, S.; Kanellopoulos, D.; Pintelas, P.; et al. Handling imbalanced datasets: A review. *GESTS international transactions on computer science and engineering* **2006**, *30*, 25–36.

32. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Information Processing & Management* **2009**, *45*, 427–437. https://doi.org/https://doi.org/10.1016/j.ipm.2009.03.002.

33. Batista, G.E.; Prati, R.C.; Monard, M.C. A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explorations Newsletter* **2004**, *6*, 20–29.