PREDICTION OF SUSCEPTIBILITY FOR OLD TREES (> 100 YEARS OLD) TO FALL IN BOGOR BOTANICAL GARDEN

Faozan Indresputra[1], Rizmoon N. Zulkarnaen[2], Muhammad R. Hariri[2], Fitri F. Wardani[2], Prima Wahyu K. Hutabarat[2], Dwi Setyanti[3], Widya A. Pratiwi[4], Lutfi Rahmaningtiyas[3], and Dina Safarinanugraha[4]

[1]Directorate of Development Policy for Environment, Maritime, Natural Resources and Nuclear Energy, National Research and Innovation Agency, Indonesia
[2]Research Center for Plant Conservation, Botanical Garden and Forestry, National Research and Innovation Agency, Indonesia
[3]Directorate of Scientific Collection Management, National Research and Innovation Agency, Indonesia
[4]Directorate of Research and Innovation Partnership, National Research and Innovation Agency, Indonesia

Received: 25 March 2021, Revised: 15 September 2022, Accepted: 21 October 2022

PREDICTION OF SUSCEPTIBILITY FOR OLD TREES (> 100 YEARS OLD) TO FALL IN BOGOR BOTANICAL GARDEN. Since the establishment of the Bogor Botanical Garden (BBG) in 1817, the protection of the tree collections, even the loss of aging trees (> 100 years old), has been one of its most important tasks. Abiotic factors such as intense extreme events, i.e., heavy rainfall and strong winds, as well as biotic factors from human activities, pests and diseases, and the deterioration of the health of the plant collection with age, has threatened the survival of the old tree collections. As the BBG has many functions for conservation and human ecological activities, tree fall accidents have become a primary concern in preventing the loss of biodiversity and human life. Therefore, disaster map zonation is required to prevent and minimize such accident together with a prediction of which individual specimen is likely to fall. We examined the health of 154 to determine the falling probability of 1106 aged trees based on several factors that might cause the fall in the past and to make model predictions generated by nine supervised machine learning algorithms to get a binary value of falling probability and then classified into four categories (neglectable, low, moderate, and high probability of falling). Inverse Distance Weighted interpolation method was used to depict a zone map of trees prone to fall in BBG. We found 885 susceptible trees, of which 358 individual trees were highly susceptible to fall (red zone color), dominated by families from Fabaceae, Lauraceae, Moraceae, Meliaceae, Dipterocarpaceae, Sapindaceae, Rubiaceae, Myrtaceae, Araucariaceae, Malvaceae, and Anacardiaceae. This result was based on Random Forest model due to its highest accuracy among algorithms and its lowest *false negative* (FN) value. The FN value was important to minimize error calculation on aged trees that were not prone to fall but turned out to be prone to fall. The dominant factor contributing to high falling intensity was hollow and brittle on the tree trunks where many were found to have pests inside damaged parts such as termites, wood-borers, and bark-eaters. Several trees were found to have combined damages with more than a single causative factor that exacerbated tree's health and increased falling probability.

Keywords: Aged trees, 100 years old, probability to fall, model predictions

PREDIKSI KERENTANAN KOLEKSI POHON (>100 TAHUN) YANG RAWAN TUMBANG DI KEBUN BOGOR. Sejak berdiri pada tahun 1817, perlindungan koleksi pohon, termasuk pohon tua (> 100 tahun), telah menjadi salah satu tugas terpenting Kebun Raya Bogor. Faktor abiotik seperti iklim ekstrem seperti hujan lebat dan angin kencang yang terjadi terlalu sering dan faktor biotik dari hama dan penyakit serta kegiatan aktivitas wisata dan faktor usia yang memang menurunkan kesehatan tanaman dapat mengancam kelangsungan hidup koleksi pohon tua. Karena BBG memiliki banyak fungsi tidak hanya untuk konservasi tetapi juga untuk kegiatan ekologi manusia, kecelakaan akibat pohon

* Corresponding author: aozan@mail.ugm.ac.id

tumbang menjadi perhatian utama untuk mencegah hilangnya keanekaragaman hayati dan keselamatan pengunjung. Oleh karena itu, prediksi terkait specimen pohon mana yang mungkin tumbang dan adanya peta zonasi kerawanan tumbang diperlukan untuk mencegah dan meminimalkan kejadian tersebut. Kami memeriksa kesehatan 154 pohon untuk meprediksi kemungkinan tumbang dari 1106 pohon tua berdasarkan beberapa faktor yang menyebabkan tumbang di masa lalu dan untuk membuat prediksi model yang dihasilkan oleh sembilan algoritme supervised machine learning untuk mendapatkan nilai biner terkait probabilitas kerawanan tumbang dan kemudian diklasifikasikan menjadi empat kategori (dapat diabaikan, rendah, sedang, kemungkinan jatuh tinggi). Metode interpolasi Inverse Distance Weighted digunakan untuk menggambarkan peta zona pohon rawan tumbang. Hasil prediksi menunjukkan bahwa ditemukan 885 pohon rawan tumbang dengan 358 individu diperkirakan tumbang dengan probabilitas tinggi (warna zona merah) yang didominasi dari famili Fabaceae, Lauraceae, Moraceae, Meliaceae, Dipterocarpaceae, Sapindaceae, Rubiaceae, Myrtaceae, Araucariaceae, Malvaceae, dan Anacardiaceae. Hasil ini diperoleh dari model Random Forest yang memiliki akurasi tertinggi dengan nilai negatif palsu (FN) terendah. Nilai FN sangat penting untuk meminimalkan kesalahan perhitungan pada pohon tua yang tidak rawan tumbang tetapi ternyata rawan tumbang. Faktor dominan penyebab tingginya intensitas pohon yang tumbang adalah dikarenakan batang pohon yang berlubang dan rapuh dimana banyak ditemukan hama di daerah luka seperti rayap, penggerek kayu dan pemakan kulit kayu. Beberapa pohon ditemukan memiliki kerusakan kombinasi dengan lebih dari satu faktor penyebab yang memperburuk kondisi dan meningkatkan probabilitas tumbang.

Kata kunci: Panjang serat, diameter serat, ketebalan dinding sel, runkel ratio, serat elastis

I. INTRODUCTION

Bogor Botanical Garden (BBG) is the oldest Botanical Garden in Southeast Asia, until December 2018 it is conserving 12,141 specimens of plant collections from small-sized plants such as herbs to big trees with a total of 3,156 species (excluding orchid species) (Darussalam et al., 2021). BBG, an ex-situ plant conservation, also functions as a tourist place. Since it was established in 1817, several rare plants were included in the red list of the International Union for Conservation of Nature (IUCN) which have lived through the time (Ariati et al., 2019; Rachmadiyanto, et al., 2020).

In the twenty-first century, the survival of plant collections in botanical gardens is considered as priority target. The collections are needed to solve many obstacles, not only because they were a national heritage from the mid-sixteenth to the nineteenth century or essential for study, introduction, acclimatization, and cultivation of high economic value and medicinal plants, focus on conservation and biodiversity, but they are now critical to tackling all odds of global issues outside the boundaries of those past times (Krishnan & Novy, 2016). They are now so important in contributing to solve issues such as climate change, food scarcity, sustainable conservation, restoration ecology, informal education, horticultural therapy, and human well-being (Chen & Sun, 2018; Galbraith et al., 2011; Krishnan & Novy, 2016; Ocak & Kurtaslan, 2017; Powledge, 2011).

By possessing thousands of trees for more than two centuries, BBG has been facing challenges in maintaining its collection, particularly the old tree collections. The older collections of trees have a greater risk of decay, falling, and getting-associated disturbances from pests and diseases. In addition, more frequent or intense extreme events such as heavy rainfall and strong winds from global warming may also increase the falling rate. These abiotic and biotic effects will undoubtedly influence the survival of old collection trees. For instance, a 194-years old tree, Litchi chinensis, fell in 2017. Although it looked healthy on the outside, it had open roots and an undetected-big hollow caused by termites inside its trunks (Affandi et al., 2020; Darussalam et al., 2021). Moreover, intensive human activities in the garden, such as educational and recreational activities, would increase the impact of falling branches or trees (Affandi et al., 2020). Some cases of falling trees that happened in 2005, 2014 and 2015 caused injuries and even deaths. Therefore, preventive actions are needed to preserve these trees even longer by first forestalling which of these ancient historical trees are prone to fall so that incoming damages can be mitigated and minimized.

These cases indeed have become a great concern for the management of BBG as more and more visitors come to BBG. As of 2016, the number of tourists who visited BBG was approximately 1.43 million and was kept growing by around 32% in the next year, with an approximate average of around one million visitors each year during the last five years (Affandi et al., 2020; Darussalam et al., 2021). This increasing trend of visitors, combined with around 1,196 aged trees (more than 100 years old) that have the potential to fall could indeed increase the probability of falling tree accidents which can cause fatalities if no further action is implemented. Moreover, this threat not only causes harm to people but also influences the existence of the rare collections of species especially those that only have one or a few specimens in BBG.

Tree risk assessment and evaluation are the keys to protect existing tree collections from biotic and abiotic disturbances. In urban forests and botanical gardens, tree assessment is an active process and done annually with more frequent performances depending on the surrounding traffic areas, reports and feedback from many stakeholders, and after destructive weather events (Australian National Botanic Gardens, 2016). There are weaknesses and advantages in each type of tree risk assessment methods regardless of what tools are used. They generally use almost similar core principles and methods for assessing tree's health by visual assessment, sounding hammers, and various types of tools that basically could determine decay and hollow level inside tree trunks such as increment boring, cordless drills, resistograph, sonic tomography, arbosonic, or tree Radar Unit. What differs are the ways they score initial ratings of the tree defects, inspection metadata, calculating methods, and final numerical or descriptive hazard rates to make conclusions and recommendations for a tree hazard rating (Koeser et al., 2016). Although they have different assessments, caution and speed of data collection speed must be considered to ensure accuracy and consistency for precise risk ranking.

Preservation efforts to prevent tree collection loss have been performed since 2018 by establishing a health tree monitoring team in BBG. The team is appointed to do regular health checks on every tree collection, especially those that visually seem to be defected and damaged. A few suspected decaying trees inside the trunk are examined with a sonic tomography scanner to get physical wood information. However, these efforts have weaknesses because the checking activity will primarily be performed when the damages are detected visually. As a result, it is too late for prevention as the detrimental effects have already occurred, and in most cases have already been severe.

In addition, the number of personnel and the scanners to do the regular checking are still limited compared to the total number of tree collections in BBG which reaches 5,703 specimens (Ariati et al., 2019). This monitoring effort is indeed far from security aspects for visitors and employees in BBG, considering that approximately 1,196 old trees are categorized as prone to fall. Many previous cases of fallen trees in BBG created a further need for preventive action for a list of aged trees predicted to fall and a vulnerability map zone of fallen trees. This research aims to predict the collection of specimens and families of aged trees (more than 100 years) that are relatively prone to fall with the best-selected model and later to be depicted as a predicted disaster map with several susceptibility levels of zonation based on tree's position and predictive falling probability. The research results are hopefully be used as a preventive action because the impact could also be detrimental to people's lives. Also, this research is expected to dismantle which families of tree collections are classified as susceptible to fall so that biodiversity losses could be avoided.

II. MATERIAL AND METHOD

A. Study Site

The research was conducted in an 87 ha area inside BBG (106°47'40"–106°48'18" East Longitude and 6°35'32"–6°36'13" South Latitude), Research Center of Plant Conservation, Botanical Garden and Forestry, National Research and Innovation Agency (BRIN) (Figure 1). During the study, daily temperatures ranged from 22 to 30°C, with humidity ranging from 60 to 80% in normal day but reaching 95% during rainy day, and with rainfall of 263.48 mm/month. The Botanical Garden is dominated by tropical rainforest species collections, encompassing a generally flat topographic gradient (227-284 m above sea level; asl) with a slightly steep area along the Ciliwung River that divides BBG.

B. Sample Collection

The primary data set for creating and evaluating model predictions was collected from the tree's health monitoring list from January 2018 to June 2021. 154 trees were examined based on reports from field supervisors in each region. They had undertaken routine field monitoring for the tree collections in all parts of the BBG. The reports were then followed by tree health monitoring teams for field checks. The team identified several types of visual signs and symptoms, followed by sounding with hammers on its trunk, and if there were sounds indicating that the trunk is hollow and brittle, that will be checked with PICUS *sonic tomograph* to get several scanning images (Figure 2). This tool measures sound waves inside the trunk by hammering each nail point placed around the trunk. The sound waves travel inside the trunk and detect the wood density. The sound waves gather information and geometric data inside the trunk by calculating the speed of sound between transmitting and receiving points in healthy wood and damaged wood such as fractures, cavities, rot, hollows, and brittleness that can reduce wood's elasticity and density. By using this information, the software will calculate its solid parts and damage percentage. To determine whether the aged trees have probability of falling or not, experiences from the tree health monitoring team and cases of fallen trees over the last three years were used as the basic standard for damage calculations such as fungal bodies on the stems and around

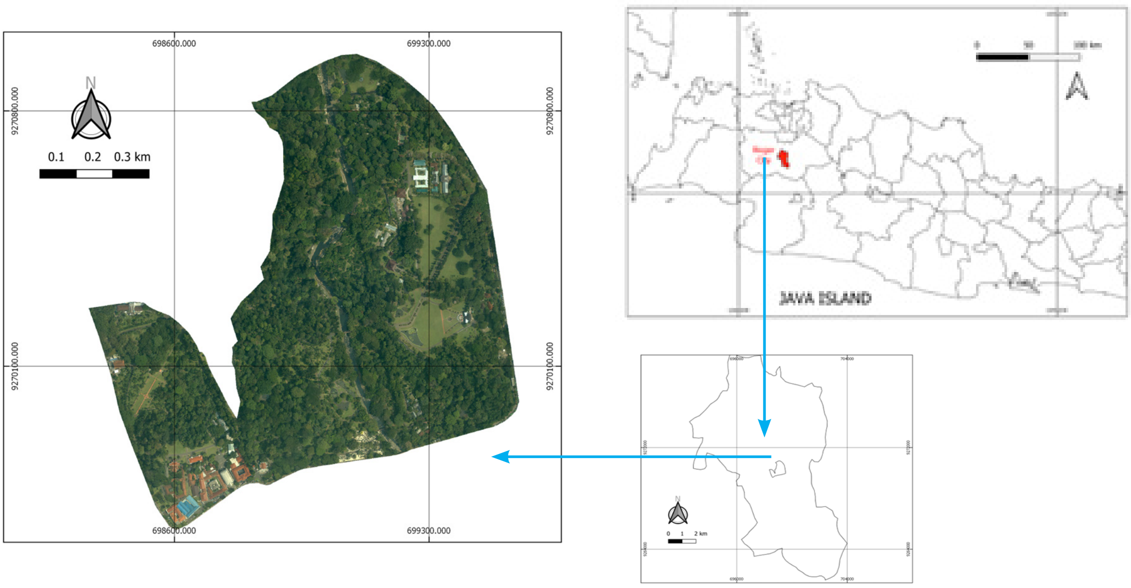

Figure 1. Map of Bogor Botanical Garden

the roots (e.g. *Ganoderma* spp.), hollow, and brittle trunks due to termites, standing dead tree, stranglers, and the position of the trees in relation to the surrounding environment that is likely to be highly exposed to strong winds such as nearby open areas, parks, roads, or the Ciliwung river which were later categorized to have high or low incidence of falling. Trees with high proneness to fall were later treated by reducing the canopy or stem loads, chemical spraying (e.g., insecticides, fungicides) supporting iron or bamboo poles, etc. To predict aged trees (more than 100 years old), we used secondary data of 1,108 aged trees obtained from the Registration Division, BBG. Information attached to each tree specimen such as genus, family, internal geographic position inside BBG (called "vak", and used internally by BRIN), and position of trees to highly exposed strong wind was used as variable predictors. Mainly, to get information about tree positions related to wind, estimation was acquired by a survey for each specimen on the high resolution of BBG's aerial map photography with a 1-meter precision of 15,683 pixels with additional sampling for field survey confirmation at each vak or vague location to get better estimation directly in the field.

C. Analysis

Machine learning is now often used in a variety of sectors including business and sales prediction, spam filtering, health, e-commerce, social media sentiment and hoax, products and customers, mapping classification, banking and finance, mining industry, image classification (computer vision) and many others (Agarwal et al., 2019; Ashari et al., 2020; Balakumar et al., 2018; García-Gonzalo et al., 2016; Singh et al., 2019; Wang et al., 2020; Wibawa et al., 2019). This borderless use makes it possible to be used in predicting the susceptibility of falling trees. Machine learning algorithms are capable of predicting the probabilities by automatically learning the data with statistical formulas to calculate, identify, and analyse patterns in the data, then producing an outcome from experience obtained from previous data. The primary data set used for this study was randomly divided into data training consisted of 80% of the data set and the remaining data (20%), data testing, to make scoring. The data training set was used to create the prediction and build the calculation models, while data testing was used to evaluate each model based on the value of dependent variable (actual occurrence

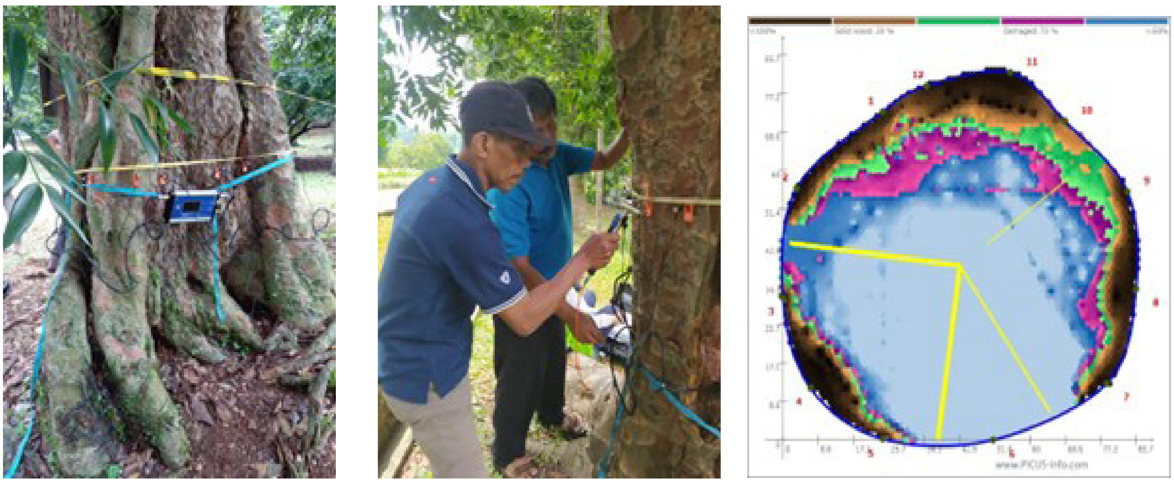

Figure 2. Scanning observation with a PICUS sonic tomograph. Trunk was measured, its geometry, nailed each marked point (left image) and tapped each nailed point with an electronic hammer to create sonic waves (middle image). The tomograph calculated the speed of acoustic waves through its trunk based on the wood's properties to display residual wall thickness and various properties of the wood. Black and brown are solid parts, while blue and purple are the damaged parts. Percent solid and damaged wood are shown above the image (right image)

of the tree falling probability). Prediction models were generated using nine supervised machine learning algorithms, namely Logistic Regression, Gaussian Naïve Bayes, Bernoulli Naïve Bayes, XGBoost, Decision Tree, Random Forest, Stochastic Gradient Descent, and K-Nearest Neighbors, and Support Vector Machine.

C.1. Logistic Regression

Logistic regression is used to make a probability model with a finite number of outcomes typically classified as the first category (class 0) or the opposite category (class 1). Logistic regression requires one or more features as predictors. If there are as many as "n" features (x), then to obtain an output it will require a coefficient (b) of n + 1 as follows (Kurniawan, 2020):

$$\text{Output} = b_0 + b_1x_1 + b_2x_2 + \ldots + b_nx_n \quad \text{...............(1)}$$

The above output is applied to become a probability number with the help of the logistic function below:

$$p(\text{class } 0) = \frac{1}{(1 + e^{-\,output})} \quad \text{........................(2)}$$

This function shows the probabilities of an input number that is classified as class 0 with e is Euler's constant (approximately 2.71828). With this function, positive or negative number, will result in 0 or 1.

C.2. Naïve Bayes

Naïve Bayes is a set of classification algorithms built on Bayes Theory (Reverend Thomas Bayes) for calculating probability and statistical methods. It is a famous model in machine Learning applications due to its simplicity which accommodates all variable attributes to an equal final decision (Wibawa et al., 2019). This model allows explaining "conditional probability" which is a possibility of an A event will occur if B event also occurs since A event depends on B event (Kurniawan, 2020). Mathematically, the Bayes Theorem formula is written as follows:

$$P(A|B) = \frac{P(B|A)\,.\,P(B)}{P(B)} \quad \text{................................(3)}$$

Where,
P = Conditional Probability
B = Data with unknown class
A = Hypothesis X is a specific class
P(A|B) = Probability of the Q hypothesis refers to x
P(A) = Probability of the hypothesis Q (prior probability)
P(B|A) = Probability X in the hypothesis Q
P(B) = Probability X

C.2.1. Gaussian Naïve Bayes

In this study, because we dealt with numerical data classification, Gaussian Naïve Bayes is used where the function is shown as below (Agarwal et al., 2019; Wibawa et al., 2019; Kurniawan, 2020):

$$P(X_i|y) = \frac{1}{\sqrt{2\Pi\sigma^2}}\exp\left(-\frac{(X_i - \mu_y)^2}{2\sigma_y^2}\right) \quad \text{....................(4)}$$

Here, X_i is i-attribute of data instance and y is related class while μ_y and σ_y are mean and variance of predictor distribution.

C.2.2. Bernoulli Naïve Bayes

Also, we used Bernoulli Naïve Bayes because it is possible to predict the occurrence of the fallen trees in terms of presence and absence over the predictors where the function is shown as below (Ashari et al., 2020; Singh et al., 2019; Kurniawan, 2020):

$$P(p|n) \propto P(p) \prod_{1\le k\le nd} P(t_k|\,p)(1 - P(tk'|p)) \quad \text{..(5)}$$

Where n is attribute predictor and p is class variable while $(t_k|p)$ is the conditional probability of predicted term t_k in the attribute predictor of class variable. Meanwhile, $tk'|p$ is the conditional probability of non-predicted term tk' in the attribute predictor.

C.3. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting is a regression and classification algorithm with an ensemble method. It is one of Tree Gradient Boosting variants capable of optimizing ten times faster than any other Gradient Boosting (Chen & Guestrin, 2016; Rahman, 2020). This method is versatile because it is capable of coping with all thinly dispersed data patterns in unified data (Chen & Guestrin, 2016) with an equation:

$$F = argmin_f \; E_{x,y}[L(y, f(x))] \quad \text{..................(6)}$$

The purpose of XGBoost process is to get a function that is closest to F to its construction function $f(x)$ by minimizing lost function value L $(y, f(x))$ and formulated by the above equation. In the training process, to minimize loss function according to initial function $F_0(x)$, algorithm gradient boosting generally has a formula as follows (Rahman, 2020):

$$\{\gamma_m, h_m\} = argmin \sum_{m=1}^{M} L(y_1, f^{(m-1)}(x_i) + \gamma_m h_m(x_i)) \quad (7)$$

Where m is the number of each classification while h and γ are residual output and weight.

C.4. Stochastic Gradient Descent (SGD)

Stochastic Gradient Descent is a modification of gradient descent method techniques in machine learning and deep learning. It calculates the degree of variable change due to changes in other variables. This method performs with specific matter using the entire training data set. The algorithm can be written as follows (Achlioptas & Stanford, 2017; De Sa et al., 2015; Ruder, 2016; Song et al., 2013):

$$w_t + 1 = w_t - \alpha_t \nabla_{wt}\, l(w_t; z_{it}) \quad \text{.....................(8)}$$

Here, w_t is the example parameter training at time t and $l\,(w; z_i)$ is the loss function of w_t on the example label z_{it} during learning rate at time as αt.

C.5. K-Nearest Neighbors (KNN)

K-Nearest Neighbors is used to determine clustering of uncharacterized observations based on grouping data points of the most similar characterized examples from the training data set. It uses Euclidean distance to measure the distance between two points with the following equation (Zhang, 2016):

$$\text{Dist} = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^n} \quad .(9)$$

Where p and q are subjects that are compared with n characteristics.

C.6. Decision Tree

Decision tree is an algorithm used to classify and identify patterns in various fields for constructing a reversed tree-like decision to make a discrete output (Figure 3).

The decision starts from the root node where the data is separated based on the entropy and information gain. Data separation continues until Boolean outcomes at the bottom, leaf nodes, are achieved or there is no data left after the last leaf node. Entropy is used to measure two categories (usually 0 and 1) of the dataset's impurity in each node based on attributes of the data. On the other hand, information gain is calculated to choose the best feature to split into each node. The equation of entropy and information gain are shown below (Charbuty & Abdulazeez, 2021):

$$\text{Entropy (S)} = \sum_{i=1}^{c} P_i \log 2^{P_i} \quad \text{...............................(10)}$$

Where, P_i is the sample number ratio of the subset in the i-th attribute value.

$$\text{Gain (S,A)} = \sum_{v \in V(A)} \frac{|S_v|}{|S|} \ \text{Entropy } (S_v) \quad \text{........(11)}$$

Where, V(A) is the range attribute of A, while S_v is a subset of set S that is equal to the value attribute v.

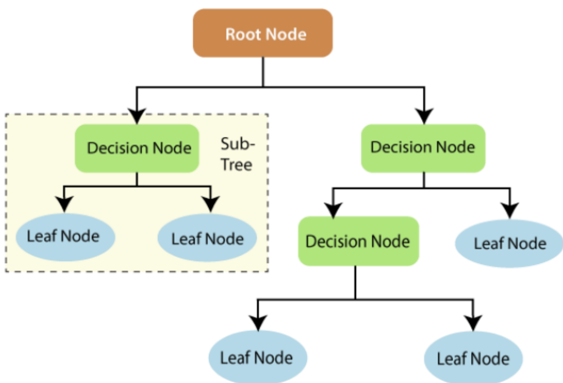

Figure 3. Model of Decision Tree

C.7. Random Forest

Random Forest is an assembled learning technique that builds multiple decision tree models using bootstrapped data sets of original data and randomly selects a subset of variables at each step of the decision tree to make a discrete output. The model then selects all predictions of each decision tree by relying on the majority 'Wins' model to reduce the risks of error from the individual tree. For classification, the equation frequently used is (Cutler et al., 2011; Louppe, 2015):

ensemble prediction $f(x) = \underset{y \in Y}{arcmax} \sum_{j=1}^{J} I(y = h_j(x))$ (12)

Where the possible value of y is denoted by Y while j is the base learner that a tree denoted $h_j(x)$.

C.8. Support Vector Machine (SVM)

Support Vector Machine is a model (Figure 4) that carries an objective to find the best *hyperplane* in n-dimensional space that can distinctly classify data points into two class categories.

SVM will determine the most optimum *hyperplane* located precisely in the middle position of both classes so that it has the farthest distance to the outermost data in the two classes. A *hyperplane* in SVM can be denoted as (Rahman, 2020):

$w.x + b = 0$ ....(13)

where w is the weight of a vector and b is the bias value to get x (the best *hyperplane*) to distinct positive class (1) and negative class (-1). Then, the classification function in SVM can be defined:

$f(x) = sign\ (w.x + b)$ ....(14)

to get the highest margin, it can be found by maximizing the distance value between hyperplanes and the nearest point that is $\frac{1}{||w||}$. This problem is usually called *Quadratic Programming* (QP), by finding minimum point from:

$\min \tau(w) = \frac{1}{2} ||w||^2$ ....(15)

with constraint (y) limitation:

$y_i (x_i . w + b) - 1 \geq 0, i = 1,2,3,\ldots l$ ....(16)

this calculation is solved with *lagrange multiplier* with equation:

$L(w,b,a) = \frac{1}{2}||w||^2 - \sum_{i=l}^{l} \alpha_i y_i (x_i.w + b) - 1, i = 1,2,3,\ldots,l$

Where $ai \geq 0$ is the coefficient value of *lagrange multiplier*. Then optimum value from the above equation can be calculated by minimizing the value of L to w and b while maximizing L to a_i. By considering characteristic on optimum gradient point $L = 0$, it is possible to modify as problem maximation so that it only contains a_i

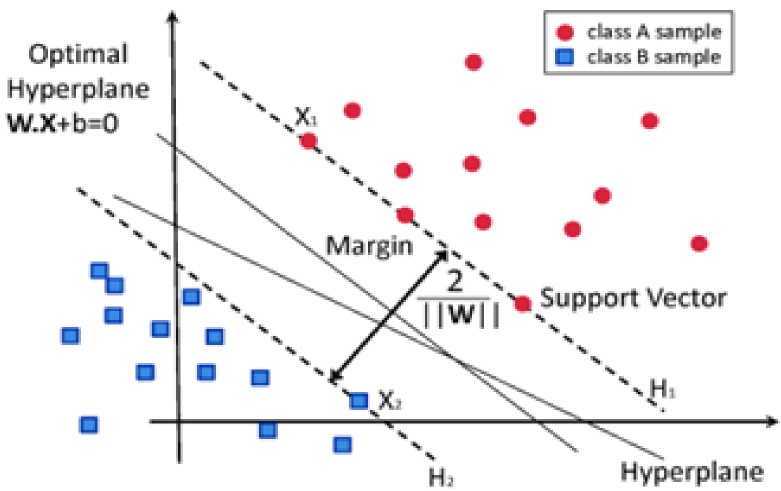

Figure 4. Data classification of SVM (García-Gonzalo et al., 2016)

as below equation:

$$\sum_{i=l}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=l}^{l} \alpha_i \alpha_j y_i y_j x_i x_j \quad \text{..................(18)}$$

Subject to:

$$a_i \geq 0 \; (I = 1,2,\ldots, l) \quad \sum_{i=l}^{l} \alpha_i y_i = 0 \quad \text{................(19)}$$

From this calculation process, $a_i > 0$, a positive value, is obtained as *support vector* and the rest would have $a_i = 0$.

We ran all the prediction models for 154 trees from primary data using Jupyter Notebook as an environmental web tool to run the Python programming language version 3.8.3. We used all the algorithm models from Sckit-learn library. Accuracy of the entire predicting models from data testing was calculated and compared between the predicted data from the data testing (20%) with the actual value occurrence of tree falling from field observation. The accuracy results were later ranked to get the best performance. Accuracy can be defined as below:

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{True\ Positive\ (TP) + True\ Negative\ (TN) + False\ Positive\ (FP) + False\ Negative\ (FN)} \quad (20)$$

Confusion matrix, *Precision* $(\frac{TP}{TP+F})$, and *Area Under Curve* (AUC) on *Receiver Operating Characteristic* (ROC) were also calculated to measure the model performance of the highest accuracy model. We also made classifications among families according to types of damage with *Correspondence Analysis* (CA) and which families of the tree are prone to fall based on primary data using Bray and Curtis index (1957) on Multidimensional Scaling (MDS). We used the Bray-Curtis index because it is better at handling the large proportion of zeroes (e.g., absences of certain damage types) commonly found in this study. This measure does not consider shared absences as being similar, which makes sense biologically. We visualized plots using R software version 4.0.3 (R Core Team 2020) with required package libraries such as factominer, factoextra, magrittr, dplyr, and ggpubr. Tree families that were prone to fall were later categorized into three classes: low, moderate, and high chance of falling. The result of these categories was used as additional information to determine the risk of trees prone to fall as a second layer level prediction to make a zone map from the first model prediction. We used Inverse Distance Weighted (IDW) interpolation method to depict a zone map of trees prone to fall because it could estimate unknown values of the surrounding environment of tree points using known values (tree points) with weighted values (risk level to fall). The IDW formula is given as (Chen & Liu, 2012; Razali & Wandi, 2019; Hamzah & Prayogo, 2014):

$$R_p = \sum_{i=l}^{N} w_i R_i \quad \text{..(21)}$$

Where R_p is the unknown area of probability to fall or not, and R_i is the known data of trees prone to fall or not at the scatter points. While N is the number of scatter points in the set and the weight function assigned to each point is $w_i = \frac{1}{d_{ij}{}^{p}}$, with d is *Euclidian distance* from the unknown data to the known data point, p is power in Shepard's rule which has default value of 2. This method was compatible for measuring the vulnerability zone of tree fall by using a binary value of 1 because the aged trees had a probability of falling and a binary value of 0 as they did not have a probability of falling or was negligible. The zone was later categorized into green areas (a binary value of 0) and other zones used a binary value of 1 which was classified into three different risk levels using MDS which were light green zone (low falling probability), yellow zone (moderate falling probability), and red zone (high falling probability) using QGIS software version 3.10 A Coruna. The result of these categories was used as additional information to determine the risk of trees prone to fall as second layer level prediction to make a zone map from the first model prediction.

III. RESULT AND DISCUSSION

A. Accuracy of all used models to predict tree falling probability within data testing

The result showed that Decision Tree and Random Forest model predictions had the highest accuracy of the nine-machine learning supervised algorithms (Table 1). They both scored 0.774194, which means that the closer the accuracy is to 1, the more accurate the model prediction will be. Generally, models that have more than 0.7 can be classified as having well-performance predictions (Kurniawan, 2020). As the two models had similar accuracy, values of confusion matrixes from both models had to be considered to get the best model performance because accuracy could not distinguish between error values of *false positive* (FP) and *false negative* (FN). These values were essential to get more information by detecting and focusing on a binary classification that was very sensitive in the case of predicting trees where the predicted label was not prone to fall but the true label showed proneness to fall. This model should have the least of FN's values because it gives priority to visitors' and employees' safety from fallen trees as well as the sustainability of aged tree collection as the management of BBG certainly does not want to have any victims or losses in tree collections.

The confusion matrixes showed that FN was lower in the random forest model with a value of 0.05, while in the decision tree it was only 0.1 (Figure 5). Although the decision tree model had a lower value on FP (0.45) than the

Table 1. Accuracy scores of all prediction models

Output [108]:

Prediction Models	Accuracy Score
Decision tree	0.774194
Random forest	0.774194
XGBoost	0.741935
Gaussian Naïve bayes	0.709677
K-Nearest Neighbors	0.645161
Stochastic Gradient Descent	0.645161
Bernoulli Naïve Bayes	0.645161
Logistic Regression	0.645161
Support Vector Machine	0.645161

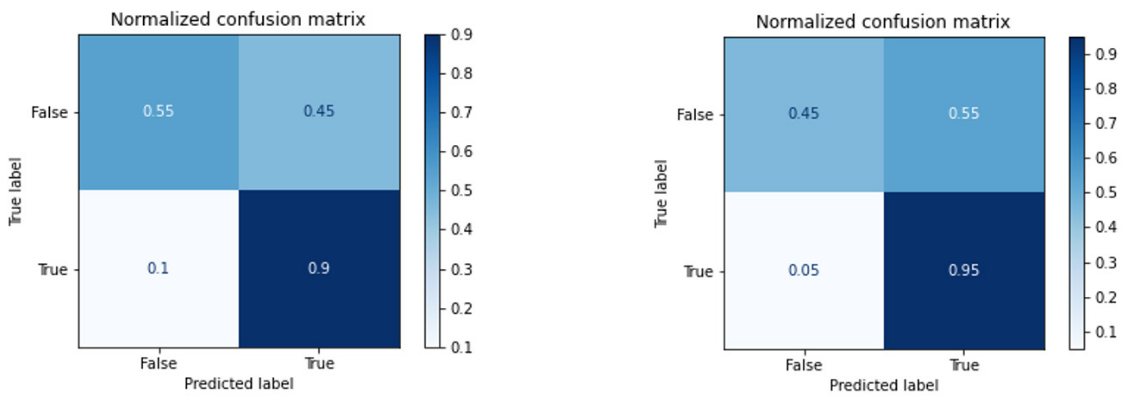

Figure 5. Confusion matrix of the two highest accuracy models: Decision tree (Left) and Random Forest (Right)

random forest (0.55), FN was a more important value because it was related to human safety and sustainability of tree collections than FP which had not that issue if the reality was not prone to fall but the model predicted it would fall. The model must avoid error calculation to the trees that are not prone to fall but turn out to be prone to fall by obtaining as low a value of FN as possible (Kurniawan, 2020). Meanwhile, XGBoost was a good model, but it still had high FN. Similarly, the Gaussian Naïve Bayes model could also become the best model with its zero FN but its FP was higher than the previous three models. In this case, although FP is not so important, higher FP will underrate the model and cause more checking workload to handle wrongly-predicted trees prone to fall. Meanwhile, the five last models had the same accuracy of 0.645161 with zero FN except for SGD which conversely had the highest value of FN. The four models were good at predicting FN but again, they all had the highest value of FP (1), making their prediction worse than Gaussian Naïve Bayes.

Better performance of the random forest model could also be confirmed by its precision and ROC curve. Its precision was 0.76 while the decision tree only had 0.75. Precision indicates how the model can capture positive predictions from TP's value, and it should have a negative correlation with FN. Meanwhile, the ROC curve of both models showed higher value of AUC on the random forest model with a value of 0.7386 than the decision tree model which had 0.7 (Figure 6). Ideally, AUC's value of 1 is a perfect model, but it is practically impossible. Realistically, the more space under the curve, the better the model (Kumar & Indrayan, 2011). The random forest merges with several decisions from tree models (ensemble) and executes them in parallel to get a better final result. This model is sometimes used to solve weaknesses of decision tree model that are often overfitting, unstable (small changes in the data set lead to drastic transformations in the model), and have too many unclear relationships and uncorrelated variables (Kurniawan, 2020). It uses bagging or bootstrapping aggregation tools to get different training data sets in each tree model to perform classification analysis. So, the result is that the whole forest has low variance, although one or several trees have high variance. Each decision tree in the random forest can be called a class prediction that uses a voting mechanism to collect the most popular tree as the final class (Cutler et al., 2011; Louppe, 2015; Wang et al., 2020). From these results from accuracy, confusion matrix, precision, ROC curve space, AUC, and its capability to include all predictions of each decision tree to get the "best" model for reducing error from the individual tree, the compatible model for predicting aged trees that were prone to fall in this study was the random forest model. It

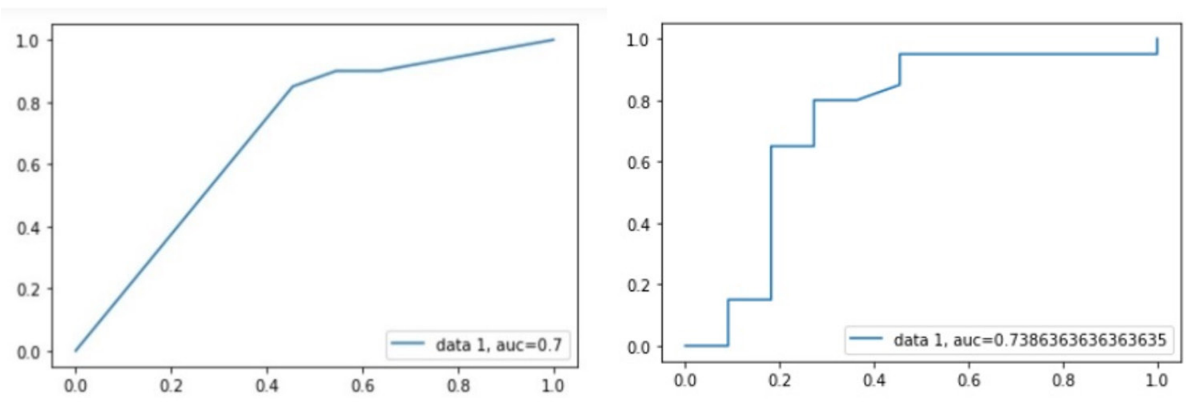

Figure 6. *Receiver Operating Characteristic* (ROC) plot with *Area Under Curve* (AUC) values from both models of Decision tree (Left) and Random Forest (Right)

had a prone probability of falling as many as 855 trees, while 204 trees were safe or had a negligible chance of falling.

B. Susceptible families

The susceptible families were divided into three categories (Figure 7) that are low (green), moderate (blue), and high (red) prone to fall based on MDS result. Among the families of aged tree collections that are prone to fall, Fabaceae was the most prone to fall (orange color). Although Fabaceae was shown as the only individual family creating a stand-alone classification, it would later be included as a high prone probability to fall (red color) in the zoning map. This family tree is known to be susceptible to several pests such as stem borer by *Aristobia horridula* on tree species *Pterocarpus indicus* (FAO, 2007). *Crossotarsus lecontei*, *Hyloconis*, and *Metriona trisignata* are also identified as pests to *P. indicus* (CAB International, 2002). Different types of damage also come from fungal diseases as shown by CA analysis (Figure 8), where Fabaceae and fungal diseases are close to their pinpoints. Angsana wild disease caused by *Fusarium oxysporum* and Sandragon disease caused by an unknown fusarium species were responsible for the decay. Both have similar symptoms that are premature loss of leaves and dieback. However, only Sandragon disease has wilt symptoms, while other fungal diseases such as Cercospora and *Ganoderma lucidum* are also responsible for the causal death of this species (CAB International, 2002; FAO, 2007). Other tree species within the Fabaceae family are also threatened by different pests such as defoliators (*Lamprosema lateritialis*) on *Periscopsis elata* (an endangered tree species) and black wattles (*Lygidolon laevigatum*) on *Acacia mearnsii* (Govender, 2002; Wagner et al., 2008). Moreover, over 66 pests such as defoliators, sap-feeders, wood and shoot borers could harm Fabaceae tree species like Acacia species and Myrtaceae family (red color; highly prone to fall) such as *Eucalyptus* spp. that are susceptible to those pests (Hurley et al., 2017).

While MDS shows falling vulnerability, CA shows types of damage that cause the family to be prone to fall. Families such as Dipterocarpaceae, Sapindaceae, Rubiaceae, Sapotaceae, Myrtaceae, Fabaceae, Meliaceae, Araucariaceae, and Myristicaceae had a higher probability of getting combined damages as their point locations were close to the damages from pests and fungal diseases, stranglers, and heavy wind. Despite having light-hardwood species, Dipterocarpaceae was classified as a highly prone to fall tree family (red color in MDS). In

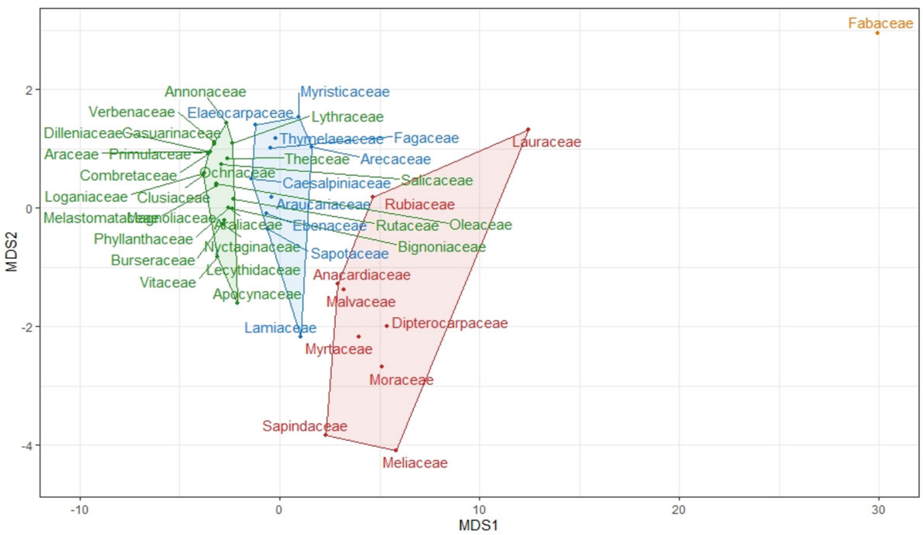

Figure 7. *Multidimensional Scaling* (MDS) of family tree collections. Orange and red-colored families are at a high risk of falling, blue-colored families are at a medium risk, and green-colored families are at low risk.

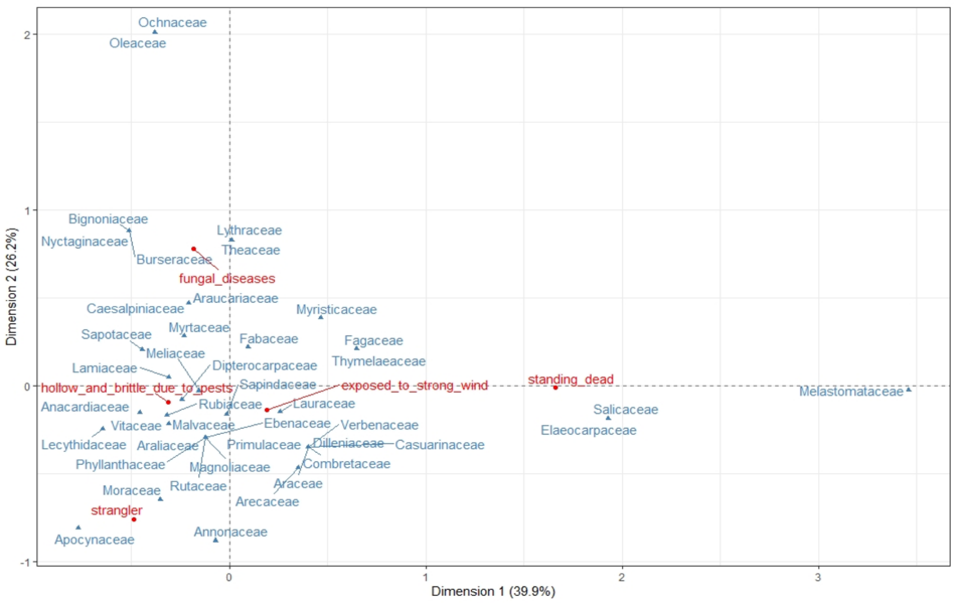

Figure 8. *Correspondence Analysis* (CA) among families according to types of damage. Blue-colored dots are families while red-colored dots are damage types causing the possibility of falling. The closer the point of a family to the point of a type of damage, the greater probability of a tree family that is prone to fall to that damage causal point

BBG, many of the damage cases came from pests that caused hollows and brittleness in the tree's trunk. Although resin in dipterocarps may reduce pest threats, several are still capable of infecting this family, such as *Lymantria Mathura*, which disrupts physiology and photosynthetic activity on *Shorea javanica* in Sumatra, Indonesia (Torquebiau, 1984). Many pests belonging to Coleoptera and Lepidoptera are mainly causal problems in Indonesia for defoliating, wood-boring, and root sucking (Appanah & Turnbull, 1996). Although many of the insects that attack dipterocarp plantations in Indonesia are polyphagous, cases such as borer outbreaks on *Shorea robusta* in India and hairy caterpillars on *Shorea albida* in Sarawak and Brunei could serve as a warning that many of them are already capable of attacking dipterocarp in Indonesia on a large scale (Nair, 2000).

Moraceae was another tree family in the red zone prone to fall. A study by Khan et al. (2021) showed that tree species such as the Jackfruit tree (*Artocarpus heterophyllus*) has many pests and diseases. Several major pests such as shoot and fruit borer (*Diaphania caesalis*) were responsible for around 27.44% of jackfruit trees in Bangladesh including damages caused by trunk borer (*Batocera rufomaculata*). There were also other minor pest damages infested by bud weevil (*Ochyromera artocarpi*), mealybug (*Drosicha mangiferae*), spittlebugs (*Cosmoscarta relata*), bark-eating caterpillar (*Indarbela tetraonis*), caterpillars of leaf webbers (*Perina nuda* and *Diaphania bivitralis*), aphids (*Greenidea artocarpi* and *Toxoptera aurantii*), thrips (*Pseudodendrothrips dwivarna*), and scale insects (*Ceroplastes rubina*) (Ahmed et al., 2013; APAARI, 2012; Khan et al., 2021; Prakash et al., 2009; Soumya et al., 2015; Kallekkattil & Krishnamoorthy, 2017). Some major diseases also infected this species like stem and fruit rots (*Rhizopus artocarpi*), fruit bronzing (*Pantoea stewartia*), bacterial dieback, dieback (*Colletotrichum gloeosporioides*), gummosis (*Phomopsis artocarpi*) while others caused minor damages such as leaf spot (*Phomopsis artocarpina*, *Pestalotia quepini*, *Colletotrichum lagenarium*, *Septoria artocarpi*), gray blight (*Pestalotia elasticola*), anthracnose, and rust (*Rhizopus artocarpi* and *Lasiodiplodia* sp.) (Kallekkattil & Krishnamoorthy, 2017; Khan et al., 2021;

Mohammed et al., 2012; Rahman & Afroz, 2016; TFNet, 2012, Zulperi et al., 2017).

Aside from pests and diseases, low leaf biomechanical properties such as leaf toughness, thickness, and strength could influence their vulnerability to the attack of pests and diseases. These traits are shown in CA analysis where Myrtaceae and Meliaceae are close to pest damages. According to a study conducted by Read et al. (2003), the trend of strength and specific strength on intercostal lamina, secondary vein, and average lamina tend to be lower in species from the Myrtaceae and Meliaceae families, such as *Acmena smithii* and *Toona ciliate* than on species from the Monimiaceae and Eucryphiaceae such as *Doryphora sassafras* and *Eucryphia moorei*. In addition, *T. ciliate* has other traits that show its vulnerability such as higher SLA (Specific Leaf Area) and water content with lower total phenolics as a chemical defense against herbivores. A similar pattern also happens on species of Nothofagus moorei from Fabaceae (blue color, moderate prone to fall tree family) with higher specific strength than *A. smithii* and *T. ciliate* on midrib. Biomechanical properties on leaves are necessary to prevent more damages as herbivores require more force and energy to chew (Read et al., 2003). Genus Cecropia, from Moraceae, also has weak resistance to pests due to high water and N content, as well as low fiber content and strength in its leaves, which makes it vulnerable to herbivores (Coley, 1983).

In Rubiaceae, *Neonauclea calycina* was also a tree species found to be highly prone to fall in BBG. After the Krakatau eruption in 1883, this pioneer species, along with *Ficus pubinervis* (Moraceae), *Timonius compressicauli* (Rubiaceae), and *Macaranga tanarius* (Euphorbiaceae), became dominant in Rakata Besar and Rakata Kecil Islands (Whittaker et al., 1989). Pioneer species tend to be acquisitive type, resulting in being susceptible to the attack of pests and diseases, because they generally have investment strategies (tradeoffs) on obtaining short-term carbon gain and fast RGR (Relative Growth Rate) with long-term leaf persistence and survival (Kitajima & Myers, 2008; Onoda et al., 2017). These strategies also apply to most of the species in Euphorbiaceae family where we did not find the family in MDS but it was predicted to have a low probability of falling by the algorithm. A study by Nakamura et al. (2020) summarized that, within Euphorbiaceae, *macaranga trachyphylla* is fast-growing pioneer species with a lower C/N ratio, higher N, lower lignin, and higher SLA. Geekiyanage et al. (2018) also found that tree species such as *Cleistanthus sumatranus* and *Drypetes perreticulata* from Euphorbiaceae had a lower LMA, lamina thickness, and LDMC (Leaf Dry Matter Content) compared to hilltop species that tend to have opposite functional traits such as *Ligustrum* sp. (Oleaceae) and *Sinsosideroxylon pedunculatum* (Sapotaceae) which are included in MDS as green and blue color respectively. With so many threats from pests and diseases as well as many have acquisitive strategies, they could become more susceptible to being prone to fall.

Different factors such as type and physical soil properties in the surrounding area also influence Dipterocarpaceae's vulnerability to pests and diseases. These affect tree affinity in poor or rich-nutrient conditions and their responses to biotic threats as clearly shown by Palmiotto et al. (2004) from his study in tropical rain forests of Sarawak, Malaysia. *Shorea laxa* and *Swintonia schwenkii* tended to grow on poor-nutrient humult soil, while *Dryobalanops lanceolata* and *Hopea dryobalanoides* generally grew on high-nutrient udult soil. This difference leads to their survival strategies. In the poor-soil condition, plants usually adapt with limited nutrient availability by having higher trait values close to conservative types such as low SLA, low LAR, low growth rates, high LMA (Leaf Mass Area), high C/N ratio, low N (nitrogen) mass and P (Phosphor) mass, low photosynthetic rate, greater life span, thicker cell wall, greater bulk density, high fiber content and high Silica (Geekiyanage et al., 2018; Kitajima & Myers, 2008; Nakamura et al., 2019; Onoda et al., 2011; Onoda et al., 2017). This conservative type's characteristics would help the plants to protect themselves from herbivores (Hanley et al., 2007). Meanwhile, the two species on the rich-nutrient udult soil had significantly higher growth, SLA, and LAR than the two other species on poor-nutrient humult soil indicating they are opportunistic types (fast growth and low survival). These tradeoffs between opportunistic and conservative species, with the presence of biotic threats and availability of resources in the environment, can result in their survival (Kitajima & Myers, 2008; Palmiotto et al., 2004). Threats from climate change such as strong winds and outbreaks of pests and diseases could also lessen the survivability even for conservative types. Moreover, as the trees get older and bigger, more mechanical loads from their wood and canopy would also increase the falling probability.

Other tree collections, Annonaceae and Apocynaceae were included as having a low probability of falling as they seem to have several defense mechanisms. Furthermore, there is no report of pests and diseases on *Alstonia scholaris* (Apocynaceae) many of which were planted in BBG. The damage only came from newly harvested logs by pinhole borers from families Scolytidae and Platypodidae respectively (Nair, 2000). A study by Coley (1983) reported that species from both families, especially those which are persistent characteristic species (shaded condition), have lower water content, higher leaf toughness, more phenolic substances, more lignin, and more cellulose than those of pioneer species which are often spotted in light gaps. As herbivores tend to choose species with more nutritional properties (such as high N and water) and low physical and chemical properties, these gap colonizers would have more grazing damages (Coley, 1983; Onoda et al., 2011; Read et al., 2003). Although the pioneer species have lower defenses resulting in more damages, as long as they can thrive in a rich nutritious environment and not lose too much tissue from herbivores, they can generally recover faster because of their high RGR (Kitajima & Myers, 2008). Some opportunistic types also generally produce a lot of small-sized seeds with rapid growth and high phenotypic plasticity as a strategy to overcome seed predators with a large number of seeds and to adjust to their surrounding environment. As total carbohydrate reserve in tiny seeds is relatively low, their seedlings must have high leaf SLA to get more energy from photosynthesis. That is why the opportunistic type tends to be fast-growing and germinate quickly to avoid early predatory activity at seed or seedling level. On the contrary, the conservative type has few numbers and is large-seeded with well-protected seeds physically or chemically to achieve their survivability (Kitajima & Myers, 2008). These trait tradeoffs would eventually affect plant-animal interactions within an ecosystem within the environment on what strategies are used by the opportunistic or conservative plant species (Hanley et al., 2007; Onoda et al., 2011). These leaf investment strategies on mentioned species of several families that tend to be acquisitive type results in low survivability if they get threats from biotic and abiotic attacks. Leaf construction costs per unit leaf area, nutrient concentrations, and carbon fixation that are part of the leaf economic spectrum would play as tradeoffs between acquisitive and conservative types that later decide their response to disturbances (Kitajima & Myers, 2008; Onoda et al., 2011; Onoda et al., 2017; Zhao et al., 2017)

C. Prediction map of tree zone prone to fall

The prediction map was assembled by considering the results from the random forest algorithm where as many as 204 trees had a binary value of 0 with no prone probability to fall or negligible. In contrast, all binary values of 1 were ranked for their vulnerability using the result from MDS based on family clustering. It showed 334 aged trees with low probabilities, 210 aged trees with moderate probabilities, and 358 aged trees with high probabilities of falling (Figure 9). Trees from family Dipterocarpaceae, Moraceae, and Myrtaceae looked like clumps together and became a dangerous area with a high probability of falling together with

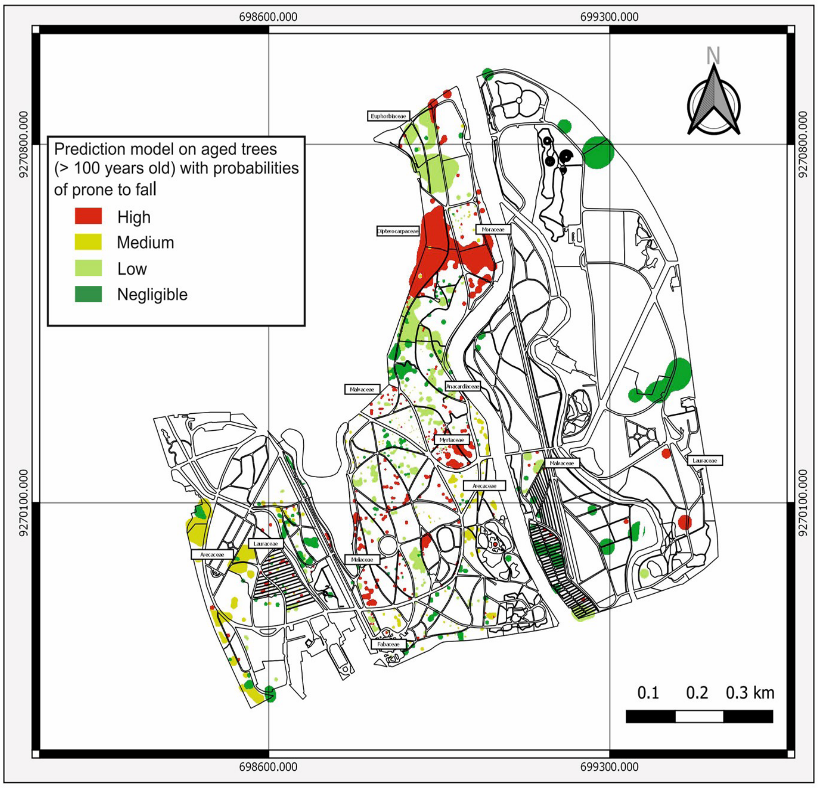

Figure 9. Map zone prediction of aged trees (> 100 years old) with several falling probabilities

Malvaceae, Meliaceae, Fabaceae, and Lauraceae that had a little bit different patterns as they were scattered mainly in the middle area. However, looking at the vulnerability-prone pattern to fall from families Fabaceae and Lauraceae, regardless of their scattered patterns, both families still become potential threats to visitors' safety. Similar to Fabaceae, Lauraceae has potential pests such as *Epimecis hortaria* and *Spodoptera exiqua* that attack *Lindera benzoin*. This family susceptibility is also shown in the study by Kong et al. (2016 & 2015) based on its higher root N concentration on *Cryptocarya chinensis* (Lauraceae), *Acacia auricliformis* (Fabaceae), and *Endospermum chinense* (Euphorbiaceae) that had a negative correlation with root tissue density on thin absorptive roots for resource acquisition and root C fraction as it is constructed with more acid-soluble C compounds. Not only on roots, more enzyme peroxidase on leaves also affects less susceptibility to herbivores such as *Spodoptera littoralis* as an inducible defensive protein (War et al., 2012). These trait correlations give proof of opportunistic-conservation tradeoffs strategy that are already on different plant organs such as leaves, seeds, roots, and stems i.e., the plant economic spectrum (Díaz et al., 2016; Kitajima & Myers, 2008; Onoda et al., 2017; Verbeeck et al., 2019; Zhao et al., 2017).

IV. CONCLUSION

The random forest model algorithm was the best predictor for estimating trees prone to fall since it has the highest accuracy with a larger area of under curve (AOC). It also had a lower FN value which is very important to classify trees prone to fall because it will be dangerous if the model predicts the tree is healthy but, in reality, it is prone to fall. There were 885 of 100-year-old trees predicted to fall. There were 358 aged trees with a high probability of falling dominated by families from Fabaceae, Lauraceae, Moraceae, Meliaceae, Dipterocarpaceae, Sapindaceae, Rubiaceae, Myrtaceae, Araucariaceae, Malvaceae, and

Anacardiaceae. Hollow and brittle trunks due to pests were dominating factors that became causal factors for the family's trees of high probability of falling. A combination between prediction and following up by the use of sonic tomograph to get scanning image inside tree's trunk of the trees with high probability to fall would be important point to get deeper and more promising accuracy about brittleness and fragility, to prevent loss of biodiversity as well as to secure the safety of visitors and workers in the botanical garden.

ACKNOWLEDGEMENT

We would like to personally say thank for the provision of materials and facilities by the Bogor Botanical Garden. The authors also gratefully acknowledge the Tree Health Monitoring team, staff, and technicians who have assisted during field checking.

REFERENCES

Achlioptas, P., & Stanford. (2017). Stochastic Gradient Descent in Theory and Practice. *Stanford*, pp. 1–16.

Affandi, S. D., Halimatussadiah, A., & Asrofani, F. W. (2020). Visitors' preferences on the characteristics of bogor botanical gardens. *Sustainability (Switzerland)*, *12*(22), 1–18. https://doi.org/10.3390/su12229489.

Agarwal, S., Jha, B., Kumar, T., Kumar, M., & Ranjan, P. (2019). Hybrid of Naive Bayes and Gaussian Naive Bayes for Classification: A Map Reduce Approach. International *Journal of Innovative Technology and Exploring Engineering (IJITEE)*, *8*(6), 266–268.

Appanah, S., & Turnbull, J. M. (1996). A review of Dipterocarps. In *Center for International Forestry Research* (Vol. 7). https://doi.org/10.1021/acs.jnatprod.5b00916

Ariati, S. R., Astuti, R. S., Supriyatna, I., Yuswandi, A. Y., Setiawan, A., Saftaningsih, D., & Pribadi, D. O. (2019). *An Alphabetical List of Plant Species Cultivated in The Bogor Botanic Gardens*.

Ashari, H., Arifianto, D., & Faruq, H. A. Al. (2020). *Perbandingan kinerja algoritma multinomial Naïve Bayes (Mnb), multivariate bernoulli dan rocchio algorithm dalam klasifikasi konten berita hoax berbahasa Indonesia Pada media sosial* (Universitas Muhammadiyah Jember). Retrieved from http://repository.unmuhjember.ac.id/id/eprint/5387.

Australian National Botanic Gardens. (2016). *Tree management strategy 2016-2026*. Retieved from https://www.anbg.gov.au/gardens/about/management/policy-docs/ANBG-tree-management-strategy-2016-2026.pdf.

Balakumar, B., Raviraj, P., & Sivaranjani, K. (2018). Prediction of Survivors in Titanic Dataset: A Comparative Study using Machine Learning Algorithms. International Journal of Emerging Research in *Management and Technology*, *6*(6), 1. doi://10.23956/ijermt.v6i6.236.

Charbuty, B., & Abdulazeez, A. (2021). Classification Based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, *2*(01), 20–28.

Chen, F. W., & Liu, C. W. (2012). Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. *Paddy and Water Environment*, *10*(3), 209–222. doi://10.1007/s10333-012-0319-1.

Chen, G., & Sun, W. (2018). The role of botanical gardens in scientific research, conservation, and citizen science. *Plant Diversity*, *40*(4), 181–188. doi://10.1016/j.pld.2018.07.006

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August*, 785–794. doi://10.1145/2939672.2939785.

Coley, P. D. (1983). Herbivory and defensive characteristics of tree species in a lowland tropical forest. *Ecological Monographs*, *53*(2), 209–234. doi://10.2307/1942495.

Cutler, A., Cutler, D. R., & Stevens, J. R. (2011). Chapter: Random Forests. In *Machine Learning* (pp. 1–20). doi://10.1007/978-1-4419-9326-7.

Darussalam, A. D., Sugiyanto, D. R., & Lubis, D. P. (2021). Analisis krisis public relations pada peristiwa tumbangnya pohon di Kebun Raya Bogor. *Profesi Humas Jurnal Ilmiah Ilmu Hubungan Masyarakat*, *5*(2), 251. doi://10.24198/prh.v5i2.29439.

De Sa, C., Olukotun, K., & Ré, C. (2015). Global convergence of stochastic gradient descent for some non-convex matrix problems. *32nd International Conference on Machine Learning, ICML 2015, 3*, 2322–2331.

Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., Dray, S., … Gorné, L. D. (2016). The global spectrum of plant form and function. *Nature*, *529*(7585), 167–171. doi://10.1038/nature16489.

FAO. (2007). Overview of Forest Pests, Thailand. Forestry Department. Retrieved from https://www.fao.org/publications/card/en/c/96d227ec-977d-5f16-858f-878c8e370a36/.

Galbraith, D. A., Iwanycki, N. E., McGoey, B. V., McGregor, J., Pringle, J. S., Rothfels, C. J., & Smith, T. W. (2011). The evolving role of botanical gardens and natural areas: A floristic case study from royal botanical gardens, Canada. *Plant Diversity and Resources*, *33*(1), 123–131. doi://10. 3724/ SP. J. 1143. 2011. 10235.

García-Gonzalo, E., Fernández-Muñiz, Z., Nieto, P. J. G., Sánchez, A. B., & Fernández, M. M. (2016). Hard-rock stability analysis for span design in entry-type excavations with learning classifiers. *Materials*, *9*(7), 1–19. doi://10.3390/ma9070531.

Geekiyanage, N., Goodale, U. M., Cao, K., & Kitajima, K. (2018). Leaf trait variations associated with habitat affinity of tropical karst tree species. *Ecology and Evolution*, *8*(1), 286–295. doi://10.1002/ece3.3611.

Hamzah, R., & Prayogo T. (2014). Interpolation methods for sea surface height mapping. *International Journal of Remote Sensing and Earth Sciences*, *11*(1), 33–40.

Hanley, M. E., Lamont, B. B., Fairbanks, M. M., & Rafferty, C. M. (2007). Plant structural traits and their role in anti-herbivore defence. Perspectives in *Plant Ecology, Evolution and Systematics*, *8*(4), 157–178. doi://10.1016/j.ppees.2007.01.001.

Hurley, B. P., Slippers, B., Sathyapala, S., & Wingfield, M. J. (2017). Challenges to planted forest health in developing economies. *Biological Invasions*, *19*(11), 3273–3285. doi://10.1007/s10530-017-1488-z.

Khan, A. U., Choudhury, M. A. R., Maleque, M. A., Dash, C. K., Talucder, M. S. A., Maukeeb, A. R. M., … Adnan, M. (2021). Management of insect pests and diseases of jackfruit (Artocarpus heterophyllus l.) in agroforestry system: a review. *Acta Entomology and Zoology*, *2*(1), 37–46. doi://10.33545/27080013.2021.v2.i1a.29.

Kitajima, K., & Myers, J. A. (2008). Seedling ecophysiology : strategies toward achievement of positive net carbon balance. In: M. A. Leck, V. T. Parker, and R. L. Simpson. *Seedling Ecology and Evolution, Chapter 8*(Cambridge University Press), 172–188.

Koeser, A. K., Hasing, G., Mclean, D., & Northrop, R. (2016). Tree Risk assessment methods: A comparison of three common evaluation forms. *University of Florida*. Retrieved from https://edis.ifas.ufl.edu/pdf/EP/EP48700.pdf

Kong, D. L., Wang, J. J., Kardol, P., Wu, H. F., Zeng, H., Deng, X. B., & Deng, Y. (2016). Economic strategies of plant absorptive roots vary with root diameter. *Biogeosciences*, *13*(2), 415–424. doi://10.5194/bg-13-415-2016.

Kong, D., Wang, J., Kardol, P., Wu, H., Zeng, H., Deng, X., & Deng, Y. (2015). The root economics spectrum: divergence of absorptive root strategies with root diameter. *Biogeosciences Discussions*, *12*(15), 13041–13067. doi://10.5194/bgd-12-13041-2015.

Krishnan, S., & Novy, A. (2016). The role of botanic gardens in the twenty-first centuryy. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources, 11*(December 2016). doi://10.1079/PAVSNNR201611023.

Kumar, R., & Indrayan, A. (2011). Receiver Operating Characteristic (ROC) Curve for Medical Researchers. *Indian Pediatrics*, *48*, 277–287. doi://10.1007/s13312-011-0055-4.

Louppe, G. (2015). *PhD dissertation: Understanding random forests: From theory to practice*. Retrieved from http://arxiv.org/abs/1407.7502.

Nair, K. S. S. (2000). Insect pests and diseases in Indonesian forest: an assessment of the major threats, research efforts and literature. In *Center for International Forestry Research*. doi://10.17528/cifor/000700.

Nakamura, R., Cornelis, J. T., de Tombeur, F., Yoshinaga, A., Nakagawa, M., & Kitajima, K. (2020). Diversity of silicon release rates among tropical tree species during leaf-litter decomposition. *Geoderma*, *368*(February), 114288. doi://10.1016/j.geoderma.2020.114288.

Nakamura, R., Ishizawa, H., Wagai, R., Suzuki, S., Kitayama, K., & Kitajima, K. (2019). Silicon cycled by tropical forest trees: Effects of species, elevation and parent material on Mount Kinabalu, Malaysia. *Plant and Soil*, *443*(1–2), 155–166. doi://10.1007/s11104-019-04230-7.

Ocak, R. Ö., & Kurtaslan, B. Ö. (2017). Education Function of Botanical Gardens. International *Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, *9*(June 2015), 6.

Onoda, Y., Westoby, M., Adler, P. B., Choong, A. M. F., Clissold, F. J., Cornelissen, J. H. C., … Yamashita, N. (2011). Global patterns of leaf mechanical properties. *Ecology Letters, 14*(3), 301–312. doi://10.1111/j.1461-0248.2010.01582.x.

Onoda, Y., Wright, I. J., Evans, J. R., Hikosaka, K., Kitajima, K., Niinemets, Ü., … Westoby, M. (2017). Physiological and structural tradeoffs underlying the leaf economics spectrum. *New Phytologist*, *214*(4), 1447–1463. doi://10.1111/nph.14496.

Palmiotto, P. A., Davies, S. J., Vogt, K. A., Ashton, M. S., Vogt, D. J., & Ashton, P. S. (2004). Soil-related habitat specialization in dipterocarp rain forest tree species in Borneo. *Journal of Ecology*, *92*(4), 609–623. doi://10.1111/j.0022-0477.2004.00894.x.

Powledge, F. (2011). The evolving role of botanical gardens. BioScience, 61(10), 743–749. https://doi.org/10.1525/bio.2011.61.10.3

Rachmadiyanto, A. N., Wanda, I. F., Rinandio, D. S., & Magandhi, M. (2020). Evaluasi kesuburan tanah pada berbagai tutupan lahan di Kebun Raya Bogor. *Buletin Kebun Raya*, *23*(2), 114–125. doi://10.14203/bkr.v23i2.263.

Rahman, I. F. (2020). *Implementasi metode support vector machine, multilayer perceptron dan xgboost pada data ekspresi gen*. Retrieved from https://dspace.uii.ac.id/handle/123456789/23679.

Razali, M., & Wandi, R. (2019). Inverse distance weight spatial interpolation for topographic surface 3D modelling. *TECHSI - Jurnal Teknik Informatika*, *11*(3), 385. doi://10.29103/techsi.v11i3.1934.

Read, J., Gras, E., Sanson, G. D., Clissold, F., & Brunt, C. (2003). Does chemical defence decline more in developing leaves that become strong and tough at maturity? *Australian Journal of Botany*, *51*(5), 489–496. doi://10.1071/BT03044.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. Retrieved from http://arxiv.org/abs/1609.04747.

Schulman, L., & Lehvävirta, S. (2011). Botanic gardens in the age of climate change. *Biodiversity and Conservation*, *20*(2), 217–220. doi://10.1007/s10531-010-9979-6.

Singh, G., Kumar, B., Gaur, L., & Tyagi, A. (2019). Comparison between multinomial and Bernoulli Naïve Bayes for text classification. 2019 *International Conference on Automation, Computational and Technology Management, ICACTM 2019*, (May 2020), 593–596. doi://10.1109/ICACTM.2019.8776800.

Song, S., Chaudhuri, K., & Sarwate, A. D. (2013). Stochastic gradient descent with differentially private updates. *2013 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2013 - Proceedings*, (December 2013), 245–248. doi://10.1109/GlobalSIP.2013.6736861.

Verbeeck, H., Bauters, M., Jackson, T., Shenkin, A., Disney, M., & Calders, K. (2019). Time for a plant structural economics spectrum. *Frontiers in Forests and Global Change*, *2*(August), 1–5. doi://10.3389/ffgc.2019.00043.

Wagner, M. R., Cobbinah, J. R., & Bosu, P. P. (2008). Forest entomology in West Tropical Africa: Forests insects of Ghana. In *Forest Entomology in West Tropical Africa: Forests Insects of Ghana*. doi://10.1007/978-1-4020-6508-8.

Wang, Y., Zhang, Y., Lu, Y., & Yu, X. (2020). A Comparative Assessment of Credit Risk Model Based on Machine Learning : a case study of bank loan data. *Procedia Computer Science*, *174*, 141–149. doi://10.1016/j.procs.2020.06.069.

War, A. R., Paulraj, M. G., Ahmad, T., Buhroo, A. A., Hussain, B., Ignacimuthu, S., & Sharma, H. C. (2012). Mechanisms of plant defense against insect herbivores. *Plant Signaling and Behavior*, *7*(10), 1306–1320.

Whittaker, R. J., Bush, M. B., & Richards, K. (1989). Plant recolonization and vegetation succession on the Krakatau Islands, Indonesia. *Ecological Monographs*, *59*(2), 59–123. doi://10.2307/2937282.

Wibawa, A. P., Kurniawan, A. C., Murti, D. M. P., Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., & Nugraha, Y. R. (2019). Naïve Bayes Classifier for Journal Quartile Classification. *International Journal of Recent Contributions from Engineering, Science & IT (IJES)*, *7*(2), 91. doi://10.3991/ijes.v7i2.10659.

Zhang, Z. (2016). Introduction to machine learning: K-nearest neighbors. *Annals of Translational Medicine*, *4*(11). doi://10.21037/atm.2016.03.37.

Zhao, Y. T., Ali, A., & Yan, E. R. (2017). The plant economics spectrum is structured by leaf habits and growth forms across subtropical species. *Tree Physiology*, *37*(2), 173–185. doi://10.1093/treephys/tpw098.