

**FACULTY OF INFORMATION SCIENCE AND COMMUNICATIONS**

**DEPARTMENT OF INFORMATION AND COMMUNICATION TECHNOLOGY**

**TO : MR E. TUNG’ANDE**

**FROM : MAYAMIKO KADZANJA**

**REG NUMBER : BSCDS 1320**

**LEVEL : TWO**

**COURSE TITLE : MACHINE LEARNING**

**PROGRAMME : BSCDS**

**SUBJECT : PROJECT INDIVIDUAL REPORT**

**ASSIGNMENT : ASSIGNMENT**

**INTRODUCTION**

In this project, we developed a web-based flood prediction model using machine learning algorithms. The purpose of the model is to provide early warning of floods in Malawi, which will help to prevent property damage and save lives.

**RELATED WORK**

We searched for various machine learning techniques used for flood prediction, including regression analysis, and decision trees. We also reviewed similar studies in literature and the state-of-the-art techniques used for flood prediction.

**DATASET**

We collected and used a dataset containing historical rainfall data, water level data, and weather forecast data from various location in Malawi.

We used various preprocessing techniques such as feature selection, data cleaning, and data normalization to prepare the data for our model.

**METHODOLOGY**

We used a random forest regression model to predict the water levels of a particular location based on the input features. We trained the model on the historical data and evaluated its performance using root mean squared error (RMSE) and the coefficient of determination(R-squared).

We also performed feature importance analysis to identify the most significant predictors of flood risk.

**RESULTS**

Our model achieved a good performance, with an accuracy of 0.92.

Feature importance analysis showed that rainfall intensity, duration, and frequency were the most important predictors of flood risk.

**DISCUSSION**

Our results showed that a random forest regression model can achieve a good performance in predicting flood risk.

The most important predictors identified by our model were consistent with previous research. However, our model may not be generalizable to other regions or time periods and further research is needed to investigate the external validity of our findings.

**CONCLUSION**

In this project, we successfully developed a flood prediction system using machine learning techniques. Our results suggest that rainfall intensity, duration, and frequency are important predictors of flood risk. Our system can help to provide early warning of floods which can help to prevent damage and save life.

**PERSONAL CONTRIBUTION**

In this project, I took part in data cleaning and normalization of the data after it has been collected by my colleagues. The preprocessing of data was necessary to remove the irregularities which was achieved with a python code. The main aim of preprocessing was to remove the missing values or any outliers which could affect the accuracy of the model.

I and one of my colleagues were responsible for creation of the relation database of which the web-based prediction system will be operating with. I achieved this with the help of MySQL.