Software Requirements Specification

For

“Landslide Prediction Using Machine Learning”

Prepared by:

Name

Date:

--/--/----

Table of Contents

1. Introduction

1.1 Purpose

The purpose of this project is to find or precisely predict the landslide susceptibility of landslide prone area using the geological and environment data

1.2 Scope of the Project

Being able to identify the landslide prone areas and their susceptibility or vulnerability index as a random measure on any hour of the day using established Machine Learning Algorithms help us to be more aware of the calamities and related disaster possibilities.

2. Overall Description

2.1 Project Perspective

This paper introduces the current machine learning approach to solving problems in the domain of landslide susceptibility assessment. Three different machine learning algorithms can be compared: Support Vector Machines, Decision Trees and Logistic Regression. Any three landslide slopes in a selected hilly area in the State of Kerala will be selected to perform the entire modeling procedure, from attribute and referent data preparation/processing, through the classifiers' implementation to

the evaluation, carried out in terms of the model's performance and agreement with the referent data.

2.2 User Classes and Characteristics

Anyone interested in the study of disaster management, calamity mitigation can make use of these prediction techniques. A serious investment in terms of time and effort based on this Machine Learning Algorithm based research would help the concerned departments – geology, morphology, ecology etc. as well.

2.3 Operating Environment

The scenario of people succumbing to landslide and erosions is a regular scene in Kerala especially during the rainy seasons. The selected areas and the study of their geographical significance in terms of elevation and soil composition and water absorption capacity can give us precise predictions as to ascertain when an incident could possibly occur post which we can take effective measures like relocating or timely evacuation of people, animals and other resources

2.4 Assumptions and Dependent Factors

We assume that the data made available from the possible sources – both historical and approximated measures are near to perfection. Before feature extraction, a fair amount of work will have to be done into preparing the data. The dataset would come in as a collection of all the geographical and spatial features contributing to a landslide in general.

3. Requirements

3.1 Dataset

The spatial, geographical features of the three selected landslide prone areas will be fed for iterations. The values generated on procedures will be furthered for comparisons using the three models of Machine Learning.

Location	**L1**	**L2**	**L3**
Geographical Features	**X1**	**X2**	**X3**
	Y1	**Y2**	**Y3**
	Z1	**Z2**	**Z3**

3.4 Programming Language

Python, the high-level programming language is used to run the Machine Learning Algorithms

4. Algorithm and Methods

1. Support Vector Machines (SVM)

The Support Vector Machine (SVM) is an increasingly popular learning procedure based on statistical learning theory, and involves a training phase in which the model is trained by a training dataset of associated input and target output values. The trained model is then used to evaluate a separate set of testing data. The characteristics compared are *slope angle*, *slope aspect*, *elevation*, *profile curvature of slope and topographic wetness index* etc. which will be used as environmental parameters which influence the occurrence of landslides.

2. Random Forests

Random Forest is a supervised learning algorithm. It involves the "bagging" method. The general idea of the bagging method is that a combination of learning models increases the overall result. It can be used for both classification and regression equations. Here we can employ this method for classification of the data and regression as well.

3. AdaBoost

AdaBoost, like Random Forest Classifier is another ensemble classifier. Here we can expect this model to retrain the algorithm iteratively by choosing the training set based on accuracy of previous training.

5. Observation and Recordings for Analysis and Comparative Study

The analysis report generated out of the recordings and the representations based on the algorithms run will be recorded for comparing between the models in order to suggest the most ideal of all. The idea is to identify the best techniques that can precisely predict the susceptibility index of the locations in common.

METHOD	**SVM**	**Random Forests**	**AdaBoost**
PRECISION			

Module Description

- ❖ Admin
  - View Emergencies
  - View Disaster Report
  - Send Notification
  - View Feedback
- ❖ User
  - Detect Landslide based on Location
  - Report Location Landslide
  - View Notification
  - Send Feedback
  - View other Disasters