Integrating Multilayer Perceptron Neural Nets with Hybrid Ensemble Classifiers for Deforestation Probability Assessment in Eastern India

Sunil Saha [1], Gopal Chandra Paul [1], Biswajeet Pradhan [2,3,4,*], Khairul Nizam Abdul Maulud [4,5], and Abdullah M. Alamri [6~~5~~]

[1] Department of Geography, University of Gour Banga, Malda, West Bengal, India, Email-sunilgeo.88@gmail.com (SS), gopalpaul.0321@gmail.com (GCP)

[2] Centre for Advanced Modeling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and Information Technology, University of Technology Sydney, New South Wales, Australia

[3] Department of Energy and Mineral Resources Engineering, Sejong University, Choongmu-gwan, 209, Neungdong-ro Gwangjin-gu, Seoul, 05006, Republic of Korea

[4] Earth Observation Center, Institute of Climate Change, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia

[5] Department of Civil Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

[6~~5~~] Department of Geology and Geophysics, College of Science, King Saud Univ., P.O. Box 2455, Riyadh 11451, Saudi Arabia

[*] Corresponding author: Biswajeet.Pradhan@uts.edu.au; biswajeet24@gmail.com

Abstract~~:~~

The rapid expansion of human settlement, agricultural land and roads because of population growth in several regions of the world has contributed to the depletion of forest land. In this study, novel ensemble intelligent approaches using bagging, dagging and rotation forest (RTF) as meta classifiers of multilayer perceptron (MLP) were used to predict spatial deforestation probability (DP) in Gumani Basin, India. The success rate and correctness of prediction of the ensemble models were compared with ~~those of the~~ MLP. A total of 1000 deforested pixels and 14 deforestation determining factors (DDFs) were used. The ~~novel~~ ensemble models were trained using 70% of the deforested pixels and validated with the remaining 30%. DDFs were chosen by applying the information gain ratio and Relief-F test methods. Distance to settlement, population growth and distance to roads were the most important factors. The results of DP modelling demonstrated that nearly 16.82%–12.64% of the basin had very high DP. All four models created

DP maps with reasonable prediction accuracy and goodness of fit, but the best map was produced by MLP-bagging. The accuracy of the MLP neural net model was increased 2-3% after ensemble with the hybrid meta classifiers (RTF, bagging and dagging). The proposed method could be used for deforestation prediction in other areas having similar geo-environmental conditions. Furthermore, the findings might be used as a basis for future research and could help planners in forest management.

Keywords: deforestation probability; hybrid ensemble techniques; machine learning; GIS; remote sensing; India

1. Introduction

Deforestation is a quasi-natural phenomenon occurring on our planet's surface (Wan Mohd Jaafar et al 2020). Worldwide, forests are affected by several threats, including population increase in urban areas, expansion of farming land and amenities, illegal mining and unregulated property rights (Newman et al. 2014; Gaveau et al. 2009; Robinson et al. 2014). The conservation of biodiversity and the removal of substantial carbon sink may help reduce carbon dioxide concentrations (Buchanan et al. 2008; Wang et al. 2009). Climate change, ambient carbon cycle imbalance and ecosystem degradation are the main environmental threats correlated with deforestation. Deforestation is considered as one of the most remarkable aspects of modifications in land use/land cover. Forest is a vital natural resource that provides a large range of ecological goods and facilities and plays a ~~critical~~ crucial role in balancing the atmospheric condition and, thus, climate change; therefore, forest cover change has become a global concern (Kumar et al. 2014). The effects of the growing strain on the environment have culminated in habitat destruction, deforestation and depletion for biodiversity (Sun et al. 2013; Nandi et al. 2007). Furthermore, the increased rate of soil erosion due to loss of forest cover may increase the environmental risks, such as landslide, water pollution and degradation of wetland ecosystem, which may have a major detrimental effect on the well-being of humans on a large scale (Glade, 2003; Körner et al. 2005; Wahab et al. 2019). Thus, identifying the underlying forces behind forest cover modification is crucial for recognising the transformation in our planetary ecosystem and reducing the speculation regarding spatial and temporal deforestation probability (DP) (Bax et al. 2016). The deforestation

process occurs in a haphazard fashion. On the basis of a set of suitable and desirable characteristics of physical and anthropogenic factors, forested lands are converted into other land use. For instance, forest patches near roads may have a high chance of being deforested. Similarly, low-elevation and gentle slope areas are favourable for cultivation, and farmland ~~has a higher possibility to be expanded~~ than rough terrain (Lambin et al. 2001; Turner et al. 2001). Understanding the causes of deforestation is, therefore, important in the formulation of effective mitigation steps and policies (Hosonuma et al. 2012). Causes of deforestation and the severity of their effects differ considerably from one region to another region and change over time. Most causes have been described as leading to rather than accelerating deforestation (Geist et al. 2007). Some deforestation research has focused on anthropogenic forces, although the analysis of deforestation processes requires considering the natural and anthropogenic aspects of the ecosystem (Bax et al. 2016; Wan Mohd Jaafar et al 2020).

Traditional approaches used for analysing deforestation suffer from a series of limitations, such as follows: 1. correlation cannot be regarded as a clear indicator of the source; 2. statistical models selected for prediction may have minimal explanatory importance; 3. relationships can be nonlinear. With recent advances in remote sensing (RS), geographic information systems (GIS) and various statistical techniques, spatial DP can be forecasted ~~precisely~~more accurately (Arekhi, 2011; Houet et al. 2006; Pontius et al. 2001; Maya Liyana Hamzah et al. 2020; Siti Nor Maizah Saad et al. 2020). In the Carpathian Mountains, the increasing accessibility to large temporal satellite imagery and the development of GIS and RS tools have facilitated the comprehensive study of past human-induced forest depletion. Many areas have also been studied at national (Munteanu et al. 2014; 2015) and international scales (Kaim et al. 2018; Sobala et al. 2017; Szymura et al. 2018; Wan Mohd Jaafar et al. 2020). Several scholars have prepared DP models based on logistic regression algorithms in tropical areas (Kumar et al. 2014; Bavaghar, 2015; Kucsicsa et al. 2019). Traditional unsupervised techniques, including regression analysis (Ludeke et al. 2019), change vector analysis (Nackaerts et al. 2005) and principal component analysis (Ortega et al. 2020; Deng et al. 2008), have been widely used to detect changes in forest cover. Artificial intelligence (AI) and machine learning (ML) algorithms have been widely adopted for mapping different hazards and potentiality~~,~~ such as gully erosion susceptibility, landslide susceptibility (Roy et al. 2019), flood susceptibility (Khosravi et al. 2018), land subsidence (Tien Bui et al. 2018); Individual tree crown detection and delineation (Wan Mohd Jaafar et al. 2018)

and groundwater potentiality mapping (Tien Bui et al. 2019)~~, have been widely adopted for mapping different hazards and potentiality~~. In all those cases, ML and AI methods have shown good capability in modelling hazards. ML techniques ~~are~~ have been ~~currently~~ used for the prediction of deforestation. Ortega et al. (2020) used the deep learning technique and support vector machine to detect deforestation. Saha et al. (2020) used random forest and reduced error pruning trees (REPTree) for modelling the DP. Dlamini (2016), Krüger and Lakes (2015) and Mayfield et al. (2017) used Bayesian networks for assessing DP, which provided reasonable results.

In recent years, several authors have used hybrid ensemble methods for mapping landslides (Fang et al. 2020), gully erosion (Roy et al. 2020) and groundwater potentiality (Rahmati et al. 2018) and~~;~~ these techniques have achieved better results than individual models. Ensemble method is a learning in which several models, such as classifiers, are systematically produced and integrated to solve a specific computational intelligence problem. Ensemble method is mainly used to enhance a ~~model 's~~model's efficiency (classification, estimation, etc.) or minimize the possibility of an unexpected selection of a weak one. The ensemble of hybrid meta classifier and artificial neural network is still not used in the field of deforestation modelling. ~~These ensemble methods provided better results than single ML model.~~ On the basis of the accuracy of the hybrid ensemble models used in the above-mentioned fields, the current work addressed the question that hybrid ensemble methods are equally accurate for DP modelling or not. We selected ensembles of multilayer perceptron ~~(MLP)~~ neural nets (MLPnn) and three hybrid ensemble models, i.e. MLP-bagging, MLP-dagging and MLP-rotation forest (RTF), to prepare DP maps of the study area.

The novelty of this work is that the employed hybrid ensembles of MLPnn and bagging, dagging and RTF models ~~used~~ have~~had~~ not been used for deforestation modelling. This work not only included these methods but also used Friedman and Wilcoxon signed-rank tests for judging the difference among the DP maps produced by these models, which are also relatively new in this field. Information about the forest cover changes of this area remains limited. In this situation, RS is a vital source of data for the effective monitoring of this region. The forest cover changes were demarcated using the normalised difference vegetation index (NDVI). The DP maps would help the researchers and decision makers of this region. In addition, these sorts of methods have not yet been used in this area, as well as in India for the evaluation of DP. The detailed explanation of all of these methods and parameters would direct future researchers working in this field.

The purpose of this research is to evaluate the DP in the Gumani River Basin, India by applying the hybrid ensemble frameworks of MLPnn and ensemble strategies, i.e. bagging, dagging and RTF. Preparation of the probability map for deforestation is helpful ~~for~~ to policymakers~~ing~~ ~~to~~ for identifying the areas susceptible to deforestation and evaluating~~e~~ the current forest management.

2. Description of the Study Area

The Gumani River is located in the fringe area of the Chhota Nagpur Plateau of India. It is the tributary of the Ganga River having a length of 120.09 km. Geographically, this basin extends from 24°37'39"N–25°7'19"N lat. and 87°21'20"E–87°54'20"E lon. (Figure 1), encompassing an area of 1274.57 km^2. The forested area has been decreased from 24.11% (1990) to 14.33% (2020) of the total area of the basin (Landsat TM 1990 and OLI 2020 images of the USGS Earth Explorer). The lower part of the basin is agriculturally prosperous, whilst the upper part has a high concentration of population and settlement. Population growth is high in this study area; the total population was 560,000 in 1991 and increased to 750,000 in 2011 (Census of India, 2001, 1991). Therefore, population increase has a detrimental effect on the forest cover, whilst attention should be given to geographical context and other criteria of forest depletion. Geologically, this area comprises Rajmahal Traps, lower Vindhya system, lower Gondwana system and new alluvium. This basin often has different geomorphological nature because the upper portion belongs to the undulating plateau and the lower portion is a plain area. The elevation of the study area ranges from 17 m to 581 m from the mean sea level. The climate varies from subtropical humid to subhumid (Chandniha et al. 2017). Rainfall in this basin mainly occurs between June and September (Chandniha et al. 2017). The mean annual rainfall is 1,300 mm (Chandniha et al. 2017). According to the National Bureau of Soil Survey and Land Use, the prevalent soils are fine loamy, loamy skeleton and clay skeleton. The forest concentration is mainly high in the upper portion of the basin and low in the lower portion. For ~~the protection of~~protecting the remaining forest areas in the basin, prediction of deforestation area and formulation of suitable strategies by the local government are necessary. Our work would help the decision makers in this respect.

Figure 1. SOMEWHWRE HERE

3. Background Theory of Methods Employed

3.1. Ensemble Model for DP Assessment

DP models using ensemble structures of MLPnn and bagging, dagging and RTF for spatial DP were obtained through four key stages (Figure 2).

1. *Selection of deforestation determining factors (DDFs):* After the survey of the published literature, the DDFs were selected. The selected parameters were justified using two statistical methods, i.e. information gain ratio (IGR) and Relief-F. Deforestation affecting factors were divided into two classes, namely, natural factors (viz. altitude, slope, forest density, distance to forest edge, proximity to river, aspect and topographic position index, [TPI]) and anthropogenic factors (viz. population density, agricultural land density, distance from agricultural land, proximity to road, settlement density, proximity to settlement and population growth) in the~~is, for DP~~ DP analysis.

2. *Collection and preparation of data layers:* Data regarding deforested locations and DDFs were collected to predict spatial DP. In January 2020, an intensive field investigation with a handheld global positioning system was conducted to validate the deforested locations ~~from~~ collected through the interpretation of Google Earth images and NDVI ~~prepared from the Landsat imageries~~.

3. *Assessment of the contribution of the DDFs:* A frequency ratio (FR) model was used, and the percentage shear of the sample deforestation points was calculated for judging the significance of the DDFs.

4. *Preparation of deforestation models and DP maps:* To construct deforestation models, ensemble methods were firstly, implemented to refine the training data set. Input configured data were then utilised to categorise the groups for the probability of spatial deforestation by using the MLPnn base classifier. Finally, frameworks of ML ensemble were built for DP models.

5. *Validation and comparison of models:* Using the ROC, efficiency, accuracy, MAE and RMSE DP maps were validated and compared in consideration of the training and testing data ~~-~~sets. Friedman and Wilcoxon statistical signed-rank tests were performed to check whether differences exist amongst the DP models or not.

Figure 2. SOMEWHWRE HERE

3.2. Data Used

3.2.1. Deforestation Map

The forest cover change (1990–2020) was considered a dependent variable (Figure 3) for DP modelling. NDVI was measured from the Landsat images of 30 m × 30 m resolution for 1990 (Figure 3a), 2000 (Figure 3b), 2010 (Figure 3c) and 2020 (Figure 3d) via GIS tools, and NDVI values greater than 0.3 were considered forest (~~Gayen et al. 2018;~~ Weier et al. 2000). During these decades, nearly 9% of forest cover was lost. The forest cover areas are 24.11%, 20.96%, 16.56% and 14.33% of the total basin area for the years of 1990 (3a), 2000 (3b), 2010 (3c) and 2020 (3d), respectively. NDVI map of 1990 of the study area was considered as the base map for this study. A binary map with the groups of 'deforestation' and 'non–deforestation' was produced by subtracting the forest cover from 1990 to 2020 (Figure 3e) for the duration of 1990–2020. For preparing the DP models and obtaining enhanced result, 1000 pixels for both classes, i.e. deforested and non–deforested, were randomly selected from the total deforestation and non-deforestation pixels (Süzen et al. 2004). Amongst them, 70% were considered for modelling, and 30% were selected for validating the models.

Figure 3. SOMEWHWRE HERE

3.2.2. Preparation of DDFs

For constructing the DP models, seven natural factors (i.e. altitude, slope, forest density, distance from forest edge, proximity to river, aspect and TPI) and seven anthropogenic factors (i.e. density of population and agricultural land, distance from agricultural land, proximity to road, settlement density, proximity to settlement and population growth rate) were selected (Table 1). These factors were considered as independent factors, and a thematic layer for each variable was prepared. In Table 1, methods of preparing the factors and sources of data have been presented.

Table 1. SOMEWHWRE HERE

The regional topography condition plays an important role in the forest cover change. Spatial variation in the deforestation process is influenced by slope, altitude, aspect and TPI (Bax et al. 2016; Szymura et al. 2018). The slope classes determine the spatial variability in deforestation process (Kumar et al. 2014; Bavaghar, 2015; Bax et al. 2016; Siles, 2009; Szymura et al. 2018; Vanonckelen et al. 2015). A slope map (Figure 4a) was extracted from ASTER DEM with a resolution of 30 m × 30 m (Table 1). Aspect (Figure 4j) controls the amount of sunlight and rainfall

of a particular region (Kumar et al. 2014; Bavaghar, 2015; Bax et al. 2016). It affects the composition and development of forest cover. The degree of deforestation is also indirectly connected to slope face (Bayat, 2000). The DEM of the basin was considered the altitude map (Figure 4k). In high-altitude areas, natural hazards, such as weathering, aeolian flooding and landslide, are the main drivers of deforestation; in low-altitude areas, deforestation is induced mostly by anthropogenic activities (Ercanoglu et al. 2002). Distance to the river is a parameter that determines the stability and instability of slope, indirectly influencing the forest cover change (Yalcin, 2008; Saha et al. 2002). Waterbodies may be exposed to forested areas and reflect secondary routes for timber collection (Nackaerts et al. 2005). For distance to river, a thematic layer was prepared in a GIS environment by using the Euclidean distance buffer tool (Figure 4c). The distance from the margins of forest is an important factor that can regulate deforestation (Matlack, 1994). This factor is an intermediate area from which forest destruction continues at the border of existing forest (Kumar et al. 2014; Arekhi, 2011). DP is determined using the nature and features of forest edge in the core forest region. This thematic layer was also produced using the Euclidean distance buffer tool (Figure 4f). An inverse relationship exists between forest density and DP (Bouldin, 2008). A forest density map was prepared by dividing the forested area by total area based on the forest map of 2020 (Figure 4b). Topographic Position Index (TPI) compares the elevation of each cell in a DEM to the mean elevation of a specified neighborhood around that cell. TPI classes affect the spatial variability in the deforestation process (Kumar et al. 2014; Bavaghar, 2015; Bax et al. 2016; Siles, 2009; Szymura et al. 2018; Vanonckelen et al. 2015; Wilson et al. 2005). TPI was created on the basis of DEM and applied for extracting the slope position classes (Jennes, 2006). According to Weiss (2001), TPI was classified into six categories in this study area (Figure 4n), namely, 1) ridge (TPI > 1SD); 2) upper slope (0.5SD < TPI ≤ 1SD); 3) middle slope (−0.5SD < TPI < 0.5SD, slope > 5°); 4) lower slope (−1SD < TPI ≤ −0.5SD); 5) flat (−0.5SD < TPI < 0.5SD, slope ≤ 5°); 6) valley (TPI ≤ −1SD).

Different sociocultural and economic practices are mainly responsible for the degradation and loss of forest (Boudreau et al. 2005). The potentiality of deforestation is multiplied as the population continues to grow near a forested area (Szymura et al. 2018; Vanonckelen et al. 2015). As a result, population growth (Figure 4m), population density (Figure 4g), distance to settlement (Figure 4h) and settlement density (Figure 4i) are the main reasons for deforestation. A reciprocal relationship exists between forest cover change and settlement density. As settlement density

(Figure 4i) increases, the probability of deforestation in its neighbouring parts will be increased and vice versa. The installation of road systems across land cover proceeds to divide the forest land and is the first move towards forest depletion. The road network is a vital deforestation-triggering factor because the forest close to the road is highly prone to degradation and vice versa (Chomitz et al. 1993). The chances of deforestation are high in accessible areas (Bavaghar, 2015). Here, a distance-to-road map was produced using the Euclidean distance buffer tool (Figure 4e). Rapid population increase is the main cause of deforestation (Michalski et al. 2008). Much inhabitants need substantial food and house and, hence, considerable land for farmland and houses (Cropper et al. 1994). Overpopulation is considered the major cause of forest destruction in accordance with international organisations, including FAO. The population density map of the study area was constructed on the basis of data from the 2011 census (Figure 4g). Agricultural land density (Figure 4l) is an important factor for assessing the DP of a particular region because it identifies the concentration of agricultural land in a particular area. The chances of deforestation are high where the density of agricultural land is high. The distance to agricultural land (Figure 4d) is also an important land use predictor for determining DP. The chances of deforestation will be increased as the distance decreases and vice versa because a high probability of building or other human land usage will occur near an agricultural field. Population growth can be followed by a high rate of forest cover change (Szymura et al. 2018; Vanonckelen et al. 2015). The population growth (Figure 4m) data were collected from the Census of India (2011). High rates of population growth lead to the increase in settlement and agricultural area in the area of forest cover (Minetos et al. 2010).

Figure 4. SOMEWHWRE HERE

3.32. Factor Selection

The selection of conditioning variables is a challenging task in any study because no specific criteria are available. Bui et al. (2016) and Roy et al. (2020a) identified effective factors by using statistical models for natural hazard assessment. Gayen et al. (2018) used multicollinearity analysis for selecting DDFs. Different statistical methods, such as correlations, regressions, Relief-F tests, IGR, probabilistic models and ML models, can also be used to select DDFs. In this study, the IGR and Relief-F methods were applied for selecting the important deforestation determining factors. IGR solves the weakness of information gain related to attributes that can take on a vast range of

different values that could learn the set of training too well. IGR has been used to assess which of the factors are perhaps the most significant. Relief-F algorithms have often been used as a form of selecting features that is implemented in a pre-possessing period well before the model is trained and is one of the most powerful pre-processing algorithms.

3.~~3~~2.1. Information Gain Ratio (IGR)

For DP, anthropogenic and natural factors do not have the same diagnostic power and may even reduce the predictive capacity of a model. If we remove the irrelevant DDFs from the model, enhanced findings and prediction can be obtained (Martínez-Álvarez et al. 2013). IGR is amongst the most effective factor selection strategies (Tien Bui et al. 2016). Information is gained on the basis of an intelligent principle that helps reduce variance and shows the importance of influencing variables. In data mining, IGR is an important strategy for quantifying factor predictability (Witten et al. 2011). Quinlan (1993) established the IGR, in which a high ratio means a great predictive capacity. In the supplementary material section, equations used to calculate IGR are mentioned (S1). In this study for identifying as well as selecting the important DDFs IGR was used. ~~Here, IGR was calculated using Equation 1.~~

~~Given training data S consisting of n input samples, $n(L_i, S)$ is the number of samples in the training data S belonging to class L_i (deforestation, non–deforestation). The information (entropy) needed to classify S was calculated as~~

$$Info(S) = -\sum_{i=1}^{2} \frac{n(L_i, S)}{|S|} \log_2 \frac{n(L_i, S)}{|S|} \quad (1)$$

~~The amount of information needed to divide S into $(S_1, S_2, \ldots, S_m)$ regarding the conditioning factor for land subsidence A was estimated as~~

$$Info(S, A) = \sum_{j=1}^{m} \frac{S_j}{|S|} Info(S) \quad (2)$$

~~The IGR for a certain conditioning factor for land subsidence A was computed as~~

$$Information\ Gain\ Ratio = \frac{Info(S) - Info(S, A)}{SplitInfo(S, A)} \quad (3)$$

~~where $SplitInfo(S, A)$ reflects the information gained by separating training data S into subsets. $SplitInfo(S, A)$ was calculated as~~

$$SplitInfo(S,A) = -\sum_{j=1}^{m} \frac{|S_j|}{|S|} \log_2 \frac{|S_j|}{|S|} \quad (4)$$

3.32.2. Relief-F Test Method

The Relief-F method, implemented by Kira and Rendell (1992), iteratively changes the weights of features in accordance with their capacity to distinguish between adjacent shapes. The principal concept of the Relief-F algorithm is similar to the specific rules of the k-nearest neighbour algorithm (Altun et al. 2007). Being in the same class is likely to yield a distance close to a given distance. If the attribute is useful, the closest distances of the same class are expected to be closer to the range given throughout this attribute than the closest distances of all other classes (Altun et al. 2007). Mathematically, X is assumed to be a randomly drawn sample of the outcomes of a binary test. Two closest neighbours, one from the same class (strike or NH) and the other from another class (miss or NM) should be evaluated. Then, the weight (wi) for the i-th feature is updated via a heuristic computation (Cai et al. 2012), i.e.

$$w_i \leftarrow w_i + \left|x^i - NH^{(i)}\right| - \left|x^i - NM^{(i)}\right| . \quad (51)$$

Further information on the algorithm is provided in the paper of Liu and Motoda (2008).

3.3. Data Used

3.3.1. Deforestation Map

The forest cover change (1990–2020) was considered a dependent variable (Figure 3) for DP modelling. NDVI was measured from the Landsat images of 30 m × 30 m resolution for 1990 (Figure 3a), 2000 (Figure 3b), 2010 (Figure 3c) and 2020 (Figure 3d) via GIS tools, and NDVI values greater than 0.3 were considered forest (Weier et al. 2000). During these decades, nearly 9% of forest cover was lost. The forest cover areas were 24.11%, 20.96%, 16.56% and 14.33% of the total basin area for the years of 1990 (3a), 2000 (3b), 2010 (3c) and 2020 (3d), respectively. For the duration of 1990–2020, a binary map with the groups of 'deforestation' and 'non-deforestation' was produced by subtracting the forest cover from 1990 to 2020 (Figure 3e). For preparing the DP models and obtaining enhanced result, 1000 pixels for both classes, i.e. deforested and non-deforested, were randomly selected (Süzen et al. 2004). Amongst them, 70% were considered for modelling, and 30% were selected for validating the models.

Figure 3. SOMEWHWRE HERE

3.3.2. Preparation of DDFs

For constructing the DP models, seven natural factors (i.e. altitude, slope, forest density, distance from forest edge, proximity to river, aspect and TPI) and seven anthropogenic factors (i.e. density of population and agricultural land, distance from agricultural land, proximity to road, settlement density, proximity to settlement and population growth rate) were selected. These factors were considered independent factors, and a thematic layer for each variable was prepared. ,Data were gathered from different sources, such as ASTER digital elevation model (DEM), Landsat images of 30 m × 30 m from the USGS, topographical sheets of 1:50000 scale from the Survey of India and population data from the Census of India, to produce the thematic layers of the selected DDFs.

The regional topography condition plays an important role in forest cover change. Spatial variation in the deforestation process is influenced by slope, altitude, aspect and TPI (Bax et al. 2016; Szymura et al. 2018). The slope classes determine the spatial variability in deforestation process (Kumar et al. 2014; Bavaghar, 2015; Bax et al. 2016; Siles, 2009; Szymura et al. 2018; Vanonckelen et al. 2015). A slope map (Figure 4a) was extracted from ASTER DEM with a resolution of 30 m × 30 m. Aspect (Figure 4j) controls the amount of sunlight and rainfall of a particular region (Kumar et al. 2014; Bavaghar, 2015; Bax et al. 2016). It affects the composition and development of forest cover. The degree of deforestation is also indirectly connected to slope face (Bayat, 2000). The aspect map (Figure 4j) was extracted from the ASTER DEM. The DEM of the basin was considered the altitude map (Figure 4k). In high-altitude areas, natural hazards, such as weathering, aeolian flooding and landslide, are the main drivers of deforestation; in low-altitude areas, deforestation is induced mostly by anthropogenic activities (Ercanoglu et al. 2002). Distance to the river is a parameter that determines the stability and instability of slope, indirectly influencing the forest cover change (Yalcin, 2008; Saha et al. 2002). Waterbodies may be exposed to forested areas and reflect secondary routes for timber collection (Nackaerts et al. 2005). For distance to river, a thematic layer was prepared in a GIS environment by using the Euclidean distance buffer tool (Figure 4c). The distance from the margins of forest is an important factor that can regulate deforestation (Matlack, 1994). This factor is an intermediate area from which forest destruction continues at the border of existing forest (Kumar et al. 2014; Arekhi, 2011). DP is determined using the nature and features of forest edge in the core forest region. This thematic layer was also produced using the Euclidean distance buffer tool (Figure 4f). An inverse relationship exists between forest density and DP (Bouldin, 2008). A forest density map was

prepared by dividing the forested area by total area (Figure 4b). TPI classes affect the spatial variability in the deforestation process (Kumar et al. 2014; Bavaghar, 2015; Bax et al. 2016; Siles, 2009; Szymura et al. 2018; Vanonckelen et al. 2015; Wilson et al. 2005). TPI was created on the basis of DEM and applied for extracting the slope position classes (Jennes, 2006). According to Weiss (2001), TPI was classified into six categories in this study area (Figure 4n), namely, 1) ridge (TPI > 1SD); 2) upper slope (0.5SD < TPI ≤ 1SD); 3) middle slope (−0.5SD < TPI < 0.5SD, slope > 5°); 4) lower slope (−1SD < TPI ≤ −0.5SD); 5) flat (−0.5SD < TPI < 0.5SD, slope ≤ 5°); 6) valley (TPI ≤ −1SD).

Different sociocultural and economic practices are mainly responsible for the degradation and loss of forest (Boudreau et al. 2005). The potentiality of deforestation is multiplied as the population continues to grow near a forested area (Szymura et al. 2018; Vanonckelen et al. 2015). As a result, population growth (Figure 4m), population density (Figure 4g), distance to settlement (Figure 4h) and settlement density (Figure 4i) are the main reasons for deforestation. A reciprocal relationship exists between forest cover change and settlement density. As settlement density (Figure 4i) increases, the probability of deforestation in its neighbouring parts will be increased and vice versa. The installation of road systems across land cover proceeds to divide the forest land and is the first move towards forest depletion. The road network is a vital deforestation-triggering factor because the forest close to the road is highly prone to degradation and vice versa (Chomitz et al. 1993). The chances of deforestation are high in accessible areas (Bavaghar, 2015). Here, a distance-to-road map was produced using the Euclidean distance buffer tool (Figure 4e). Rapid population increase is the main cause of deforestation (Michalski et al. 2008). Much inhabitants need substantial food and house and, hence, considerable land for farmland and houses (Cropper et al. 1994). Overpopulation is considered the major cause of forest destruction in accordance with international organisations, including FAO. The population density map of the study area was constructed on the basis of data from the 2011 census (Figure 4g). Agricultural land density (Figure 4l) is an important factor for assessing the DP of a particular region because it identifies the concentration of agricultural land in a particular area. The chances of deforestation are high where the density of agricultural land is high. The distance to agricultural land (Figure 4d) is also an important land use predictor for determining DP. The chances of deforestation will be increased as the distance decreases and vice versa because a high probability of building or other human land usage will occur near an agricultural field. Population growth can be followed

~~by a high rate of forest cover change (Szymura et al. 2018; Vanonckelen et al. 2015). The population growth (Figure 4m) data were collected from the Census of India (2011). High rates of population growth lead to the increase in settlement and agricultural area in the area of forest cover (Minetos et al. 2010).~~

~~**Figure 4.** SOMEWHWRE HERE~~

3.4. Deforestation Occurrence in Relation to DDFs and Analysis of Its Influence

The percentage of deforestation samples and the FR of subclasses of each factor were calculated to understand the influences of the selected DDFs on the deforestation process. The percentage of deforested sample in subclasses of each explaining variable was calculated by overlaying each raster representing independent variables with the randomly selected deforestation pixels. FR provides a proportion of deforestation pixels in a specific category for each input layer (Lee et al. 2006). FR values (Equation 2) based on the frequency of deforestation samples were calculated using the following equation:

$$FR = \frac{f/tf}{x/tx}, \quad (26)$$

where, f refers to the pixels of deforestation in the explanatory variable subclass, tf indicates the total deforestation pixels, x denotes the total pixels in the explanatory variable subclass, and tx is the total number of pixels.

3.5. Base Classifier of MLPnn

MLPnns are regarded as the techniques of artificial neural networks (ANN) and are commonly utilised in classification (Haykin, 2009). MLPnn is a feedforward neural network and for the training process, it uses backpropagation. No decision has been reached about the relative values of individual input variables, the plurality of inputs is set on the basis of weight adjustment throughout the training phase, and the distribution of the training data set is independent of the pre-assumptions in these techniques (Gardner et al. 1998). Three main sequences exist for creating the neural networks in MLP, i.e. input, hidden and output layers (Figure 5). In accordance with a specific application, every layer in a network contains adequate neurons. The input layer is inactive and rarely gathers data (e.g. data from various DDFs). Hidden and output layers analyse information on a constant basis. Input layers are known as variables influencing deforestation, output layers are regarded as the graded outcomes of inferring deforested or ~~non-deforested~~non-deforested groups, and hidden layers are the categorising layers for converting inputs into outputs.

MLP Neural Nets have shown to be ~~a stronger~~performing better than conventional classification methods (Benediktsson et al-. 1990). There are some benefits of using this approach: (1) there are no pre-assumptions as to the distribution of the training dataset, (ii) there is no need to decide on the relative importance of the various input measures, and (iii) the weights are changed to choose the most input measures during the training process (Gardner and Dorling 1998).

MLPnns are ~~subject to~~of two key phases: (I) inputs are transmitted via the hidden layers to the output values, then the output values are compared with the pre-values to approximate the differentiation; (II) in achieving the best performance, weights are balanced to eliminate the disparity. Let x = xi, i = 1, 2,..., 14 is the vector of the 14 factors impacting deforestation, and y = 1 (deforested) or 0 (non-deforested). The number of neurons in the input and output layers is generally calculated via operation. The number of hidden layers and their neurons is quantified by trial and error (Gong, 1996). For a classification question, MLPnn data processing includes three stages: learning, weighting, and classification stages. The learning phase happens with the issuance of random initial relational weights, which are continuously revised until the correct training efficiency is achieved. Subsequently, the modified weights derived from the prepared network are often used to process test data and assess the overall precision and effectiveness of the application. The network efficiency is assessed by evaluating the consistency of training and test data in terms of the percentage and overall accuracy of classification (Congalton, 1991). Learning information from the input neurons is considered to acquire the information of the output neurons by using the hidden neurons. Neuron $_j$ obtained from neuron $_i$ in its corresponding input layer in the first hidden layer can be represented as:

$$x' = \sum_{i=1}^{t} w_{ij} p_i \quad , \qquad (37)$$

where w_{ij} reflects the weight of the association between input neuron i and hidden neuron j, pi is the data at input neuron I, and t is the input neuron number. The output value generated in the concealed neuron j, p_j, is the transfer function, f, which is evaluated as the amount provided in neuron j, x'. f, the transfer function, can be described as

$$p_j = f(x') = \frac{1}{1+e^{-x'}} \quad . \qquad (48)$$

Function f is typically a nonlinear sigmoid feature that is implemented to the weighted sum of input data until the data are transferred to the next stage.

The sum of the squared differences between the expected and actual output neurons E values is defined as follows (Subasi, 2007):

$$E = \frac{1}{2}\sum_{j}(Y_{dj} - Y_j)^2 \quad (5)$$

where Y_{dj} is the expected output neuron j and Y_j is the actual output neuron. Each w_{ji} weight is adjusted to lessen the value E based on the training algorithm used. In this study, MLPnn was fitted with 500 epochs, ~~5~~ 1 hidden layers and ~~validit~~validation~~y~~ threshold of 20 generated from the trial-and-error process to avoid overfitting cases.

Figure 5. SOMEWHWRE HERE

3.6. ML Ensemble Techniques

3.6.1. RTF

RTF is an ensemble approach assembled with individual decision trees (Kuncheva et al. 2007) and initially proposed for classification by Rodriguez et al. (2006). It is based on the concept of a random forest approach aimed at creating reliable and flexible classifiers (Rodriguez et al. 2006). An individual tree is configured inside the RTF with compressed data sets associated with the space rotated using ~~a function~~Principal Component Analysis (PCA). In this model, bootstrap samples are used as a training set for specific classifiers (Kuncheva et al. 2007). Throughout this process, points are derived from training datasets using base classifier to generate learning sub-training datasets (Pham et al., 2016b). The function of DDFs in this analysis is $x=(x_1, x_2, \ldots\ldots x_n)$. $Y=(y_1, y_2)$ denotes the main vector divisions, deforested or not deforested. *D* stands for the training data. $F_1, F_2, \ldots F_n$ are categori~~z~~sed in accordance with the ensemble. T specifies a certain set of DDFs and is divided into sub-classes k. A new training nonempty subset X'_{ij} is prepared by applying the bootstrap method where F_{ij} is the j^{th} subset of features to run classifier D_i. Further, a linear transformation is used to X'_{ij} to prepare coefficients of matrix C_{ij} wherein size of each matrix of X'_{ij} is M×1 with the coefficients of $r_{ij}^{(1)} \ldots\ldots r_{ij}^{(k)}$. Ensemble RTF is established on the basis of the rotation matrix formed using the basic methods of characterisation and conversion (Xia et al. 2007). The rotation matrix is obtained by rearranging Ri matrix.

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

$$R_i=\begin{bmatrix} ri_1^{(1)},ri_1^{(2)},....ri_1^{(S1)} & 0 & ..... & 0 \\ 0 & ri_1^{(1)},ri_1^{(2)},....ri_1^{(S2)} & ..... & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & ..... & ri_1^{(1)},ri_1^{(2)},....ri_1^{(Sk)} \end{bmatrix} \quad (6)$$

In this matrix, columns of R are reorganized as per original feature and a novel reorganized rotation matrix is called as R_r^i wherein xR_r^i signify the altered training set for classifier Di and all classifiers are to be run in a similar method.The obtained coefficients that are created for each entity class are organised using a sparse rotation matrix called R_i via the average mixture strategy.

$$\mu_j^{(x)}=\frac{1}{n}\sum_{i=1}^{n} d_{ij}(x\,R_i^r),\ j=1,2.....c, \quad (79)$$

where $\mu_j^{(x)}$ is the chief confidence allocated to the class of y_j, the likelihood allocated by the classifier Di and the regression dij is $d_{ij}(x\,R_i^r)$. In this hypothesis, x is from class y_j, and c is the number of classes (Rodriguez et al., 2006).~~shows the generated probability of Ci classifier regarding hypotheses, and k class is activated using c. c is attributed to the highest support group.~~

3.6.2. Dagging

Dagging is a well-known re-sampling ensemble approach that produces and integrates a number of classifiers utilizing the same learning algorithm for base-classifiers. Ting and Witten proposed dagging in 1997. The procedure varies in many respects from the process of boosting and bagging. For example, based on the outcome of the previously generated classifiers, the boosting technique adapts the training data set in terms of distribution, while bagging modifies it stochastically and boosts the basis of the success of each classifier as a voting weight. For multiple disjoint experiments, dagging is used as a replacement for bootstrap experiments to obtain base classifiers (Ting and Witten, 1997; Kotsianti et al. 2007). Furthermore, strong empirical indications prevail that dagging in noisy settings is far more resilient than boosting. A resampling ensemble strategy is used to merge multiple classifiers for ensuring improved predictive performance of base classifiers dependent on majority voting (Kotsianti et al. 2007). For this purpose, we created an ensemble in this research using dagging ensembles with MLPnn base classifier through voting methodology.

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

3.6.3. Bagging

Bagging, designed by Breiman (1996), combines several cases of training dataset and uses bootstrap aggregation technique to achieve results of strong predictive precision centered on a based classifier (Wu et al. 2020). It was used to provide a precise mapping of DP. For very large ensembles, bagging gives great results; having a greater number of estimators results in increasing the accuracy of these approaches in comparison to RTF model. Such ensemble is chosen because a slight change in the training data represents and enhances the capacity for estimation (Wu et al. 2020). Random selection of bootstrap samples to create a range of training subsets, generation of classifiers of several models, and combining the classifier development in the final model are the three main steps in bagging (Bui, et al. 2016). In bootstrap experiments, one third of instances are not exterminated in the early test process. Bagging classifier in the bagging system uses the displacement approach to produce a bootstrap sample from the actual training dataset. The bagging hybrid ensemble solution enhances the success to each array of classifiers by linking them to the original feature scheme for the bagging categorisation phase. These cases were recognised by Breiman (1996) as off-bag tests. A Bagging fits each base classifier on random subsets of the initial dataset and then aggregates their individual predictions to form a final prediction (either by voting or by averaging).

3.7 Construction of ~~DP~~ Models and DP Maps

DP models utilising hybrid ML ensemble frameworks were developed using training data sets to predict the deforestation in the study area. For running the ML models ~~continuous values of~~ continuous ~~factors~~ and categorical ~~values of categorical~~ factors were used. The continuous DDF were classified based on the natural break classification method for the frequency ratio model as to know the influence of the sub-categories of the DDF through FR model. Deforested and forested pixels were considered as the training datasets. Pixels (70%) from both classes were randomly set as training datasets for running the models. The deforestation and non-deforestation were characterised as 0 and 1 codes, respectively. Once all the four models were effectively run in the training phase, the relational weights of the models were applied to compute the DP indices for all pixels. The measuring variables were standardised by training via the trial-and-error method to construct such DP models. Generally, 1 to 2 hidden layers are enough for pixel based mapping. For modelling the DP in this study using ensemble models ArcGIS and R-studio were used. Caret, rpart, ipred, rotationForest, neuralnet packages of R studio were used for predicting the

deforestation probability ~~in this research~~. In this analysis, we used 1 hidden layers, 0.3 learning rate, 0.2 momentum, 0 seed, 500 training times and 20 validation thresholds for the MLPnn to: decide the quantity of data for reduced-error pruning, upgrade weight, add value to the weight, divide the data, and build the ensemble and finish the calibration testing (Pham et al. 2016; Onan, 2016). The validation threshold is the value being used by validation test to be terminated. A threshold function is a Boolean function which determines whether a certain threshold is crossed by the value equality of its inputs. The percentage bag size indicates the training range size (Sedano et al. 2013). Likewise, 16 iterations, 1 seed, 100% of bag size (training range size) and MLPnn as base classifiers were set for bagging. Eighteen iterations, 2 seeds and MLPnn as base classifiers and 8 iterations, 1 seed and principal component analysis as base filters were used.

3.87. Validation Techniques

3.87.1. Threshold-dependent methods

ROC curve remains the most effective and acceptable approach that can effectively test models (Kumar et al. 2011). In this study, three threshold dependent methods i.e. ROC, precision and accuracy were used for effectively evaluate the performance of the used models. The area under the curve (AUC) indicates the effectiveness and consistency of the models (Pepe et al. 2000). The ROC curve has been used in various disciplines and branches (e.g. engineering and medical). Accuracy and precision have been considered for checking the robustness of models. Equations of AUC, sensitivity, specificity, precision and accuracy are mentioned in the supplementary material section (S2). High values of AUC, precision and accuracy indicate the good capability of models. AUC values vary from 0 to 1; an AUC value is highest with 1 which suggests a perfect estimation, whereas an AUC value < 0.5 implies poor results (Can et al. 2005).

~~$$Sensitivity = \frac{TP}{TP + FN}, \quad (10)$$~~

~~$$Specificity = \frac{TN}{FP + TN}, \quad (11)$$~~

~~$$AUC = \left(\sum TP + \sum TN\right) / (P + N), \quad (12)$$~~

~~$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \quad (13)$$~~

$$Precision = \frac{TP}{TP + FP}, \quad (14)$$

where, TP, TN, FP and FN indicate true positive, true negative, false positive and false negative, respectively; P and N are the total numbers of deforestation and non deforestation pixels, respectively. AUC values vary from 0 to 1; an AUC value is highest with 1 which suggests a perfect estimation, whereas an AUC value < 0.5 implies poor results (Can et al. 2005).

3.87.2. Statistical Techniques

Statistical evaluation techniques, such as MAE and RMSE, were selected for this study to validate the models. MAE is the amount sum of difference in the total number of observations between predicted and actual DP values of anythe data sets. RMSE is defined by the square root of MAE (Supplementary material-S3). MAE and RMSE were determined using the following equations:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{predicted} - y_{actual}\right|, \quad (15)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{predicted} - y_{actual}\right)^2}, \quad (16)$$

where, n indicates the total training and test samples, $Y_{predicted}$ is the expected value, and Y_{actual} is the output value. Can et al. (2005) set a cut-off value of 0.5. A value above 0.5 suggests poor results, whereas a value less than 0.5 suggests good performance.

3.87.3. Friedman and Wilcoxon Statistical Signed-rank Tests

The focus of this sub-subsection was to review the results of ensemble ML classifiers via statistical tests on multiple data sets. The classifiers of ML ensembles were tested using the same random samples. The main objective of these tests was to determine which of used methods vary statistically in performance. In this respect, Friedman and Wilcoxon rank tests are suitable because these tests do not presume homogeneity of regular distributions or variance (Tien Bui et al. 2016).

The signed-rank tests of Friedman (2012) and Wilcoxon (1937) were applied in this work to analyse the major differences amongst model outputs. A decision was obtained in consideration of the likelihood of hypotheses (p-value); if the p-value is valid, then the alternative and null hypotheses are denied, and a considerable gap exists amongst the models and vice versa (Tien Bui et al. 2016). The signed-rank Wilcoxon determines the statistical importance of the systematic pairwise variations amongst the DP models. For this test, p-value and z-value were considered to

589 determine the important variations amongst the models. If the p-value is smaller than 0.05 and the
590 z-value reaches the threshold z values (−1.96 and +1.96), then the null alternative hypothesis will
591 be accepted and the results of the DP models will be substantially different (Tien Bui et al. 2016;
592 Chen et al. 2018).

593 **4. Results**

594 **4.1. Relief-F Test and IGR**

595 The IGR and Relief-F approaches were used to examine the relative importance of each of the
596 DDFs for modelling DP probability. IGR and Relief-F were calculated for the training data, as
597 shown in Figure 6 and Table ~~1~~2. The resulting IGR and Relief-F indicated that selected variables
598 ~~provided~~ have good predictive ~~values~~capability. Distance from settlement ~~provided~~ shown the
599 maximum prediction capability; the IGR and Relief-F values were 0.3100 and 0.0922,
600 respectively. Aspect contributed the least predictive value with IGR and Relief-F values of 0.0023
601 and 0.0052, respectively.

602 **Figure 6.** SOMEWHWRE HERE
603 **Table 2~~1~~.** SOMEWHWRE HERE
604
605 **4.2. Frequency of Deforestation in Relation to DDFs**
606 The selected input factors led to a spatial heterogeneity in deforestation process across the study
607 area. The percentage of deforestation samples and FR value in each subclass of DDFs was
608 calculated to understand the ~~different~~ influences of DDFs. The histograms (Figure 7) depict the
609 relationship of deforestation with the different DDFs.
610 For each slope class, deforestation varied (Figure 7a). The maximum deforested samples were
611 identified in the low-slope class (56.7%), followed by those in the moderate-slope class. Similarly,
612 the FR value was highest in the low-slope class, i.e. 1.08. The relationship between deforestation
613 occurrence and aspect was also analysed (Figure 7b). The percentage of deforested samples and
614 FR value (Table ~~2~~3) were maximum for the flat area. For elevation (Figure 7b), the
615 ~~number~~percentage of deforestation pixels was 67% between 17 and 145 m elevation, and it reduced
616 in the high-altitude classes. The FR value was maximum (1.13) for the 79–145 m elevation class.
617 A similar pattern could be observed in TPI (Figure 7d). The highest deforested samples were

observed on flat land (53%). Most of the forest reductions were connected with distance to forest edge. In the first 62 m buffer ring, above 46% of the overall deforested samples were concentrated and within 0.5 km, which was 92% of the samples (Figure 7j). The FR value was also maximum (1.49) for the first buffer ring (0–62 m). A remarkable relationship was found between deforestation occurrence and proximity to the river. The maximum FR value (1.29) was achieved in the 0–156 m buffer ring. The incidence of forest loss decreased with increasing distance from settlement and roads (Figures 7f and 7k). For proximity to settlement and road, 91% and 87% of the total deforested sample pixels were concentrated within 0.5 km. The FR value of the 0.10–0.50 km road buffer ring was the maximum at 2.12, and the 71–142 m settlement buffer ring had the maximum FR value of 1.11 (Table ~~2~~3). Deforestation occurrence was negatively associated with forest density (Figure 7g). The percentage shear of deforestation samples and FR value were highest for the low-forest density class.

A negative association was also found in case of distance to agricultural land (Figure 7m). A high rate of deforestation occurrence (73%) was determined at less than 200 m from agricultural land, and FR value was maximum for the 0–58 m buffer ring. The concentration of deforestation samples and FR values were high in the areas with high settlement (Figure 7i) and agricultural land density (Figure 7l). Figures 7e and 7n reveal that heavy deforestation occurred in areas marked by high population density and fast population growth.

Table ~~2~~3. SOMEWHWRE HERE

Figure 7. SOMEWHWRE HERE

4.3. ~~Construction of DP Models and Maps~~Analysing the deforestation probability

~~DP models utilising hybrid ML ensemble frameworks were developed using training data sets to predict the deforestation in the study area. Deforested and forested pixels were considered the training data sets. Pixels (70%) from both classes were randomly set as training data sets for running the models. The deforestation and non-deforestation were characterised as 0 and 1 codes, respectively. Once all the four models were effectively run in the training phase, the relational weights of the models were applied to compute the DP indices for all pixels. The measuring variables were standardised by training via the trial-and-error method to construct such DP models. In this analysis, we used 5 hidden layers, 0.3 learning rate, 0.2 momentum, 0 seed, 500 training times and 20 validation thresholds for the MLPnn to decide the quantity of data for reduced-error pruning, upgrade weight, add value to the weight, divide the data, build the ensemble and finish~~

~~the calibration testing (Pham et al. 2016; Onan, 2016),. The percentage bag size indicates the training range size (Sedano et al. 2013). Likewise, 16 iterations, 1 seed, 100% of bag size (training range size) and MLPnn as base classifiers were set for bagging. Eighteen iterations, 2 seeds and MLPnn as base classifiers and 8 iterations, 1 seed and principal component analysis as base filters were used.~~

The DP indices of all pixels were ~~created~~ calculated of the~~in~~ total area, and each pixel was allocated with a specific probability index. Probability indices for deforestation were reclassified using a statistical approach. For this analysis, the ~~methodology of~~ geometrical interval was used as a statistical tool to reclassify DP indices. The approach of geometric interval is ideal for classifying continuous data as DP indices whilst minimising variance (Frye, 2007). The DP indices were classified into five probability classes on the basis of this method, namely, very low, low, moderate, high and very high (Figure 8). The outcome of the MLP model indicated that 25.16%, 22.19%, 21.02%, 14.81% and 16.82% of the overall forest area of the basin fell under very low, low, moderate, high and very high DP classes, respectively (Table ~~3~~4). The outcomes of the MLP-RTF model showed that 34.98%, 15.67%, 18.98%, 16.87% and 13.50% of the basin's total forest area fell under very low, low, moderate, high and very high DP classes, respectively. In the MLP-dagging model, very low, low, moderate, high and very high DP classes covered 37.44%, 22.52%, 16.17%, 11.23% and 12.64% of the basin's total forest area, respectively. The land occupied by very low, low, moderate, high and very high PD classes were 33.48%, 19.15%, 17.88%, 16.00% and 13.49%, respectively, in accordance with the MLP-bagging method.

Table 4~~3~~. SOMEWHWRE HERE

Figure 8. SOMEWHWRE HERE

4.4. Validation and Comparison of DP Models

The robustness of the DP models was judged using three threshold-dependent methods (AUC of ROC, precision and accuracy), two threshold-independent methods (MAE and RMSE) and two statistical tests (Friedman and Wilcoxon signed-rank tests). The AUCs showed that the precision of the DP maps reached more than 86% (0.86) for the test and validation data sets (Table 5~~4~~). The MLP-bagging method for training and testing achieved the highest accuracy, followed by MLP-dagging, MLP-RTF and MLPnn. The AUC value~~s~~ of success rate curve (training data) and prediction rate curve (test data)~~for~~ was the highest for the MLP-bagging (0.902 and 0.943) and the

lowest for the MLPnns ~~(, MLP-bagging, MLP-dagging and MLP-RTF were~~ 0.869 and 0.885)~~, 0.921, 0.902 and 0.887~~, respectively~~; the corresponding AUC values of prediction rate curve (test data) were 0.885, 0.943, 0.928 and 0.902~~ (Figure 9). The highest values of precision and ~~;~~ accuracy, ~~MAE and RMSE of MLP~~ were ~~0.77, 0.77, 0.22 and 0.37 by using the training data sets and 0.71, 0.71, 0.29 and 0.43~~obtained by the MLP-bagging and the lowest by the MLPnn, respectively (Table 5) ~~when utilising the test data sets, respectively~~. ~~The values of precision, accuracy, MAE and RMSE of MLP-bagging were 0.87, 0.85, 0.22 and 0.29 for the training data sets and 0.87, 0.80, 0.19 and 0.34 for the test data, respectively. The values of precision, accuracy, MAE and RMSE of MLP-dagging were 0.82, 0.80, 0.22 and 0.29 for the training data sets and 0.84, 0.76, 0.24 and 0.38 for the validation data set, respectively. The values of precision, accuracy, MAE and RMSE of MLP-RTF were 0.77, 0.77, 0.20 and 0.32 by using the training data and 0.79, 0.74, 0.27 and 0.37 by using the validation data, respectively.~~ The values of statistical measures, i.e. MAE and RMSE, were calculated in consideration of the training and validation data sets. The l~~L~~owest values (0.24 and 0.38) were obtained for the MLP-bagging ensemble model. On the other hand, the highest values (0.29 and 0.43) were obtained by the MLPnn model.

Therefore, from the validation results, it was found that the accuracy of the MLP model was improved after combining with the selected three meta classifiers. On an average AUC of prediction and success rate curves was increased by 3%. The highest increase of AUC values of both curves were found in the MLP-Bagging ensemble modes i.e. 5.4% (in success rate curve) and 5.8% (In prediction rate curve) respectively. However, as per the results of ROC, precision, accuracy, MAE and RMSE, the robustness level of the MLP-bagging model was higher than those of the other MLPnn and ensemble models.

Figure 9. SOMEWHWRE HERE

Table 45. SOMEWHWRE HERE

Friedman and Wilcoxon signed-rank tests were used to ascertain the DP models. The results of the Friedman test are presented in Table ~~5~~6. The mean ranking values for the MLPnn, MLP-bagging, MLP-dagging and RFB-RTF models were 2.77, 2.22, 2.42 and 2.48, respectively.

The signed-rank test of Wilcoxon was applied to determine the gaps in pairs amongst the ML models at a relevance level of 5% (Table ~~6~~7). When p (value) < 5% (0.05) and z (value) > z (−1.96

and +1.96), the capabilities of the models in the Wilcoxon rank test varied substantially [106]. Analysis suggested (Table ~~6~~7) a substantial disparity amongst all DP models.

Table ~~5~~6. SOMEWHWRE HERE

Table ~~6~~7. SOMEWHWRE HERE

5. Discussion

The changes in the forest cover of the Gumani River Basin are well recognised, with numerous factors primarily focused on institutional, financial and economic ~~forces~~ aspects (Vanonckelen et al. 2015), the low performance of protected areas (Bălteanu et al. 2016) and environmental disruptions (Săvulescu et al. 2011). The estimated evaluations for DP are limited, with only a few works assessing the relative impacts of biophysical, socio-demographic and land use approaches on the changes in the forest cover at temporal scales (Munteanu et al. 2015; Vanonckelen et al. 2015). Thus, we measured the future possibility of deforestation across the Gumani River Basin in this study by using hybrid ensemble frameworks, MLP-bagging, MLP-dagging and MLP-RTF. In this analysis for preparing the DP models first, hybrid ensemble methods were used to optimize the input data using training dataset. Thereafter, optimized input data were used to categorize classes for spatial DP considering the MLPnn base classifier (Roy et al. 2020). Ultimately, frameworks of the machine learning ensemble were developed for the DP models. The results of training sets of DP were used for the creation of DP maps. Ensemble approaches are classification methods for data processing, whilst MLPnns are regarded as ANNs with excellent results in the spatial modelling of deforested areas.

The findings of this study indicated that all probability models of deforestation utilising hybrid ML ensemble increased the efficiency of the MLPnn (AUC=0.869) base classifier. This result is reasonable because DP models using hybrid ML ensemble systems are well recognised to be very successful in enhancing the efficiency of base classifiers. The DP models in this analysis produced a satisfactory result and allowed basic performance indicators (such as accuracy, precision, AUC, RMSE, MAE and Friedman and Wilcoxon signed-rank tests) to be used to evaluate the models. The outcomes produced through the ensemble modes ~~were shown the~~showed a better accuracy than the previously used individual model for the mapping the probability of deforestation (Sahana et al. 2018; Kumar et al. 2014; Bavaghar, 2015; Kucsicsa et al. 2019; Saha et al.~~,~~ 2020; Dlamini, 2016; Krüger and Lakes~~;~~ 2015; Mayfield et al. 2017). Due to the less error and very low overfitting problem, the ensemble methods provided better results than previous works done by the different

scholars (Roy et al. 2020). The quantity or overall area of deforestation is helpful for planning or zoning, but the models could not be used for measurement. Another drawback of the used models is that the assumed predictors of deforestation do not alter with time. This drawback is common amongst many ML models, but it is especially applicable to our models because deforestation predictors were chosen on the basis of predisposing risk factors for deforestation (Geist et al. 2001; Mas et al. 2004). Despite these drawbacks, the findings showed that data sets that are publicly accessible could be considered to estimate the DP within the research area.

DP models utilising ensemble frameworks were compared. The results of the evaluation of the DP maps were obtained using ROC, efficiency, accuracy, MAE, RMSE and two statistical tests, i.e. Friedman and Wilcoxon signed-rank tests (Tables 45–76). The results showed that MLP-bagging considerably outperformed the other models. MLP-bagging (AUC=0.943) had the strongest predictive capacity, followed by MLP-dagging (AUC=0.928), MLP-RTF (AUC=0.884) and MLP models (AUC=0.902). MLP-bagging is more efficient in mitigating volatility and discrimination compared with other ensemble approaches (Pham et al. 2016; Sedano et al. 2013). Feature selection approach is widely used to test the predictive capacity of variables to improve model performance by eliminating unwanted or unimportant factors in advance (Pham et al. 2016a). The Relief-F and IGR methods were utilised in this analysis for selecting and judging the predictive potentiality of different DDFs for DP models. On the basis of these methods, the distance to settlement and the distance to road and population growth showed the strongest influences on DP models because most of deforested locations were identified on or along road and settlement. The remaining factors, such as forest density, distance to forest edge, proximity to river, population density, agricultural land density, distance to agricultural land, density of settlement, altitude, slope and aspect, also indicated good contributions to DP models, as confirmed in other similar studies (Sahana et al. 2018). A relative difference of nearly 3% was determined from the comparison results of the DP models on the basis of the ROC curve, but it was substantial for the DP maps (Table 45). Therefore, even minor changes in the efficiency of DP models would contribute to increased change in the reliability of DP maps. Furthermore, the efficiency of such probability models for deforestation depends greatly on optimising the predictive parameters.

The output of this research might help researchers to analyse deforestation in other areas. Hybrid ensemble approaches could also be used to assess data and serve as reliable alternatives to

conventional computational strategies for modelling DP. The use of soft computing approaches would inspire the scientific communities to use sophisticated techniques for precisely modelling probable deforestation areas. In populated countries, such as India, this work would assist the policymakers in making strategic plans for managing the existing forest cover.

6. Conclusions

In this research, hybrid ensemble frameworks, MLP-bagging, MLP-dagging and MLP-RTF, were effectively implemented for the analysis of DP of the Gumani River Basin. ROC, accuracy, precision, MAE, RMSE and Friedman and Wilcoxon signed-rank tests were used to validate and compare four DP models. The findings indicated that DP models utilising ML ensemble systems worked well in this study, and substantial differences existed amongst the models. ~~The MLP-bagging model achieved the maximum predictive efficiency, followed by MLP-dagging, MLP-RFT and MLP. The efficiency of the base classifier of MLPnn was increased greatly by the architectures of the hybrid ensemble.~~

- Among the MLPnn, MLP-Bagging and MLP-Dagging model, the MLP-Bagging model produced the best performance in terms of accurateness (efficiency, accuracy and AUC) and reliability (RMSE and MAE). It may be concluded that to prepare ~~th~~an accurate ~~precise~~ deforestation probability map, MLP-Bagging model can be very effective.

- After ensemble of meta-~~-~~classifiers with the base classifier, the accuracy of the MLPnn model was increased significantly.

- Delineating deforestation probability areas by means of field based methods are ~~very~~ expensive and time-~~-~~consuming, especially for the large watersheds. Therefore, ~~as a very contemporary device~~as an alternative, application of ensemble machine learning models along with RS-GIS based data and interfaces could be very effective in creating deforestation probability map.

- Finally, the produced deforestation probability maps for the Gumani River basin displayed the areas having high and very-~~-~~high probability of deforestation, which could be an effective device for policymakers and environmental planners.

This research indicated that the ML models are powerful ~~devices~~ techniques that can be used for the DP evaluation of an area. The adequate precision acquired by the ensemble models and validation methods confirmed that the models have acceptable precision. The results would also provide spatial evidence to execute appropriate policies and strategies for forest managers and environmental planners. In fact, the deforestation process is closely correlated with certain natural and anthropogenic factors. The findings might be valuable for deforestation predictions in other regions having similar geo-environmental conditions. Furthermore, the findings would provide a foundation for future research. Existing DDFs might be combined with other DDFs, modified as per changes in the physical or socio-economic context of the Gumani River Basin, to enable for an improved and realistic simulation of DP.

Conflicts of Interest: The authors declare that there is no conflict of interest.

Funding: This research was supported by the Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and Information Technology, University of Technology Sydney. This APC was funded by Universiti Kebangsan Malaysia, DANA IMPAK PERDANA with grant no: DIP-2018-030. It was also supported by Researchers Supporting Project (number RSP-2020/14), King Saud University, Riyadh, Saudi Arabia.

Author Contributions: S.S. contributed in the methodology development, formal analysis, investigation, original draft preparation and manuscript review and editing; S.S. and G.P. performed the experiments, wrote the manuscript and collected the field data; S.S. wrote the manuscript and analysed the data; B.P. edited, restructured and professionally optimised the manuscript; B.P. and A.A. arranged the funding acquisition. All authors have read and agreed to the published version of the manuscript.

References

Altun H, Polat E, Polat G, Güneú T. 2007. Identifying and combining multi-modal biometric features from voice and facial image signs to improve human computer interaction, Tübitak Research Project Report., 104E179: 42- 50.

Arekhi M. 2011. Modeling spatial pattern of deforestation using GIS and logistic regression: A case study of northern Ilam forests, Ilam province, Iran. *African Jjournal of Bbiotechnology*, *10*(72), 16236-16249.

Bălteanu D, Năstase M, Dumitrașcu M, Grigorescu I. 2016. Environmental Changes in the Maramureş Mountains Natural Park. In *Sustainable Development in Mountain Regions*, 335-348. Springer, Cham.

Bavaghar MP. 2015. Deforestation modelling using logistic regression and GIS. *Journal of Forest Science 61*(5), 193-199.

Bavaghar PM. *2015*. Deforestation modelling using logistic regression and GIS. *Journal of Forest Science*., 61(5): 193-199. https://doi.org/10.17221/78/2014-jfs

Bax V, Francesconi W, Quintero M. 2016. Spatial modeling of deforestation processes in the Central Peruvian Amazon. *Journal for Nature Conservation*, *29*, 79-88. https://doi.org/10.1016/j.jnc.2015.12.002

Bayat MF. 2000. Surveying of the relationship between vegetation cover and some environmental variables (altitude, aspect and slope). *Pajouhesh-va-Sazandegi*, *4*(45), 24-27.

Benediktsson J, Swain PH, Ersoy OK. (1990.) Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, . IEEE Trans 28(4): 540–552

Boudreau S, Lawes MJ, Piper SE, Phadima LJ. 2005. Subsistence harvesting of pole-size understorey species from Ongoye Forest Reserve, South Africa: species preference, harvest intensity, and social correlates. *Forest Ecology and Management*, *216*(1-3), 149-165.

Bouldin J. 2008. Some problems and solutions in density estimation from bearing tree data: a review and synthesis. *Journal of Biogeography*, *35*(11), 2000-2011.

Breiman L. 1996. Bagging predictors. *Machine Llearning*, *24*(2), 123-140.

Buchanan GM, Butchart SH, Dutson G, Pilgrim JD, Steininger MK, Bishop KD, Mayaux P. 2008. Using remote sensing to inform conservation status assessment: estimates of recent deforestation rates on New Britain and the impacts upon endemic birds. *Biological Conservation*, *141*(1), 56-66.

Bui DT, Tuan TA, Klempe H,; Pradhan B, Revhaug I. 2016. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*, *13*(2), 361-378.

Cai H, Ng M. 2012. Feature weighting by RELIEF based on local hyperplane approximation. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (335-346). Springer, Berlin, Heidelberg.

Can T, Nefeslioglu HA, Gokceoglu C, Sonmez H, Duman TY. 2005. Susceptibility assessment of shallow earthflows triggered by heavy rainfall at three sub catchments by logistic regression analyses. *Geomorphology*, 72, 250–271.

Chandniha SK, -Meshram SG, Adamowski JF, Meshram C. 2017. Trend analysis of precipitation in Jharkhand State, India. *Theoretical and Applied Climatology*, 1;130 (1-2):261-74.

Chomitz K., Gray DA. 1999. *Roads, lands, markets, and deforestation: a spatial model of land use in Belize*. The World Bank.

Congalton RG. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment,* 37:35–46

Cropper M, Griffiths C. 1994. The interaction of population growth and environmental quality. *The American Economic Review*, *84*(2), 250-254.

D'Arco M, Liccardo A, Pasquino N. 2012. Anova-based approach for dac diagnostics. *EEE Transactions on Instrumentation and Measurement*IEEE Trans. Instrum. Meas., 61, 1874–1882.

DeFries RS, Rudel T, Uriarte M, Hansen M. 2010. Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geoscience*, *3*(3), 178-181.

Deng JS Wang K., Deng YH, Qi GJ, 2008. PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing*, 29, 4823–4838.

Dlamini WM. 2016. Analysis of deforestation patterns and drivers in Swaziland using efficient Bayesian multivariate classifiers. *Model. Earth Syst. Environ*Modelling Earth Systems and Environment, 2, 1–14.

Ercanoglu M, Gokceoglu C. 2002. Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzzy approach. *Environmental Ggeology*, *41*(6), 720-730.

Fang Z, Wang Y, Peng L, Hong H. 2020. A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. *International Journal of Geographical Information Science*, 16:1-27.

Friedman M. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J. Am. Stat. Assoc., 32, 675–701.

Frye C. 2007. About the geometrical interval classification method. *Environmental Systems Research Institute, Inc. Online verfügbar unter https://blogs. esri. com/esri/arcgis/2007/10/18/about-thegeometrical-interval-classification-method*.

Gardner MW, Dorling SR. 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric Eenvironment*, *32*(14-15), 2627-2636.

Gaveau DL, Epting J, Lyne O, Linkie M, Kumara I, Kanninen M, Leader-Williams N. 2009. Evaluating whether protected areas reduce tropical deforestation in Sumatra. *Journal of Bbiogeography*, *36*(11), 2165-2175.

Gayen A, Pourghasemi HR, Saha S, Keesstra S, Bai S. 2019. Gully erosion susceptibility assessment and management of hazard-prone areas in India using different machine learning algorithms. *Science of the total environment*, *668*, 124-138.

Gayen A, Saha S. 2018. Deforestation probable area predicted by logistic regression in Pathro river basin: a tributary of Ajay River. *Spatial Information Research*, *26*(1), 1-9.

Geist HJ, Lambin EF. 2001. What drives tropical deforestation. *LUCC Report series*, *4*, p.116.

Glade T. 2003. Landslide occurrence as a response to land use change: a review of evidence from New Zealand. *Catena*, *51*(3-4), 297-314.

Gong P. 2009. Integrated analysis of spatial data for multiple sources: using evidential reasoning and artificial neural network techniques for geological mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, Photogram *Eng Rem Sens*. 1996, 62:513–523

Hosonuma N, Herold M, De Sy V, De Fries RS, Brockhaus M, Verchot L, Angelsen A, Romijn E. 2012 An assessment of deforestation and forest degradation drivers in developing countries. *Environmental Research Letters*, 8;7(4):044009.

Houet T, Hubert-Moy L. 2006. Modeling and projecting land-use and land-cover changes with Cellular Automaton in considering landscape trajectories.

Jennes J. 2006. Topographic Position Index. tpi jen.avx, extension for ArcView 3.x; v.1.3a. Jenness Enterprises. http://www.jennessent.com/arcview/tpi.htm.

Kaim D, Radeloff VC, Szwagrzyk M, Dobosz M, Ostafin K. 2018. Long-term changes of the wildland–urban interface in the Polish Carpathians. *ISPRS International Journal of Geo-Information*, *7*(4), 137.

Khosravi K., Pham BT, Chapi K, Shirzadi A, Shahabi H, Revhaug I, Prakash I, Bui DT. 2018. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. *Science of the Total Environment*, *627*, 744-755.

Kira K Rendell LA. 1992. A practical approach to feature selection. In *Machine Learning Proceedings*, (-249-256). Morgan Kaufmann.

Kotsianti SB, Kanellopoulos D. 2007. Combining bagging, boosting and dagging for classification problems. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, 493-500). Springer, Berlin, Heidelberg.

Krüger C, Lakes T. 2015. Bayesian belief networks as a versatile method for assessing uncertainty in 621 land-change modeling. *International JJournal of Geographical Information Science*, 29, 111–131.

Kucsicsa G, Dumitrică C. 2019. Spatial modelling of deforestation in Romanian Carpathian Mountains using GIS and Logistic Regression. *Journal of Mountain Science*, *16*(5), 1005-1022.

Kumar R, Indrayan A. 2011. Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatrics*~~Pediatr.~~, 48 (4), 277–28.

Kumar R, Nandy S, Agarwal R, Kushwaha SPS. 2014. Forest cover dynamics analysis and prediction modeling using logistic regression model. *Ecological Indicators*, *45*, 444-455.

Kuncheva~~,~~ L~~.~~I~~,.;~~ Rodrı´guez~~,~~ J~~.~~J. 2007. An experimental study on rotation forest ensembles. *In International workshop on multiple classifier systems*, *Springer*, 459–468.

Lambin EF, Turner BL, Geist HJ, Agbola SB, Angelsen A, Bruce JW. 2001. The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental* ~~*Chang*~~*Change*~~,.;~~ 11: 261–269.

Lee S, Pradhan B. *2006*. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides*, 4:33–41.

Liu H, Motoda H. 2008. Computational methods of feature selection. *Chapman and Hall/CRC Press*.

Ludeke A, Maggio RC, Reid L. 1990. An analysis of anthropogenic deforestation using logistic regression and GIS. *Journal of*~~.~~ *Environmental*~~.~~ *Management*~~.~~, 31, 247–259.

Martínez-Álvarez F, Reyes J, Morales-Esteban A, Rubio-Escudero C. 2013. Determining the best set of seismicity indicators to predict earthquakes. Two case studies: Chile and the Iberian Peninsula. *Knowledge-Based Systems*, *50*, 198-210.

Mas JF Puig H, Palacio JL, Sosa-Lopez A. 2004. Modelling deforestation using GIS and artificial neural networks. *Environmental Modelling & Software*, *19*(5), 461-471.

Matlack GR. 1994. Vegetation dynamics of the forest edge--trends in space and successional time. *Journal of Ecology*, 113-123.

Maya Liyana Hamzah, Ahmad Aldrie Amir, Khairul Nizam Abdul Maulud, Sahadev Sharma, Fazly Amri Mohd, Siti Norsakinah Selamat, Othman A. Karim, Effi Helmy Ariffin, And Rawshan Ara Begum. 2020. Assessment of the Mangrove Forest Changes along the Pahang Coast using Remote Sensing and GIS Technology~~.,~~ *Journal of Sustainability Science and Management*, ~~Volume~~ 15 (5), ~~pp.~~ 43-58.

Mayfield H, Smith C, Gallagher M, Hockings M. 2017. Use of freely available datasets and machine 649 learning methods in predicting deforestation. ~~*Environ*~~*Environmental*~~.~~ ~~*Model*~~*Modelling &*~~.~~ *Softwa*~~*re*~~.~~, 87, 17–28.

Michalski F, Peres CA, Lake IR. 2008. Deforestation dynamics in a fragmented region of southern Amazonia: evaluation and future scenarios. *Environmental Conservation*, *35*(2), 93-103.

Millennium Ecosystem Assessment ME. Ecosystems and human well-being. Synthesis. 2005 Aug 27.

Minetos D, Polyzos S. 2010. Deforestation processes in Greece: A spatial analysis by using an ordinal regression model. *Forest Policy and Economics*, 12(6): 457-472.

971 Munteanu C, Kuemmerle T, Boltiziar M, Butsic V, Gimmi U, Halada L, Kaim D, Király G,
972 Konkoly-Gyuró É, Kozak J, Lieskovský J. 2014. Forest and agricultural land change in the
973 Carpathian region—A meta-analysis of long-term patterns and drivers of change. *Land Use*
974 *Policy*, *38*, 685-697.

975 Munteanu C, Kuemmerle T, Keuler NS, Müller D, Balázs P, Dobosz M, Griffiths P, Halada L,
976 Kaim D, Király G, Konkoly-Gyuró É. 2015. Legacies of 19th century land use shape
977 contemporary forest cover. *Global Environmental Change*, *34*, 83-94.

978 Nackaerts K, Vaesen K, Muys B, Coppin P. 2005. Comparative performance of a modified change
979 vector analysis in forest change detection. *International Journal of Remote Sensing*, 26,
980 839–852.

981 Nandy S, Kushwaha SPS, Mukhopadhyay S. 2007. Monitoring the Chilla–Motichur wildlife
982 corridor using geospatial tools. *Journal for Nature Conservation*, *15*(4), 237-244.

983 Newman ME, McLaren KP, Wilson BS. 2014. Assessing deforestation and fragmentation in a
984 tropical moist forest over 68 years; the impact of roads and legal protection in the Cockpit
985 Country, Jamaica. *Forest Ecology and Management*, *315*, 138-152.

986 Onan –A. 2016. Classifier and feature set ensembles for web page classification. *Journal of*
987 *Information* 643 *Science*, *42*(2), pp.150-165.

988 Ortega Adarme M, Queiroz Feitosa R, Nigri Happ P, Aparecido De Almeida C, Rodrigues Gomes
989 A. 2020. Evaluation of Deep Learning Techniques for Deforestation Detection in the Brazilian
990 Amazon and Cerrado Biomes from Remote Sensing Imagery. *Remote Sensing*, 12(6):910.

991 Pepe MS. 2000. Receiver operating characteristic methodology. *J. Am. Stat. Assoc.*, 95, 308–311.

992 Pham BT, Pradhan B, Tien Bui D, Prakash I, Dholakia MB. 2016a. A comparative study of
993 different machine learning methods for landslide susceptibility assessment: a case study of
994 Uttarakhand area (India). *Environmental Modelling & Software*, 84, 240-250.

995 Pham BT. Bui DT. Prakash I, Dholakia MB. 2016. Rotation forest fuzzy rule-based classifier
996 ensemble for spatial prediction of landslides using GIS. *Natural Hazards*, *83*(1), 97-127.

997 Pontius Jr RG, Schneider LC. 2001. Land-cover change model validation by an ROC method for
998 the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems &*
999 *Environment*, *85*(1-3), 239-248.

1000 Quinlan JR. 1993. C4.5: programs for machine learning. Morgan Kaufmann, San Mateo, CA,
1001 USA.

1002 Rahmati O, Naghibi SA, Shahabi H, Bui DT, Pradhan B, Azareh A, Rafiei-Sardooi E, Samani AN,
1003 Melesse AM. 2018. Groundwater spring potential modelling: Comprising the capability and
1004 robustness of three different modeling approaches. *Journal of Hydrology*, *565*, 248-261.
1005 https://doi.org/10.1016/j.jhydrol.2018.08.027.

Robinson BE, Holland MB, Naughton-Treves L. 2014. Does secure land tenure save forests? A meta-analysis of the relationship between land tenure and tropical deforestation. *Global Environmental Change*, *29*, 281-293.

Rodriguez JJ, Kuncheva LI, Alonso CJ. 2006. Rotation forest: A new classifier ensemble method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(10), 1619– 1630.

Roy J, Saha S, Arabameri A, Blaschke T, Bui DT. 2019. A Novel Ensemble Approach for Landslide Susceptibility Mapping (LSM) in Darjeeling and Kalimpong Districts, West Bengal, India. *Remote Sensing*, *11*(23), 2866.

Roy, J, Saha, S, 2020a. Integration of artificial intelligence with Meta classifiers for the gully erosion susceptibility assessment in Hinglo river basin, Eastern India. *Advances in Space Research*. https://doi.org/10.1016/j.asr.2020.10.013

Roy P, Chakrabortty R, Chowdhuri I, Malik S, Das B, Pal SC. 2020. Development of Different Machine Learning Ensemble Classifier for Gully Erosion Susceptibility in Gandheswari Watershed of West Bengal, India. In Machine Learning for Intelligent Decision Science, (pp. 1-26). Springer, Singapore.

Saha AK, Gupta RP, Arora MK, 2002. GIS-based landslide hazard zonation in the Bhagirathi (Ganga) valley, Himalayas. *International Jjournal of Rremote Ssensing*, *23*(2), 357-369.

Saha S, Saha A, Hembram TK, Pradhan B, Alamri AM. 2020. Evaluating the pPerformance of iIndividual and nNovel cEnsemble of mMachine lLearning and sStatistical mModels for lLandslide sSusceptibility aAssessment at Rudraprayag dDistrict of Garhwal Himalaya. *Applied Sciences*, 10(11),; 3772.

Saha S, Saha M, Mukherjee K, Arabameri A., Ngo PTT, Paul GC, 2020. Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest and REPTree: A case study at the Gumani River Basin, India. *Science of The Total Environment*, p.139197.

Sahana M, Hong H, Sajjad H, Liu J, Zhu AX. 2018. Assessing deforestation susceptibility to forest ecosystem in Rudraprayag district, India using fragmentation approach and frequency ratio model. *Science of the Total Environment*, 627, 1264-1275.

Săvulescu I, Mihai B. 2011. Geographic information system (GIS) application for windthrow mapping and management in Iezer Mountains, Southern Carpathians. *Journal of Forestry Research*, 23(2): 175-184. https://doi.org/10.1007/s11676-011-0213-5.

Sedano J, González S, Herrero Á, Baruque B, Corchado E. 2013. Mutating network scans for the assessment of supervised classifier ensembles. *Logic Journal of the IGPL*, *21*(4), 630-647.

Siles NJS, 2009. Spatial Modelling and prediction of tropical forest conversion in the Isiboro Secure National Park and Indigenous Territory (TIPNIS), Bolivia. (M.Sc. Thesis). International Institute for Geoinformation Science and Earth Observation, Enschede, The Netherlands.

Siti Nor Maizah Saad SNM,, Khairul Nizam Abdul Maulud KNA, Wan Shafrina Wan Mohd Jaafar WSWM, Aisyah Marliza Muhmad Kamarulzaman AMM, Hamdan Omar H., 2020. Tree Stump Height Estimation Using Canopy Height Model at Tropical Forest in Ulu Jelai Forest Reserve, Pahang, Malaysia. *IOP Conf. Series: Earth and Environmental Science,* 10th IGRSM International Conference and Exhibition on Geospatial & Remote Sensing 20-21 October 2020, Kuala Lumpur, Malaysia, 540 (2020) 012015.

Sobala M, Rahmonov O, Myga-Piatek U. 2017. Historical and contemporary forest ecosystem changes in the Beskid Mountains (southern Poland) between 1848 and 2014. *iForest-Biogeosciences and Forestry*, *10*(6), 939.

Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications*, 32, 1084–1093.

Sun J, Southworth J. 2013. Remote sensing-based fractal analysis and scale dependence associated with forest fragmentation in an Amazon tri-national frontier. *Remote Sensing*, *5*(2), 454-472.

Süzen, M.L.,; Doyuran, V. 2004. A comparison of the GIS based landslide susceptibility assessment methods: multivariate versus bivariate. *Environmental Ggeology*, *45*(5), 665-679.

Szymura, TH, Murak S, Szymura M. 2018. Changes in forest cover in Sudety Mountains during the last 250 years: patterns, drivers, and landscape-scale implications for nature conservation. *Acta Societatis Botanicorum Poloniae*., 87(1). https://doi.org/10.5586/asbp.3576

Tien Bui D, Pham BT, Nguyen QP, Hoang ND. 2016. Spatial prediction of rainfall-induced shallow landslides using hybrid integration approach of Least-Squares Support Vector Machines and differential evolution optimization: a case study in Central Vietnam. *International Journal of Digital Earth*, 1–21.

Tien Bui D, Shahabi H, Shirzadi A, Chapi K Pradhan B, Chen W, Khosravi K, Panahi M, Bin Ahmad B, Saro L. 2018. Land subsidence susceptibility mapping in Ssouth koreaKorea using machine learning algorithms. *Sensors*, *18*(8), 2464.

Tien Bui D, Shirzadi A, Chapi K, Shahabi H, Pradhan B, Pham BT, Singh VP, Chen W, Khosravi K, Bin Ahmad B, Lee S. 2019. A hybrid computational intelligence approach to groundwater spring potential mapping, *Water*, 11(10):2013.

Tien Bui D, Tuan TA, Klempe H, Pradhan B, Revhaug I. 2016a. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides*, 13, 361–378.

Ting KM, Witten IH. 1997. Stacking bagged and dagged models.

Turner MG, Gardner RH, O'Neill RV. 2001. Landscape ecology in theory and practice: Pattern and process. Springer, New York.

Vanonckelen S, van Rompaey A. 2015. Spatiotemporal analysis of the controlling factors of forest cover change in the Romanian ~~carpathian mountains~~Carpathian Mountains. *Mountain Research and Development*~~.~~, 35(4): 338-350. https://doi.org/10.1659/mrd-journal-d-15-00014

Wahab NA, Kamarudin MK, Toriman ME, Juahir H, Saad M, Ata FM, Ghazali A, Hassan AR, Abdullah H, Maulud KN, Hanafiah MM, Harith H. 2019. Sedimentation and water quality deterioration problems at Terengganu river basin, Terengganu, Malaysia. *Desalination and Water Treatment,* 149 (2019), 228-241.

Wan Mohd Jaafar W S, Maulud KN, Kamarulzaman AM, Raihan A, Md Sah S, Ahmad A, Saad SNM, Mohd Azmi AT, Jusoh Syukri NKA, Khan WR. 2020. The influence of deforestation on land surface temperature- a case study of Perak and Kedah, Malaysia. *Forests*, 11(6), 670.

Wan Mohd Jaafar~~,~~ W~~.~~S~~.,~~; Woodhouse~~,~~ I~~.~~H~~.,~~; Silva~~,~~ C~~.~~A~~.,~~; Omar~~,~~ H~~.,~~; Abdul Maulud~~,~~ K~~.~~N~~.,~~; Hudak~~,~~ A~~.~~T~~.,~~; Klauberg~~,~~ C~~.,~~; Cardil~~,~~ A~~.,~~; Mohan~~,~~ M. 2018. Improving ~~i~~Individual ~~t~~Tree ~~c~~Crown ~~d~~Delineation and ~~a~~Attributes ~~e~~Estimation of ~~t~~Tropical ~~f~~Forests ~~u~~Using ~~a~~Airborne LiDAR ~~d~~Data. *Forests*~~,~~ 9~~.~~(12), 759.

Wang G, Oyana T, Zhang M, Adu-Prah S, Zeng S, Lin H, Se J. 2009. Mapping and spatial uncertainty analysis of forest vegetation carbon by combining national forest inventory data and satellite images. *Forest Ecology and Management*, *258*(7), 1275-1283.

Weier J, Herring D. 2000. Measuring Vegetation (NDVI & EVI) Earth Observatory, NASA.

Weiss A. Topographic Position and Landforms Analysis. Poster presentation. *ESRI User Conference*, San Diego, CA. 2001.

Wilson K., Newton A., Echeverría C. 2005. A vulnerability analysis of the temperate forests of south central Chile. *Biological Conservation*, 122(1): 9-21.https://doi.org/10.1016/j.biocon.2004.06.015

Witten DM, Tibshirani R. 2011. Penalized classification using Fisher's linear discriminant. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(5), 753-772.

Wu Y, Ke Y, Chen Z, Liang S, Zhao H, Hong H. 2020. Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping. *Catena*~~,~~ 1;187:104396.

Xia J, Du P, He X, Chanussot J. 2014. Hyper spectral re- mote sensing image classification based on rotation forest. *IEEE Geoscience and Remote Sensing Letters*, *11*(1), 239–243.

Yalcin A. 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. *Catena*, *72*(1), 1-12.