GeoPlanning
Journal of Geomatics and Planning
Geoplanning: Journal of Geomatics and Planning, Vol. 12, No. 1, 2025, 95 - 122

e-ISSN: 2355-6544

Received: 09 March 2024;
Revised: 12 May 2025;
Accepted: 18 May 2025;
Available Online: 18 May 2025;
Published: 26 May 2025.

Keywords:
Tourism Potential Zone (TPZ), K-Nearest Neighbors Model, Analytic Hierarchy Process

*Corresponding author(s)
email: raha@gmail.com

Original Research Open access

Tourism Potential Zone Mapping for The State of Madhya Pradesh, India Using MCDM and Machine Learning Models

Shrinwantu Raha[1]*, Sayan Deb[1]

1. *Department of Geography, Bhairab Ganguly College, Belgharia, Kolkata, India*

DOI: 10.14710/geoplanning.12.1.95-122

Abstract

The rich and diverse tourism attractions of Madhya Pradesh have long been recognized, but the Tourism Potential Zones (TPZs) have yet to be clearly identified. This research aimed to uncover these hidden potentials using a combination of Multi-Criteria Decision Making (MCDM) and machine learning techniques. TPZ was predicted using a approaches, including Analytic Hierarchy Process (AHP), Linear Model (LM), Elastic Net Model (EN), and K-Nearest Neighbors (KNN). Further, by combining the above models, a new ensemble model (AHP-LN-EN-KNN ensemble) was prepared. We followed the ROC-AUC (Area Under Curve) and Root Mean Squared Error (RMSE) as evaluation measures. The findings reveal a landscape of promise, with each model with accuracy levels ranging from 81.4% to 90.6%. The AUC values for the models ranged from approximately 70% to 95%, while the RMSE values ranged from 0.8 to 1.3. The ensemble model appeared with better accuracy (for training set 0.92 and for test set 0.88), higher AUC value (for training set 94.5% and for test set 89.4%) and the lowest RMSE (i.e., 0.71) value. On the other hand, the AHP was identified with higher combined RMSE (i.e., combined RMSE 1.08) and diminished AUC (i.e., for training set 70.1% and test set 70.2%). The northern, south-western, and middle regions emerge as high-potential areas, whilst the south-western edges languish with less promise. Meanwhile, the north-western expanse offers a scene of moderate potential. These findings not only inform, inspire, laying a foundation for Madhya Pradesh's long-term tourist growth.

1. Introduction

Tourism, which is hard to define, may promote a region's environmental, social, and economic growth (Telfer & Sharpley, 2015). However, its' enigmatic character requires practicality, precision, and efficacy to fulfill its' revolutionary potential (Harianto et al., 2020). Tourism's dynamic and unpredictable tapestry is hidden between exquisite delicacy and cultural sanctity (Smith, 2015). (Atun et al, 2019) describe tourism potential as a complex web of social, cultural, economic, and infrastructural factors. This initiative attracts tourists with its' appealing melody of accessibility and lofty quest of guardianship over multiple valuable resources, producing a tapestry of attraction that captivates the adventurous soul (Saner et al., 2019). Underutilization persists across sectors and paradigms, awaiting its' transformational potential (Ramírez-Guerrero et al., 2021). Tourism potentiality goes beyond asset accumulation to turn sites into wanderlust hangouts (Raha & Gayen, 2022a). It nurtures a tourist-environment relationship by curating and protecting a destination's soul (Sarker, 2018). The evaluation of tourist potential embodies complex calculation that defines a full mosaic of criteria within the authority of the United Nations World Tourism Organisation (UNWTO) (Trukhachev, 2015). Within this

DOI: 10.14710/geoplanning.12.1.95-122

framework, geology, relief, aspect, and water proximity influence landscape appeal all interact to shape the landscape's attraction (Yamin et al., 2021). The complex accessibility network, depicted by road and railway density measures vs distance metrics, controls accessibility dynamics. Demographic variables including sex ratio, literacy, population, and growth rates suggest societal vitality, whereas tourism density enhances the region's appeal (Stępniak et al., 2019). A complex tapestry of these factors defines visitor potentiality, symbolizing the UNWTO's holistic approach to TPZ evaluation. In fact, exploring the potential of tourism becomes an appeal for countries to develop in unison, strengthening the fabric of culture and way of life. It is an enlightening journey, a pilgrimage to protect the integrity of our nature and our shared past.

The Indian Planning Commission considers tourism the nation's second-largest business since it creates low- and intermediate-skill jobs (Rajan, 2018). India's business lacks infrastructure and coordination despite its' historical and natural charms (Mamun & Mitra, 2012). In Madhya Pradesh, TPZs are poorly defined (Aijaz, 2022). Due to efficiency and simplicity, GIS and multi-criteria decision-making approaches, notably the Analytic Hierarchy Process (AHP), are increasingly employed for this purpose (Raha & Gayen, 2022a, 2022b). Statistical and machine learning models are also emerging alongside MCDM models. The Linear model (LM) is a statistical regression approach that performs well with GIS. The supervised learning algorithms along with the LM model are widely used in the landslide prediction (Dong et al., 2011), ground water potentiality assessment (Naghibi et al., 2015), eco-tourism potential assessment (Mirsanjari & Mirsanjari, 2012; Zhang et al., 2024), forest fire site detection (Tien Bui et al., 2016), and other fields. The Elastic Net (EN) model reduces overfitting and multicollinearity. EN model has been also widely used in the hazard mitigation (Suchting et al., 2019), landslide prediction (Zhu et al., 2023), flood prediction (Al-Areeq et al., 2023), non-linear tourist behaviour prediction (Brida, 2018), picture categorization (Soomro et al., 2016), and more.

The KNN algorithm swiftly categorizes fresh data (Okfalisa et al., 2017). K-Nearest Neighbour (KNN) has three advantages: 1) it is fast to compute, 2) it predicts better than other models, and 3) its' output is simple to read (Okfalisa et al., 2017). After storing all data, the KNN model classifies fresh data points using distance functions. The KNN model is widely used for the analysis of credit rankings, (Chen et al., 2011) pattern recognition (Mir & Nasiri, 2018), data mining (Mohanapriya & Lekha, 2018), intrusion detection (Aburomman & Ibne Reaz, 2016), face recognition (Nugrahaeni & Mutijarsa, 2016; Sugiharti et al., 2020), and for the analysis of health related data (Mittal et al., 2019). Varied models have varied accuracy levels, hence integrated ensemble models are needed to explain real-world events.

Out of these methods, the AHP have been applied widely for the TPZ identification (e.g., Raha et al., 2024, Sahani, 2019). However, the EN, KNN and LM methods have been applied rarely for the prediction of TPZs. Moreover, their performance has not also been compared in literature. A new ensemble model was also prepared in this research by taking average of AHP, LM, EN, and KNN models. Further, till now, the tourism potentiality of the state of Madhya Pradesh is unexplored. Therefore, the objectives of the present research are as follows: To evaluate and compare the performances of one MCDM technique (i.e., AHP technique), 3 machine learning models (i.e., LM, KNN and EN model) in the prediction of TPZ, and to prepare one ensemble model to improve the performances of the input algorithms (i.e., AHP, LM, EN and KNN models) for the prediction of TPZ for the state of Madhya Pradesh. The TPZs were explored through an integrated 9- step process in this research. The Receiver Operating Characteristic (ROC) curve and Root Mean Squared Error (RMSE) have been adopted to evaluate and compare the machine learning and MCDM models. This research beautifully utilized the GIS platform to spatially illustrate the input data as well as the TPZ.

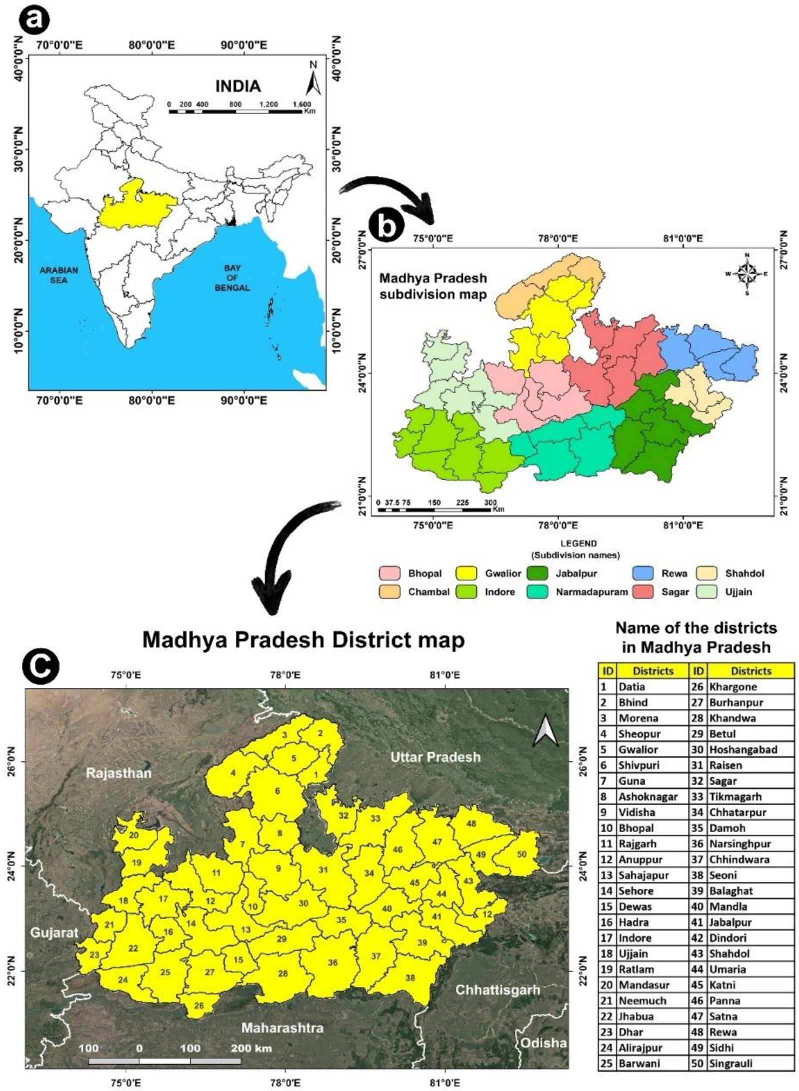
2. Data and Methods

2.1. Study Area

A rich cultural and historical history exists in Madhya Pradesh. Tourism studies favor Madhya Pradesh, the "Heart of India," in the midst of the Indian subcontinent for its historical, cultural, and natural splendor. Bhopal, Indore, Gwalior, Ujjain, and Jabalpur are major cities in Madhya Pradesh. North, southeast, south, west,

DOI: 10.14710/geoplanning.12.1.95-122

and north-west boundaries of Madhya Pradesh are Uttar Pradesh, Chhattisgarh, Maharashtra, Gujrat, and Rajasthan Figure 1.



Source: Authors, 2023

Figure 1. Location Map

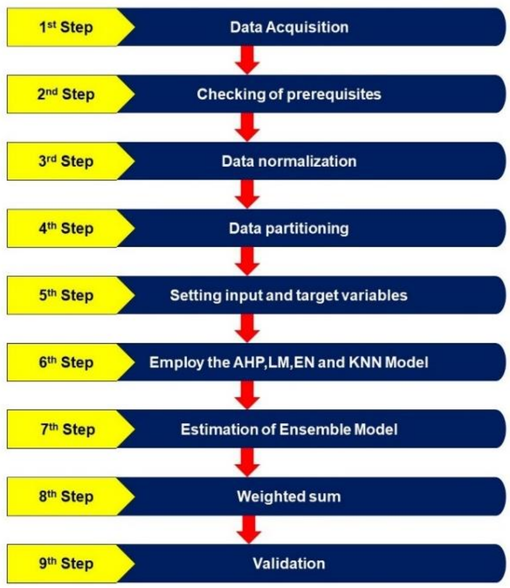
Madhya Pradesh has UNESCO World Heritage Sites including Sanchi Stupa and Khajuraho. For historians, archaeologists, and cultural anthropologists, the area is an engaging case study because of these archaeological wonders, which offer insightful information on the historical importance and architectural history of the region. Madhya Pradesh is a dynamic laboratory for comprehending the dynamic changes in infrastructure, accommodations, transportation, and tourism legislation . The state comprises a unique scenario for its' ecological and wildlife tourism through its' wealth of natural beauty and biodiversity. Kanha National Park and Bandhavgarh National Park are just two of the many national parks and animal sanctuaries that Madhya Pradesh is home to. The existence of tigers and leopards, two iconic megafaunas, has drawn a lot of attention from the fields of wildlife conservation and tourist management. The region is also very rich with the tribal traditions and indigenous populations. Many different tribal communities (e.g., Bhil, Sahariya, Kharwar, Munda,Kol, Sora, Baiga, Kolba, Andh etc.), with unique customs (e.g. Ghotul), traditions, and artistic expressions (i.e., Mandana art etc), are found in Madhya Pradesh.

DOI: 10.14710/geoplanning.12.1.95-122

There are many studies on tourism in Madhya Pradesh, including resource development (Pandey et al., 2014), social media's impact on tourism (Gohil, 2015), art and craft tourism (Kumar et al., 2023), and tourism's economic effects (Sharma, 2019), role of mass tourism (Chandravanshi & Jain, 2023; Gohil, 2015) development of sustainable tourism sector (Kishnani, 2022), eco-tourism (Ahmad & Pandey, 2016) for the state of Madhya Pradesh, but hardly any research work is available on tourism potential zone identification on this tract. Therefore, the tourism potential zone identification for the state of Madhya Pradesh is a noble attempt.

2.2. Methodology

The methodological framework was marked in the Figure 2. The 9-step methodology was used in this research to demarcate the TPZ of the Madhya Pradesh state.



Source: Authors, 2023

Figure 2. The Methodology

2.2.1. Data Acquisition

The first step was to acquire all the data to be used in this research. All of the secondary data sources were illustrated in the Table 1. Here, the Geology (GL), Relief (RL), Aspect (AS), Distance from River (DR), Road and Railway Density (RRD), Distance from Road and Railway (DRRD), Sex Ratio (SR), Literacy Rate (LR), Total Population (TP) and Growth Rate (GR) were used. The Geological data was downloaded from the USGS. This is a compiled data, which includes petroleum geology, geological provinces, oil and gas fields of the South-Asia. The data was available in a shapefile (.shp format) format, a part of the U.S. Geological Survey World Energy Project. For the efficient management and analysis of enormous amount of data; the World was classified into 8 energy regions and those are further subdivided into geologic provinces, on the basis of the natural geologic entities. These may often include a dominant structural element or a number of contiguous elements.

UNESCO World geologic maps and other tectonic geologic maps helped to delineate the boundaries of major geologic provinces. Those shapefiles were amalgamated by the USGS from the UNESCO. The geologic maps of South and East Asia of 1976 and 1990 having scales of 1:10,000,000 and 1:5,000,000 were used here. The relief and Aspect maps were prepared from the Digital Elevation Model (DEM), synthesized by the Shuttle Radar Topography Mission (SRTM) having the spatial resolution of 30m. SRTM DEM was downloaded from USGS Earth Explorer. River, road and railway shapefiles (prepared from the maps of India online portal) were required for the preparation of DRRD and DR layer. Euclidian distance method Equation.1 was utilized to estimate the DRRD and DR spatial layer:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \ldots\ldots\ldots\ldots\ldots \text{(Eq.1)}$$

where, d - Euclidian distance, (x_2, x_1) - is the point exist on the river or road, (y_2, y_1) - is the closest point of the previous.

The Euclidian Distance tool was used to estimate the distance from each cell in the raster to the closest source. The SR, LR, TP and GR were collected from the Census data prepared by Directorate of Census Operations Madhya Pradesh, Ministry of Home Affairs, Govt. of India (2011). The tourist spots were marked with the help of Google Earth and non-participant observation technique (Banik & Mukhopadhyay, 2022) (Figure 21). All of the secondary data sources used here contain an open-access licence agreement; which specify that any data can be used for the academic purposes by citing their original sources (Ruda, 2016).

Table 1. Data Sources for Analyse the Tourism Potentiality

Sl no.	**Criteria**	**Sources of data**	**Nature of data**	**Relationship with tourism**	**References**
1	Geology (GL)	https://pubs.er.usgs.gov/publication/ofr97470C	Open access (Input/Independent variable)		Rutherford et al. (2015)
2	Aspect (AS)	https://earthexplorer.usgs.gov/ (2014-09-23) Resolution (30 meters * 30 meters)	Open access (Input/Independent variable)	Inverse	Woźniak et al. (2018)
3	Relief (RL)	https://earthexplorer.usgs.gov/ (2014-09-23) Resolution (30 meters * 30 meters)	Open access (Input/Independent variable)	Inverse	Sahabi Abed and Matzarakis (2018)
4	Distance from River (DR)	https://www.researchgate.net/post/How-to-get-a-River-basin-shape-files-of-India-kindly-suggest-any/615150d9584a141e805d5a83/citation/download	Open access (Input/Independent variable)	Inverse	Woźniak et al. (2018)
5	Road & Railway Density (Road and Railway Shapefile) (RRD)	https://grpbhopal.mppolice.gov.in/railway-map & https://www.mapsofindia.com/maps/madhyapradesh/madhyapradeshroads.htm	Open access (Input/Independent variable)	Proportional	Wang et al. (2018)
6	Distance from Road & Railway (DRRD)	https://grpbhopal.mppolice.gov.in/railway-map & https://www.mapsofindia.com/maps/madhyapradesh/madhyapradeshroads.htm Road and Railway Shapefile	Open access (Input/Independent variable)	Inverse	Raha et al. (2022b)
7	Sex Ratio (SR)	Directorate of Census Operations Madhya Pradesh Ministry of Home Affairs, Govt. of India (2011)	Open access (Input/Independent variable)	Proportional	
8	Literacy Rate (2011) (%) (LR)			Proportional	Natalia et al. (2019)
9	Total Population (2011) (TP)			Proportional	Trukhachev (2015)
10	Growth Rate (2011) (%) (GR)			Proportional	Banerjee (2014)
11	Density of Tourist Spots (TS)	Non-Participant observation technique, Google Earth, Govt. Reports and relevant websites	Dependent Variable	Proportional	Raha et al. (2022a)

2.2.2. Checking of Prerequisites

a. Checking of Multicollinearity

The correlation matrix was prepared to check whether any multicollinearity exist within the acquired data. For the efficient processing of machine learning and MCDM models, the basic prerequisite is the independent nature of the variable (Vairetti et al., 2024). If the correlation coefficient value comes >0.8 then the variables are considered as the potentially correlated and dependent. If it is less than 0.8 the variables are considered as the independent and then the models could be applicable on the collected data sets.

b. Checking of Residuals vs. Fitted Plot, Normal Q-Q Plot, and Residuals vs. Leverage Plot

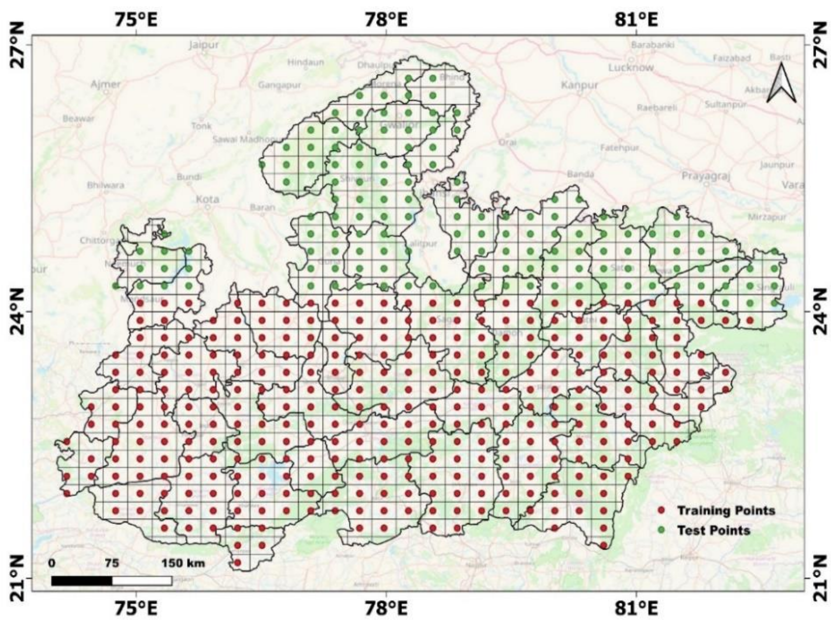
Before applying the LM model, the residuals vs. fitted plot, normal Q-Q plots, scale location plot and residuals vs. leverage plot were checked. Residuals vs. fitted plot help to detect the non-linear nature, variances (equal/unequal) of error, potential biases and possible outliers (Chavan & Momin, 2017). Normal Q-Q plot help display the theoretical distribution of associated data sets (i.e., whether normal, exponential or gaussian distribution etc.) determining if two data sets originate from populations with a similar (run-of-the-mill) distribution is another benefit of it. There may be cases, where the results are affected due to the extreme data points, which are influential. The Cook's distance was applied in the residuals vs. leverage graph to calculate the measure the reasonable range beyond which the data points might be influential (Zhao et al., 2020).

2.2.3. Data Normalization

The normalization of data is required to remove the redundancy of data, minimize the modification errors, and simplify the query process (Ayesha et al., 2020) of the data. The normalization of the data was done in the third step of the research.

2.2.4. Data Partitioning

In the fourth step, total 471 data points were extracted from the each of the raster layers. For the efficient analysis, those were subdivided into the training and test set. 70% of the total data points (i.e., 330 data points) were included in the training set and the 30% of the total data points were included in the test set (i.e., 141 data points) Figure 3.



Source: Authors, 2023

Figure 3. Location of Training and Test points

2.2.5. Setting the Input and Target Variables

The tourist spots were added in the GIS platform and using the Inverse Distance Weightage (IDW) tool, the density of tourist spots (TS) layer was created. This raster layer was considered as the target (dependent) variable which was assumed to be influenced by the other independent variables in this research. Therefore, apart from the TS raster variable, other variables were considered as input variables.

DOI: 10.14710/geoplanning.12.1.95-122

2.2.6. Employ the Analytic Hierarchy Process (AHP), Linear Model (LM), Elastic Net (EN) model and K-Nearest Neighbors (KNN) Model

a. The AHP Model

The AHP is a widely recognized multicriteria decision making tool (Saaty, 1980), which was used in this research to put the weightage of each thematic layer. AHP is an objective mathematical procedure that allows the incorporation of subjective and objective choices. AHP's broad application is a result of its ease of use, readiness, and high degree of flexibility (Sahani, 2019). In decision making, the AHP approach incorporates a hierarchical structure of different criteria in a pairwise comparison method (Saaty, 1980). AHP matrices display the uniform number of rows and columns (Raha & Gayen, 2022a). Each criterion is rated against the other criteria by assigning a relative priority scale of 1 (Equal importance), 3 (Moderate importance), 5 (very strong) and 9 (Extreme importance) to construct a pair-wise comparison matrix. 2,4,6 and 8 were the intermediate values. Relative priority scale was assigned to each of the spatial layer and their classes using the recommendation of 5 expert panel. Each expert had more than 5 years of experience in the field of travel and tourism.

The experts signed a separate informed permission letter confirming that their replies would be utilised solely for academic reasons without revealing their identities. For example, in the Table 2, SR and GR are less important than GL and RL. Therefore, SR and GR were coded as 2 and the GL and RL were coded as 9 and 8 respectively. Similarly, RRD is less important than GL. Therefore, the RRD was coded as 3 and the GL was coded as 9. Here, GL was identified with the highest priority (23.8% weightage) followed by the RL, AS (18.4% weightage), DR (11% weightage), DRRD (8.2% weightage), RRD (5.7% weightage), TP (4% weightage), LR (3% weightage), SR (2.5% weightage), and GR (1.8% weightage). All subclasses of each thematic layer was rated based on the causative factors on which the tourism phenomena triggers. Here, higher rating indicates higher tourism potential value. The weightage was estimated by the AHP using the following formula Equation. 2

$$0 < w < 1;\ \sum_{i=1}^{n} w_{ij} = 1 \ldots\ldots\ldots\ldots\ldots \textit{(Equation 2)}$$

where, Consistency Ratios (CRs) were calculated to determine, whether pairwise comparisons were consistent or inconsistent. It was calculated as follows Equation. 3 :

$$CR. = \frac{C.I.}{R.I} \ldots\ldots\ldots\ldots\ldots \textit{(Equation 3)}$$

$$\text{where, } C.I. = \frac{\lambda_{max} - n}{n-1} \ldots\ldots\ldots\ldots\ldots (\textit{Equation } 4)$$

where, *C.I.* is the Consistency Index; *R.I.* is the Random Index. If the CR is <0.1 the index was considered as the consistent (Saaty, 1980). In this research, consistency ratios of all matrices Table 4 were <0.1 so, all of the matrices were appeared as consistent.

b. Linear Model (LM)

The linear model (LM) is expressed here as follows Equation. 5:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \cdots \ldots\ldots\ldots\ldots + a_nx_n \ldots\ldots\ldots\ldots\ldots (\textit{Equation.5})$$

where, y is the target variable; which is the density of tourist spot raster. $x_1, x_2, x_3 \ldots\ldots x_n$ are the variables in the linear model specified in the Table 1. $a_1, a_2, a_3 \ldots\ldots a_n$ are the coefficients. a_0 is the parameter of the model. The parameters a and b are estimated using the ordinary least square (OLS) procedure. The OLS method is implemented by minimizing the actual and predicted value.

c. Elastic Net Model (EN)

The EN model is an enhanced iteration of the machine learning-based regression model that incorporates both lasso and ridge regression. Equation. 6, Equation. 7, Equation. 8: Modifying the Equation. 5, we can write

$$SSE_{Ridge} = \sum_{i=1}^{n}(x_i - \bar{x}_i)^2 + \lambda(a_1^2 + a_2^2 + a_3^2 + \cdots \ldots \ldots + a_n^2) \ldots\ldots\ldots\ldots\ldots \textit{(Equation 6)}$$

$$SSE_{lasso} = \sum_{i=1}^{n}(x_i - \bar{x}_i)^2 + \lambda(\lceil a_1 \rceil + \lceil a_2 \rceil + \lceil a_3 \rceil + \cdots \ldots \ldots + \lceil a_n \rceil) \ldots\ldots\ldots\ldots\ldots \textit{(Equation 7)}$$

$$SSE_{EN} = \sum_{i=1}^{n}(x_i - \bar{x}_i)^2 + \lambda[(1-\alpha)\sum_{i=1}^{n} a^2 + \alpha|a|] \ldots\ldots\ldots\ldots\ldots \textit{(Equation 8)}$$

where, Equation. 6, Equation 7, and Equation 8 are known as the ridge, lasso and Elastic Net regression model (EN) respectively. x_i is the observed model variable; $\bar{x}_i$ is the predicted model; *SSE* is the sum of the squared error. λ is the regularization parameter, that controls the amount of regularization applied. By adding the regularization term, the magnitude of the regression coefficients *(a)* are penalized by the ridge, lasso and EN models. If the α value tends to 0 the Equation 8 is transformed into Equation 6 ; and when the α value tends to 1 then Equation 8 is transformed to Equation 7.

For the effective use of the EN model; the fraction deviance plot and log lambda plot were used. The fraction deviance plot illustrates that how the model coefficients Varies with the increase or decrease of the fraction deviance. It helps to determine the larger or smaller coefficients with the changing nature of fraction deviance value. The log lambda plot is essential to know the coefficient scores as a function of log (λ). The top numbering of the plot indicates the number of predictors (variables) the model.

d. K-Nearest Neighbors Algorithm (KNN) Model

The K-nearest neighbors algorithm, sometimes referred to as KNN or k-NN, is a supervised learning classifier that employs proximity to produce classifications or predictions about the grouping of a single data point (Boateng et al., 2020). Although it may be applied to classification or regression issues, it is commonly employed as a classification method since it relies on the idea that comparable points can be discovered close to one another (Zhang et al., 2017). It generalizes well to multi-class problems and can learn complex decision boundaries when combined with ample data. Additionally, as it does no training step beyond caching the dataset, it's very effective in situations where training speed is essential and memory resources are ample. Here, the Euclidian distance metric was used to determine the distance between the given point and query point. The partitioning of various category datasets is one of the decision boundaries that this KNN model aids in determining (Bansal et al., 2019).

2.2.7. Ensemble Model (AHP-LM-EN-KNN Model)

The ensemble models were prepared by taking average of AHP, LM, EN, and KNN models (Huang et al., 2024). These models were prepared using following equations Equation 9:

$$Ensemble\_model = \frac{AHP\ model + LM\ model + EN\ model + KNN\ model}{4} \ldots\ldots\ldots\ldots\ldots \textit{(Equation 9)}$$

Before proceeding with the ensemble model preparation, we evaluated the inter-model correlation coefficient. If the correlation coefficient exceeds 0.8, suggesting a high level of correlation between the models, ensemble approaches may not be appropriate (Cankurt & Subasi, 2022).

2.2.8. Weighted Sum

The Variable Importance Plot (VIP) was used to determine the weightage of each criterion utilized in AHP, LM, EN and KNN models. VIP is a popular global method to rate the importance of criteria involved in a model by rating the criterion 0 to 100 based on the priorities in the model. 0 means the lowest priority and 100 indicates the highest priority. Those weightages are used to determine the weighted sum model; which is marked as the TPZ Equation 10. (Mitra et al., 2022):

$$TPZ = \sum_{i=1}^{n} w_i^{TPZ} \ldots\ldots\ldots\ldots\ldots \textit{(Equation 10)}$$

where, TPZ is the Tourism Potential Zone; w_i^{TPZ} is the criteria for the TPZ identification.

2.2.9. Validation of Models

Validation is one of the most critical steps in ensuring the accuracy of any model (Mitra et al., 2022, Sofaer et al., 2019). There are several ways to validate a model and here, the overall accuracy, ROC-AUC and RMSE metric were used here. The ROC-AUC curve renders the trade-off between the False Positive Rate and True Positive Rate. The false positive rate is shown on the x axis of the ROC (e.g. a two-dimensional graph), while the true positive rate is shown on the y axis. Equation 11 and Equation 12 represents the attributes of x and y axis respectively:

$$x = false\ positive\ rate = 1 - \frac{TN}{TN+FP} \text{ .... (Eq.11)}$$

$$y = true\ positive\ rate = \frac{TN}{TN+FP} \text{ .... (Eq.12)}$$

$$accuracy = \frac{(TN+TP)}{(TN+TP+FN+FP)} \text{ .... (Eq.13)}$$

where, *TN*, *FP*, *TP*, *FN*, and *FP* represent true negative, false positive; true positive; false negative; and the false positive values. The AUC (i.e., the area under the ROC curve) was applied to evaluate the performance of the AHP, LM, Elastic Net and KNN models. The proposed tourism potential map was verified high (Code 1) and moderate to low tourism potential points (Code 0). Further, the tourism potential map was validated using Root Mean Squared Error (RMSE). Lower RMSE depicts higher accuracy of the model. In this research, the RMSE was estimated using the following formula Equation 14:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \text{ .... (Eq.13)}$$

Where, x_i is the observed value; y_i is the predicted value. The number of observations is denoted as *n*. The ARC-GIS 10.4 version, R and Python were used to process the data in this research.

2. Result and Discussion

3.1 Analysis of Criterion (Input Variable)

a. Geology (GL)

Overall, the Madhya Pradesh has nine geological units, which are Carboniferous sedimentary rocks, water, Cretaceous sedimentary rocks, quaternary sediments, Tertiary and Cretaceous sedimentary rocks, Paleocene Cretaceous extrusive rocks, Tertiary igneous rocks, lower Triassic to upper Carboniferous sedimentary rocks, and undivided Precambrian rocks Figure 4a. As the scenic beauty of water bodies attract tourists; it creates smooth visual ambience on tourists (36.7% weightage, Code 9). On the other hand, the undivided Precambrian rocks hold least or equal importance (9.6% weightage, Code 1). The Tertiary and Cretaceous sedimentary rocks are noticed with moderate importance (14.4% weightage, Code 3).

b. Aspect (AS)

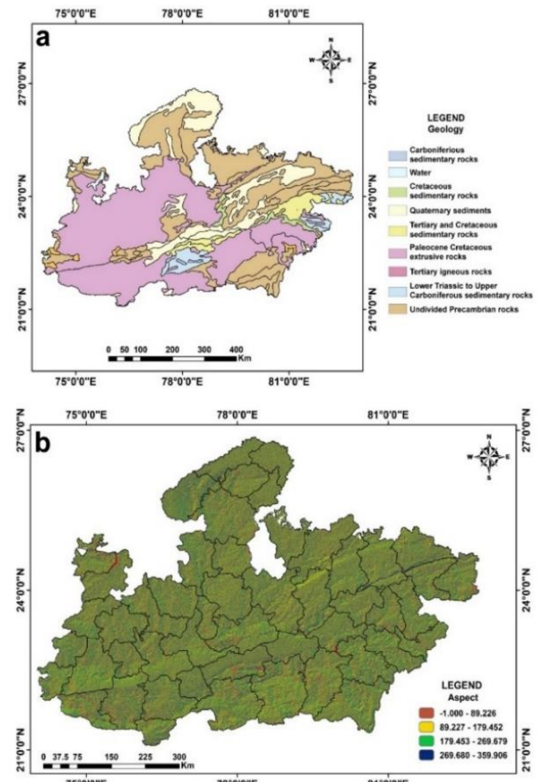
The Aspect was reclassified here into four classes; and as the class value decreases, the tourism potentiality is expected to increase and vice-versa. Following this, the lowest class was gained very strong to extreme importance (Code 8, 55% weightage). On the contrary, the highest class achieved the Equal to Moderate importance (Code 2, 9.4% weightage) Figure 4b. Other classes were marked with Strong to Very Strong importance and Moderate to Strong importance.

c. Relief (RL)

The relief of the Madhya Pradesh fluctuated from 56 metre to 1333 metre. The northern, north-eastern and south-eastern sections of the study area were marked with lower relief (i.e., 56.00 -300 metre). The Narmada River is flowing from south-western to north-eastern section and hence this portion attains a moderate to low

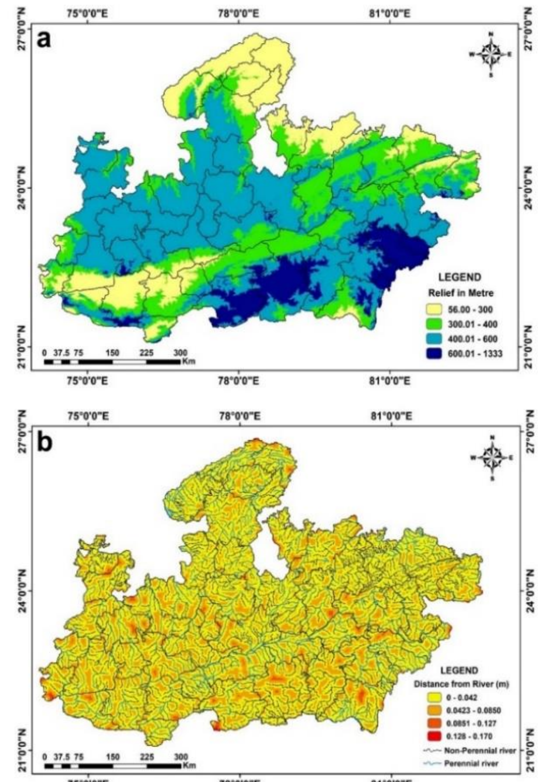
DOI: 10.14710/geoplanning.12.1.95-122

relief (i.e., 300 metre - 400 metre). The north-western and south-eastern portions were marked with a higher relief value (400.01metre- 1333 metre) Figure 5a. Low relief is suitable for tourism activity (Ovreiu et al., 2018) and thus the lowest relief was noticed with higher priority and vice-versa.



Source: Authors, 2023

Figure 4 a) Geological Map; b) Aspect Map



Source: Authors, 2023

Figure 5. a) Relief b) Distance from River

d. Distance from River (m) (DR)

The DR in the Madhya Pradesh varied from 0 to 0.170-meter Figure 5b. Riverine beauty increases the scenic beauty of a particular territory (Pouya & Başkaya, 2018). Therefore, tourism potentiality is positively enhanced by the scenic beauty of the river. Here, DR was classified into 4 classes; and as the class value increases; priority decreases and vice-versa. Hence, the lowest class (i.e., 0 to 0.042m) was marked with the highest weightage (48.7% weightage, Code 8) and the highest class (i.e., 0.128 to 0.170m) was marked with the lowest priority (5.5% weightage, Code 2). The lowest class was identified here with the highest areal coverage (%).

e. Road and Railway Density (RRD)

The RRD fluctuated in the Madhya Pradesh from 0 to 453 Km./Sq. Km Figure 6a. Excessively high RRD decreases tourism potentiality but the moderate to low RRD increases the value of tourism potentiality (Acharya et al., 2022). Here, the RRD has been subdivided into 4 classes and the highest class was identified with the highest priority and vice-versa.

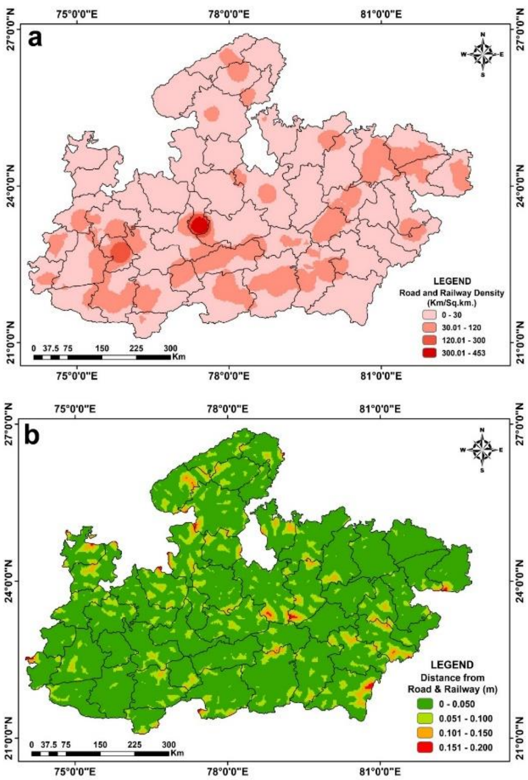
f. Distance from Road and Railway (DRRD)

The DRRD varied from 0 to 0.200 meter within the study area Figure 6b. If the distance from road and railway increases; tourism potentiality decreases and vice-versa (Dedík et al., 2022). Here, the DRRD was divided into 4 classes and the highest class was marked with the lowest weightage and the lowest class was marked with the highest weightage. The highest class was coded with 2 (Equal to Moderate importance) and lowest class was coded with 8 (Very Strong to Extreme Importance).

DOI: 10.14710/geoplanning.12.1.95-122

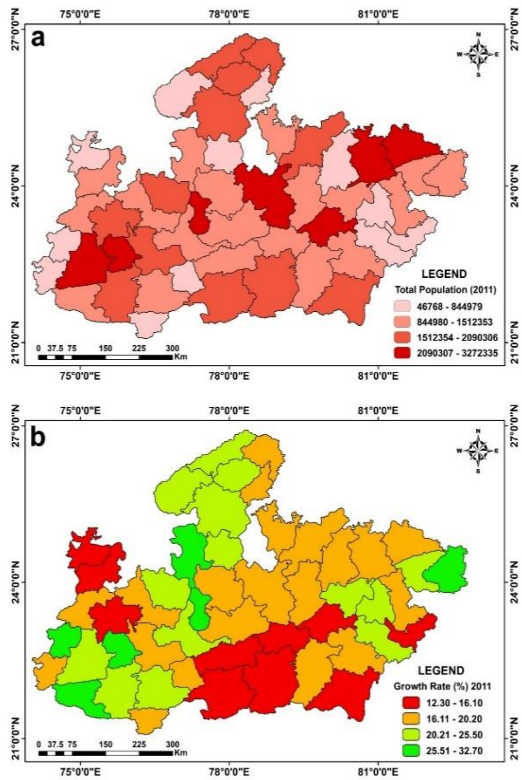
g. Total Population (TP)

The TP varied from 46768 to 3272335 within the study area (Figure 10a). The TP is comparatively low in the Datia, Sheora, Ashoknagar, Panna, Shahdol, Dindori, Anupur, Hadra, Burhanpur, Jhabua, Alirajpur and Neemuch districts. Higher TP (i.e., 2090307 to 3272335) was marked in the Satna, Rewa, Sagar, Jabalpur, Bhopal, Indore and Dhar districts Figure 7a. Comparatively low TP accelerates the scenic beauty and thus motivates the tourists to visit the Madhya Pradesh. Here, the TP was classified into 4 class, and as the class value decreases, priority increases and vice-versa.



Source: Authors, 2023

Figure 6. a) Road and Railway Density b) Distance from Road and Railway



Source: Authors, 2023

Figure 7. a) Total Population b) Growth Rate

h. Population Growth Rate (PGR)

Low PGR value (i.e.,12.30% to 16.10%) was marked for the Chhindwara, Hoshangabad, Betul, Warshinghpur, Jabalpur, Balaghat, Anuppur, Ujjain, Mandasur, and Neemuch districts. The PGR is higher (i.e., 20.21% to 32.70%) for the Guna, Bhopal, Singrauli, Indore, Barwani, Jhabua, Dhar, Khargone, Khandwa, Sehore, Rajgarh, Shivpuri, Ashoknagar, Gwalior, Sheopur, Morena, Katni, Umaria, Dindori and Sidhi districts Figure 7b. Remaining districts are marked with 16.11% to 20.20% PGR. Higher PGR degrades the environment and thus deteriorates the tourism potentiality (Raha & Gayen, 2022b). Here, the PGR was reclassified into 4 groups and as the different class value of PGR decreases; tourism potentiality increases and vice-versa.

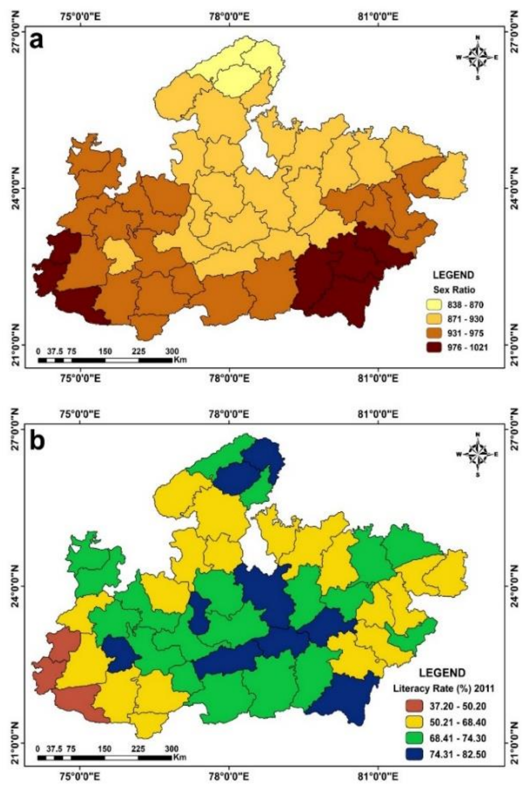
i. Sex Ratio (SR)

The Sex Ratio (SR) of the Madhya Pradesh varies from 838 to 1021. Lower SR (i.e., 838 to 870) occurred for the Morena, Bhind and Gwalior districts. Higher sex ratio (i.e., 976 to 1021) was identified for the Jhabua, Alirajpur, Barwani,Seoni,Belaghat, Mandia and Dindori districts. Remaining districts were noticed with 871 to 975 sex ratios Figure 8a. The higher sex ratio helps to flourish the tourism potentiality (Rahman, 2021) and therefore as the class value of SR increases; priority increases and vice-versa.

DOI: 10.14710/geoplanning.12.1.95-122

j. Literacy Rate (LR)

The Literacy Rate of the study area fluctuates from 37.20% to 82.50%. Comparatively low literacy rate (i.e., 37.20% to 68.40%) was marked for the Jhabua, Alirajpur, Barwani, Sheopur, Shivpuri, Guna,Ashoknagar, Rajgarh, Ratlam, Dhar, Khargone, Burhanpur, Khandwa, Tikmagarh, Chhatarpur, Panna, Mandla, Dindori, Shahdol, Umaria, Sidhi, and Singrauli districts. Higher LR was observed for the Datia, Morena, Vidisha, Anuppur,Sahajapur,Sehore,Dewas, Ujjain, Neemuch, Raisen, Betul, Damoh, Chhindwara, Seoni, Katni, Satna, Rewa, Bhind, Gwalior, Bhopal, Indore, Hoshangabad, Sagar, Narsinghpur, Balaghat, and Jabalpur districts Figure 8b. As the tourism potentiality is positively vibrated by the LR (Raha & Gayen, 2022a); the higher class of LR was marked with the higher priority and vice-versa.



Source: Authors, 2023

Figure 8. a) Sex Ratio b) Literacy Rate

3.2 Density of Tourist Spots (TS) (Target Variable)

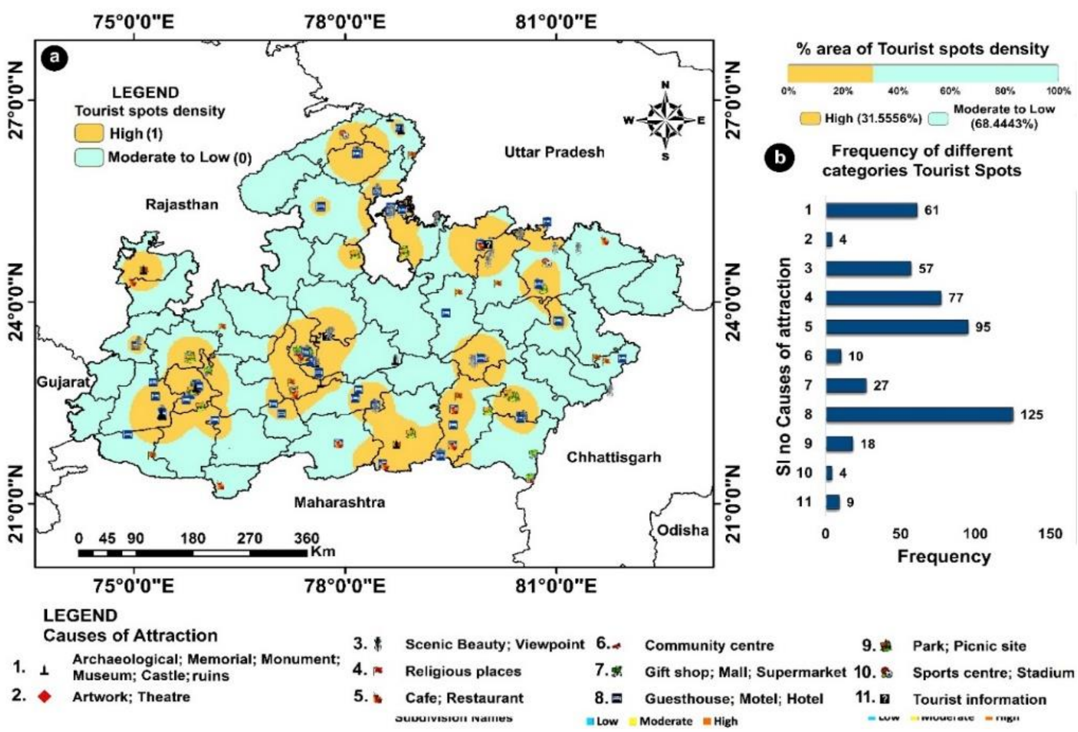
Density of tourist spots is a good indicator of tourism potential (Chen et al., 2021; Marrocu & Paci, 2013). The intensification of tourist spots increases the tourist availability and in turn creates destination loyalty. Many tourist spots create ample opportunity for tourists to identify many destinations within a single time budget friendly time frame. Total 487 tourist spots were identified in the Madhya Pradesh state; and those were categorized under 11 categories. Bhimbetka rock shelter is one of the most popular archaeological tourist sites explored here.

Apart from it, the Bhind fort (castle), Maharana Pratap Square (memorial), Matageswar Temple (monument), Rani Roopmati Mahal (monument), hanuman Statue (monument), Tribal Museum, Darya Khan's Tomb (Ruins) were explored. The region is very rich in several artwork (e.g, Raja Bhoj and Bhagawan Kala Kendra etc.), theatre centres (e.g. Davy auditorium, Bharat Bhawan amphitheatres), several attractive scenic beauty places (e.g., Purva falls, Udaigiri cavesRaneh waterfalls viewpoint etc.), water tower, Budhist (e.g. Southern gate, Western gate etc.), Christian (e.g. Nun Monestry), Hindu (e.g. Nilkanth Shiva temple, Iscon

DOI: 10.14710/geoplanning.12.1.95-122

temple), Muslim (e.g. Jama Masjid) and Sikh (e.g. Gurudwara) tombs. Café (e.g. Dominos Vijay Nagar, Cafe Coffee Day Cafe Kava, Chaifeteria etc.), restaurants (e.g. Mediterraneo Restaurant, My Kitchen Restaurant, Food Land Restaurant, Balaji Family Restaurant and Dhaba etc.), community centres (e.g., Akshat Garden, Relax Garden, Nani Maa Ki Dharamshala, Ravindra Bhavan, Hindi Bhavan, Gandhi Bhavan, Shahpura Community Hall, Sindhu Bhavan etc.), giftshop (e.g. handmade items), guesthouse (e.g. Sharma Guest House, Madai Forest Guesthouse), hotel (e.g., Gem Palace, Royal Garden, Red Maple, Pramad Palace etc.), motel, park (e.g., Kanha National Park, Rani Park, Saket Park, Dravid Nagar Colony Park, Naveen Nagar Park, Laxman Sing Gaur Udyan etc.) picnic site (e.g., Ekta Park, Siddha Ghat etc.), sports centre, stadium (e.g., Railway Stadium, Cricket Stadium, Ashbagh Stadium, Dr Bhim Rao Amedkar Stadioum etc.) and several tourist information centres (e.g., Ticket Counter Man-Singh-Palace, Orchah Nature Reserve, Ticket Counter, Asi Counter, Government of India Tourist Office, Ticket Office, M.P. Tourism Office, Kanha National Park- Kisli Gate etc.),Supermarket (e.g. Vishal Mega Mart, Aparti Super market, Aprooti Super Market) were identified and listed with latitude and longitude (with the help of a GPS).

Overall, the region is dominated by Guesthouses and hotels, followed by café, religious places, archaeological sites, giftshops and mall, park, community centres, tourist information centres, and several tourist viewpoints Figure 9b. The density of tourist spots is comparatively high at the Northern, South-western and middle South-north stretches. On the other hand, the South-western portions were marked with the relatively low density. Overall, 31.56% area of the Madhya Pradesh was marked as the highly dense with the popular tourist spots; and 68.44% area was identified as the low to moderately dense with several popular tourist spots Figure 9a.



Source: Authors, 2023

Figure 9. a) Density of tourist spots b) Frequency of different categories of tourist spots

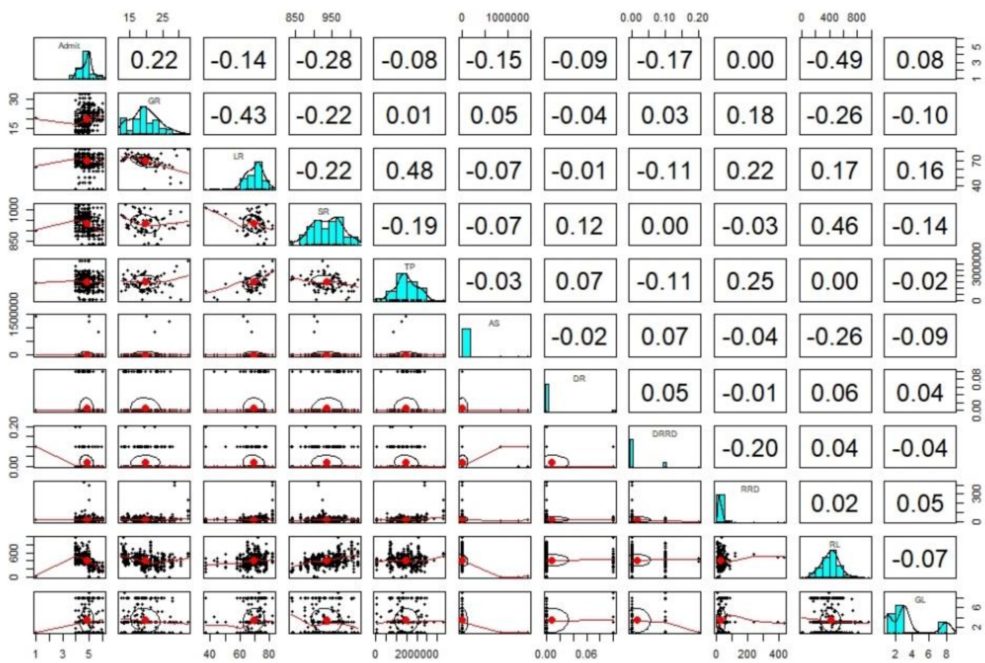
3.3 Analysis of Prerequisites

a. Assessment of Multicollinearity

The correlation matrix was portrayed in the Figure 4. It is evident that here the correlation coefficient value fluctuated from -0.49 to +0.17. As all of the correlation coefficient value comes below 0.8; it can be stated that no significant correlation coefficient value exists in the dataset. It was further verified from the scatter plot in Figure 10. In most of the cases, the points are coalesced in different portions of the plot section. Therefore, all variables used in this research is independent and machine learning models can be applied on collected data without any hesitation. Further, before proceeding to prepare the ensemble model, the multi model inter-collinearity was checked in Table 3. Here, the correlation coefficient fluctuated from 0.55 to 0.74. Therefore, models are not highly inter-correlated with each other and ensemble method could be applicable in determination of TPZs.

Table 3. Correlation Matrix between Each Model

Models	AHP	LM	EN	KNN
AHP	1			
LM	0.6	1		
EN	0.59	0.69	1	
KNN	0.55	0.73	0.74	1

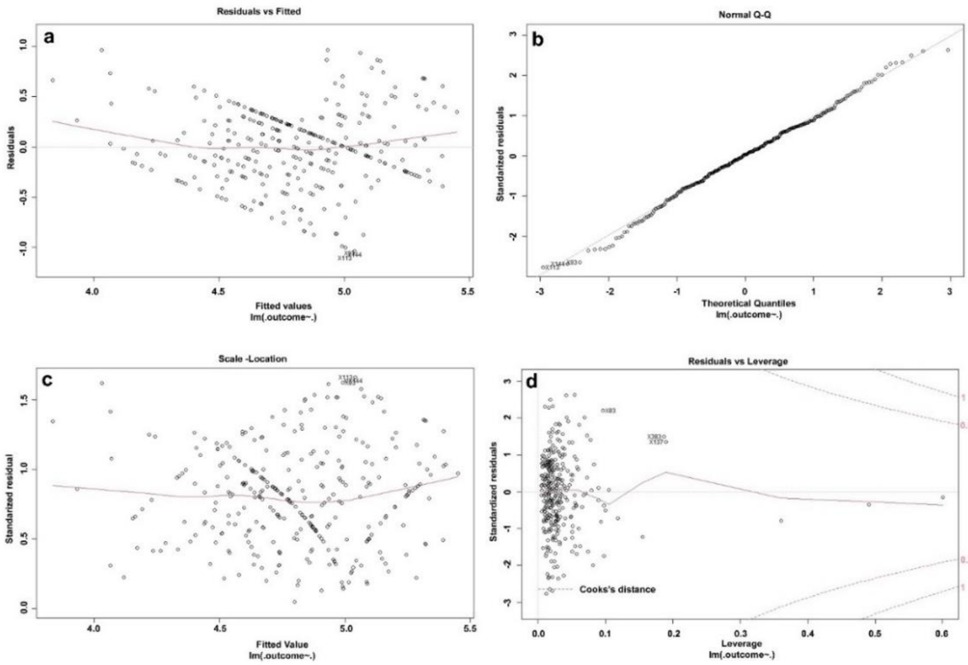


Source: Authors, 2023

Figure 10. Correlation Matrix

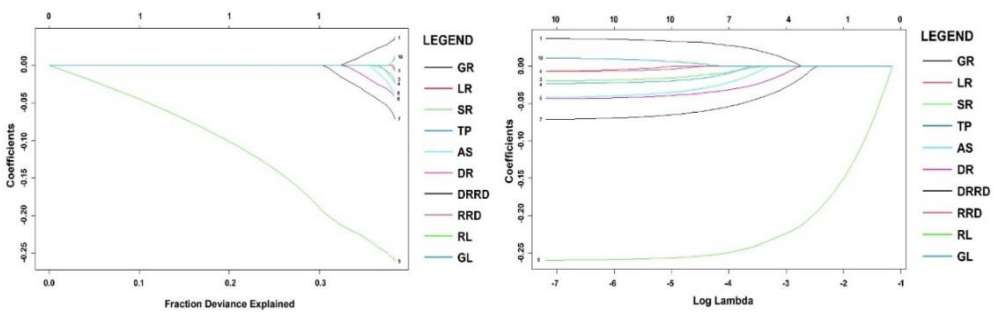
b. Residuals vs. Fitted plot, Q-Q Plot, Scale location plot, Residuals vs. Leverage plot, Fraction Deviance Plot and Log Lambda vs. Coefficients Plot

The red line seems to be fitted with the dashed line (parallel to the x-axis) in the residuals vs. fitted plot Figure 11a. In the Q-Q plot. Figure 11b the standardized residuals fit with the theoretical quantiles. Here all of the data points are exactly aligned over 45°line in the Q-Q plot. In the scale location plot Figure 11c, the residuals spread wider along with the x-axis. Moreover, the red line is almost aligned with the dashed line. That means the spread is random. There is no influential case found in the residual vs. leverage plot Figure 11d. Here, all data points are well inside the Cook's Distance line. The coefficients are large in both side of the axis, whenever the fraction deviance value is increasing Figure 12a. The coefficients are fluctuating from 0 to -0.25. The coefficients are increasing with increasing the log (λ) Figure 12b.



Source: Authors, 2023

Figure 11. Different Plots a) Residual vs. Fitted; b) Normal Q-Q; c) Scale Location d) Residual vs. Leverage



Source: Authors, 2023

Figure 12. Different Plots a) Fraction Deviance; b) Log Lambda

c. Best Tuning Results

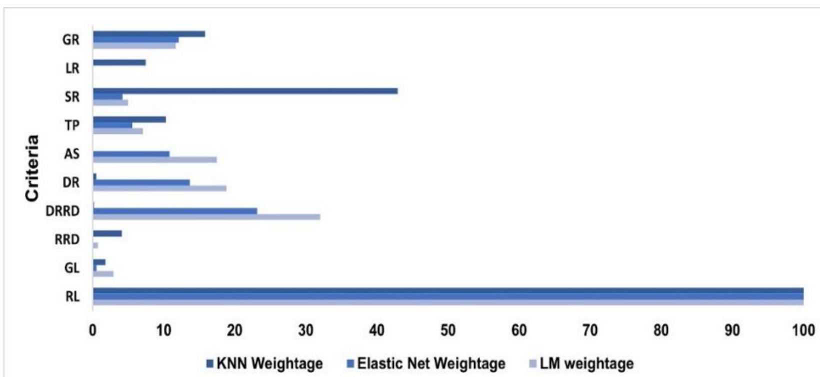
After the 5 repetition and 10-fold cross validation, here, optimal λ and α values were found. In this research, after the best tune the λ and α value was obtained as 0.0136 and 0.861 respectively.

3.4 Estimation of Weightages by Each Model

For the AHP; the GL and RL were marked with the highest weightage (23.8% and 21.6% weightages) and the Growth Rate (GR) (1.8% weightage) was marked with the lowest weightage. For the AHP, the AS was marked with 18.4% weightage; DR was marked with 11% weightage, DRRD was identified with 8.2% weightage; RRD was identified with 5.7% weightage; TP, LR and SR were marked with 4%, 3% and 2.5% weightages respectively Table 2. The layers were reclassified as 9, 8, 7,6,4,3,3,2 and 2 respectively. The GL and RL were marked with the highest code as those two layers were marked with the highest priority. Similarly, the SR and GR were identified with the lowest code (Code 2) as those were identified with the lowest priority. For the LM, KNN, and EN model, the RL was marked with the highest priority, and the LR was identified with the lowest priority. For the LM model, DRRD, DR, AS, GR, TP, SR, GL, were marked with 32.01, 18.80, 17.45, 11.70, 7.09, 4.98, 2.90, and 0.76 weightages. For the EN model, DRRD, DR, GR, AS, SR, GL, and RRD were identified with 23.14, 13.64, 12.07, 10.80, 5.56, 4.17, 0.57, and 0.02 weightages. Similarly, for the KNN model, SR, GR, TP, LR, RRD, GL, DR, DRRD, and AS were identified with 42.91, 15.80, 10.31, 7.47, 4.09, 1.77, 0.50, and 0.16 weightages. For the ensemble model, GR, LR, SR, TP, AS, DR, DRRD, RRD, RL, and GL, were marked with 6.40, 1.60, 8.40, 4.20, 7.20,6.80, 9.80, 1.60, 4.96 and 4.50 weightages Figure 13. The detailed weightages for the AHP and other models were portrayed in the Table 2.

Table 2. Estimation of Priority of Different Indicators through AHP

Parameters	GL	RL	AS	DR	DRRD	RRD	TP	LR	SR	GR	Priority (%) Weightage
GL	1	1	2	3	5	4	5	6	7	7	23.8% (9)
RL	1	1	1	2	4	5	6	7	7	8	21.6% (8)
AS	0.5	1	1	2	3	4	5	6	7	8	18.4% (7)
DR	0.33	0.5	0.5	1	2	2	3	4	5	6	11% (6)
DRRD	0.2	0.3	0.3	0.5	1	1	3	4	5	7	8.2% (5)
RRD	0.25	0.2	0.3	0.5	1	1	1	2	3	4	5.7% (4)
TP	0.2	0.2	0.2	0.33	0.33	1	1	1	2	3	4% (3)
LR	0.17	0.1	0.2	0.25	0.25	0.5	1	1	1	2	3% (3)
SR	0.14	0.1	0.1	0.2	0.2	0.33	0.5	1	1	2	2.5% (2)
GR	0.14	0.1	0.1	0.17	0.14	0.25	0.33	0.5	1	1	1.8% (2)



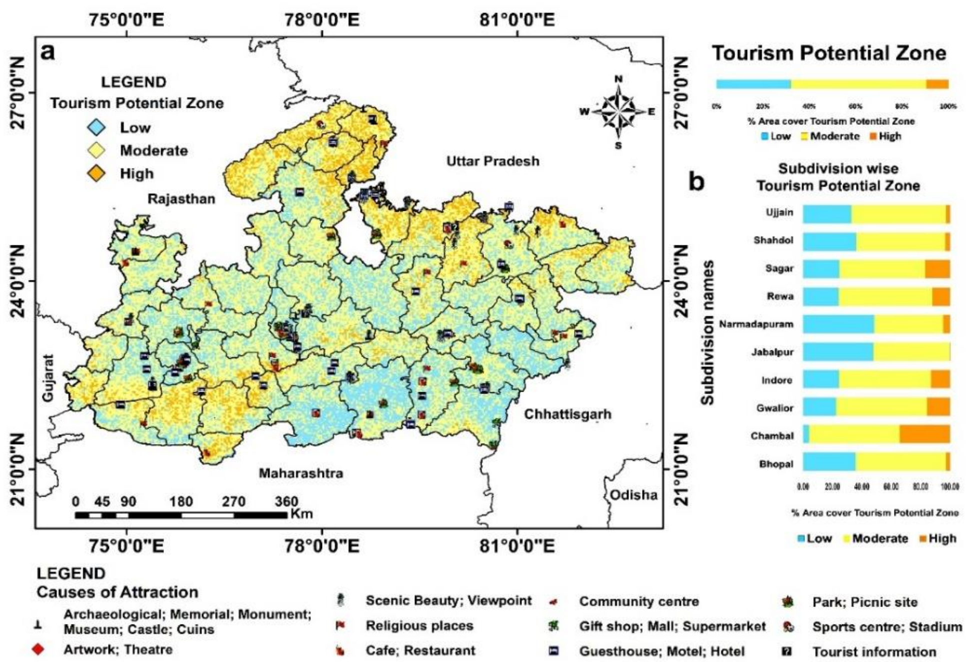
Source: Authors, 2023

Figure 13. Variable Importance Plot of different models

3.5 Tourism Potential Zone (TPZs)

For all models, the TPZs were classified into the High, Moderate and Low TPZs. Using the AHP model; 31.92% area of the Madhya Pradesh was identified as the low tourism potentiality; 58.69% area was delineated under the moderate TPZ and 9.39% area was marked with high tourism potentiality Figure 14a. In case of the LM model, 18.71% area was marked with the low tourism potentiality, 58.20% area was marked with moderate tourism potentiality, and 23.09% area was demarcated as the high tourism potentiality Figure 15a. For the EN model, 12.83% area, 66.66% area and 20.51% area were demarcated as the low, moderate and high tourism potentiality respectively Figure 16a. With the help of the KNN model, 15.56% area was marked as the low tourism potentiality, 57.16% area with the moderate tourism potentiality and 27.28% area was identified as with high tourism potentiality Figure 17a. For the ensemble model, 23.19% area was demarcated as the low TPZ, 65.11% area as the moderate TPZ, and 11.70% area was delineated as the high TPZ Figure 18a. For each case, the Northern, South-western and middle South-north stretches were identified with the high tourism potentiality. On the other hand, the South-western portions in each case were marked with the relatively low tourism potentiality. The North-western portions are demarcated with the moderate tourism potentiality in each case Figure 14a, 15a, 16a, 17a.

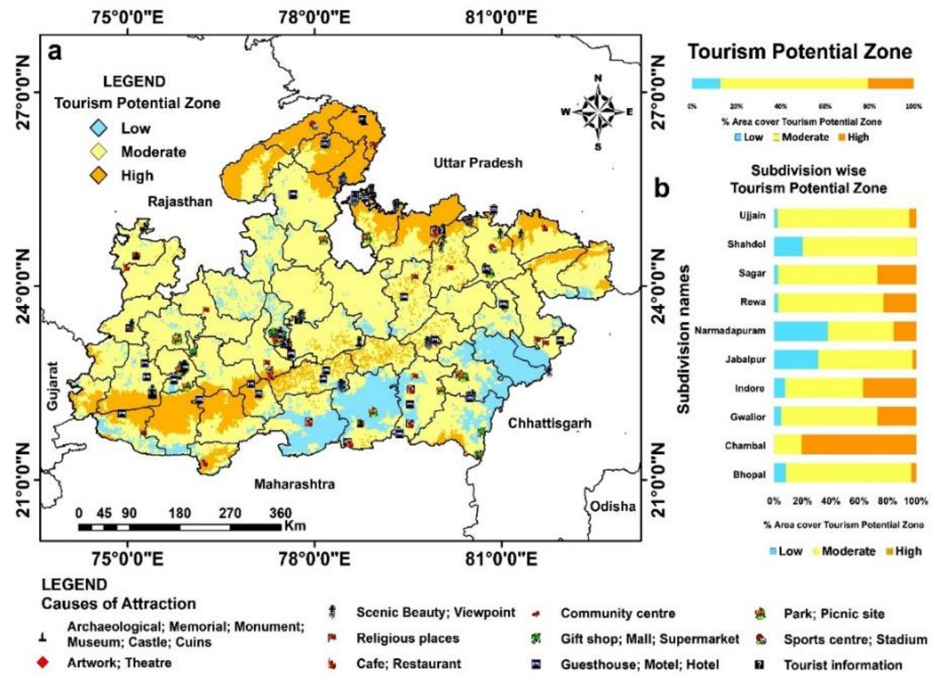
In case of AHP based TPZ; the high tourism potentiality dominated at Chambal (34.37% area); Gwalior (15.46% area), Indoor (13.02% area), Rewa (11.84% area) and Sagar (16.77% area) subdivisions. For the LM model; the high TPZ was trounced at the Chambal (81.65% area), Gwalior (27.94% area), Indore (37.56% area), Narmadapuram (17.81% area). Rewa (32.11% area), and Sagar (32.75% area) subdivisions. Similar feature exists for the KNN and Elastic Net models also; but the percentage area differs slightly. For the KNN model, Chambal, Gwalior, Indore were identified with the higher tourism potential with 94.07%, 33.26% and 32.58% area respectively. Narmadapuram, Rewa and Sagar were marked with 29.11%, 23.48% and 56.84% area respectively under the high TPZ. For the Elastic Net model, Chambal, Gwalior, Indore, Nardapuram, Rewa, and Sagar were marked with 80.63%, 27.64%, 37.45%, 16.24%, 23.27% and 27.48% area Figure 14b, 15b, 16b, 17b. For the ensemble model, 18% districts were marked under the high TPZ category. For the Bhopal, Gwalior and Rewa each, 2 districts were categorized under the high TPZ Figure 18b, 18c. These sections have moderate relief, and a better accessibility and connectivity network through high road-railway density. Apart from it, the distance from the river, the total population and population growth rate are relatively less. The sex ratio and literacy rate are also high in these sections of the study area.



Source: Authors, 2023

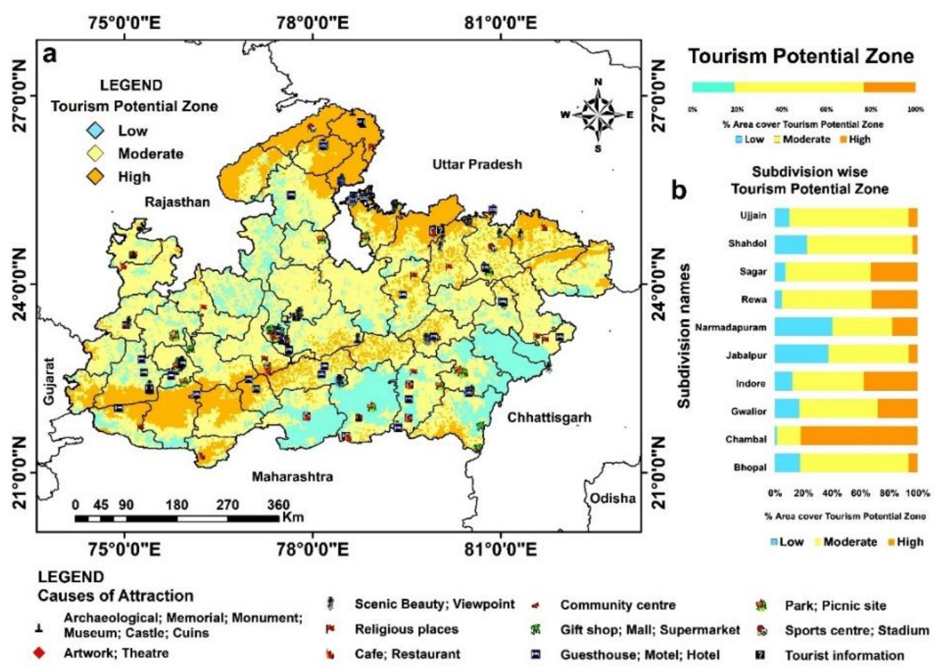
Figure 14. a) TPZ by the AHP technique b) Sub-division wise TPZ

In case of the AHP based model; the Low TPZ (LTPZ) dominated for the Jabalpur (47.83% area), Narmadapuram (48.43% area), Rewa (24.33% area), Sagar (24.75% area), Shadal (36.23% area), Bhopal (35.94% area), Gwalior (22.57% area) and Indore (24.41% area) subdivisions. The Chambal district are identified with the lowest areal coverage (4.09% area) of LTPZ. For the LM model; the low TPZ subjugates for the Bhopal (17.85% area), Gwalior (17.44% area), Indore (12.52% area), Jabalpur (37.68% area), Narmadapuram (40.39% area), Shahdol (22.45% area) and Ujjain (10.43% area) districts. The Chambal district are identified with the lowest areal coverage (1.63% area) of LTPZ. For the Elastic Net model, Jabalpur (31.02% area), Narmadapuram (37.71% area), Shadol (19.95% area) and Bhopal (8.09% area) districts were marked with higher LTPZ. For the KNN model, Indore (11.56% area), Jabalpur (64.04% area), Narmadapuram (36.78% area), and Shadol (17.13% area) were noticed with higher areal (%) coverage of LTPZ Figure 14b, 15b, 16b, 17b. Remaining districts were marked with less than 5% area in this category. For the ensemble model, 32% districts of the study area were categorized under the low TPZ. Approximately 2 to 4 districts of Bhopal, Indore, Jabalpur, Narmadapuram, and Shadol subdivisions were marked with low TPZ Figure 18b, 18c.



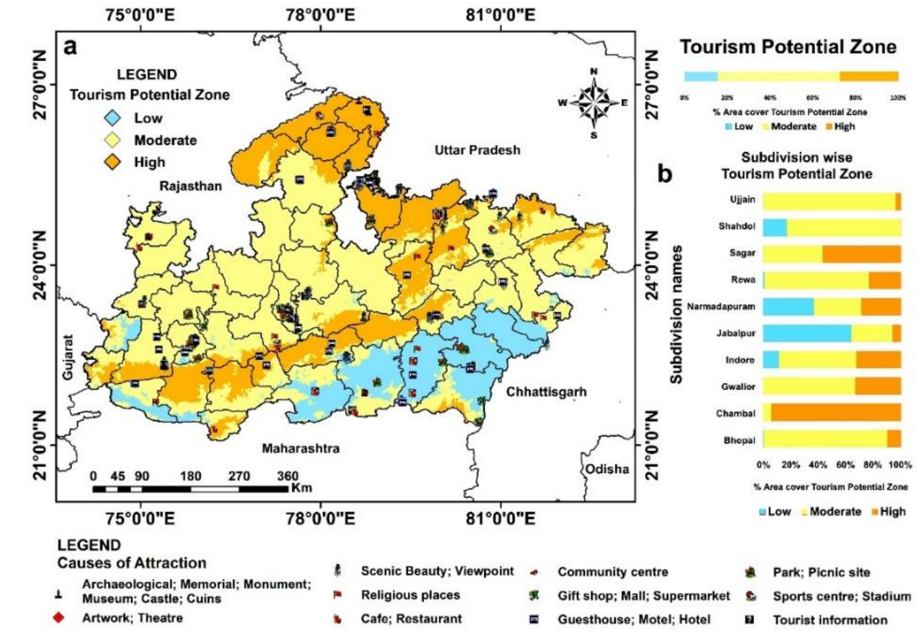
Source: Authors, 2023

Figure 15. a) TPZ by the LM model b) Subdivision wise TPZ



Source: Authors, 2023

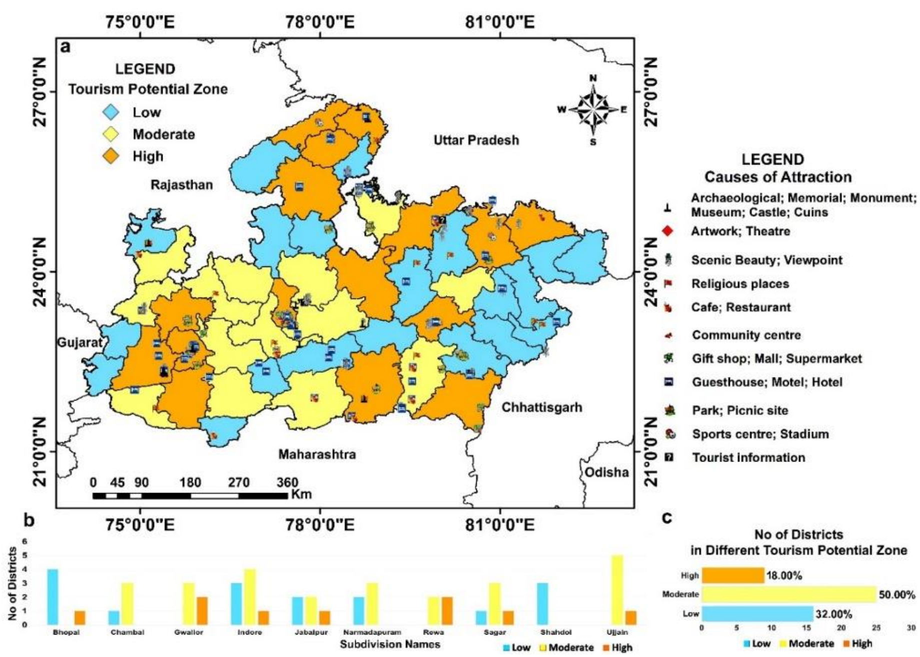
Figure 16. a) TPZ by EN model b) Sub-division wise TPZ



Source: Authors, 2023

Figure 17. a) TPZ by KNN model b) Sub-division wise TPZ

In case of the AHP based model; the lowest Moderate TPZ (MTPZ) was marked (46.97% area) for the Narmadapuram District. Other districts were noticed with 58% to 65% areal coverage. For the LM, KNN and Elastic Net model, the Chambal district was marked with the lowest areal coverage (16.71% area; 5.93% and 19.25% area). Remaining districts were identified with 40% to 80% areal coverage Figure 14b, 15b, 16b, 17b. For the ensemble model, approximately 50% area was demarcated under the moderate TPZ. Approximately, 2 to 6 districts from each subdivision (except, Bhopal and Shadol) fall under this category Figure 18b, 18c.



Source: Authors, 2023

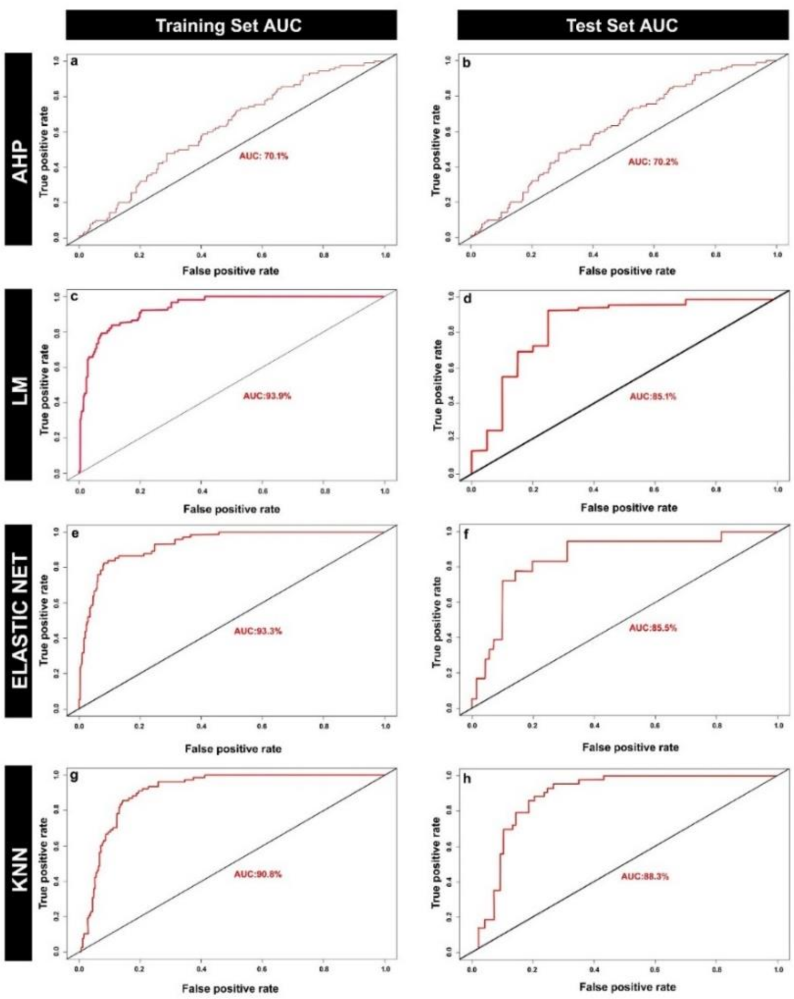
Figure 18. a) TPZ by Ensemble Model b) share of number of districts under different categories of TPZ in different subdivisions c) Percentage share of number of districts under TPZ

DOI: 10.14710/geoplanning.12.1.95-122

3.6 Accuracy Assessments

For the AHP model; the overall predictive accuracy for the training and test set were marked as 83.1% and 87.3% respectively. The AUC for the ROC curve for both training and test set were marked as 73.1% Figure 19a and 73.4% Figure 19b respectively. For the LM model; the overall accuracy of the model was 90.6% for the training data and 81.4% for the test set. For the LM model, the AUC for the ROC curve for both training and test set were marked as 93.9% Figure 19c and 85.1% Figure 19d respectively.

The overall accuracy for the Elastic net model were identified as the 90.3% and 86.5%, respectively for the training and test sets. For the Elastic Net model, the AUC for the ROC curve for both training and test set were marked as 93.3% Figure 19e and 85.5% Figure 19f respectively. For the KNN model, 83.5% and 85.7% accuracy were marked for the training and test data respectively. For this model the AUC were marked as 90.8% Figure 19g and 88.3% Figure 19h respectively. For the ensemble model, the overall accuracy was attained as 91.2%for the training set and 89.1% for the test set respectively. RMSE for the ensemble model was achieved as 0.33 and 0.46 for training and test set respectively. The AUC was marked as 94.5% area Figure 20a for the training set and 89.4% area for the test set Figure 20b. The combined RMSE was lowest for the ensemble mode (i., e., 0.79); whereas it was higher for the LM models. For the AHP; the combined RMSE value was moderate Table 4. Further, the combined AUC value was the highest for the ensemble model; followed by KNN, LM, EN and AHP model Table 5. Therefore, the ensemble model outperformed the others for the TPZ identification in this region.



Source: Authors, 2023

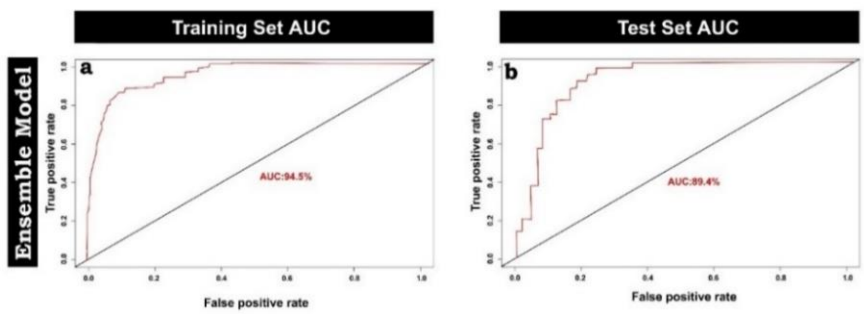
Figure 19 AUC-ROC measurement for different models

Table 4. RMSE for each model

Model	RMSE for the Training Set	RMSE for the Test Set	Combined RMSE
LM	0.36	0.96	1.32
AHP	0.37	0.71	1.08
KNN	0.37	0.46	0.83
Elastic Net	0.37	0.5	0.87
Ensemble	0.33	0.46	0.79

Table 5. AUC for each model

ROC-AUC value				
Model name	AUC for the Training Set	AUC for the Test Set	Combined AUC	Level of Accuracy
Ensemble	94.50%	89.40%	183.90%	Highest
KNN	90.80%	88.30%	179.10%	
LM	93.90%	85.10%	179%	To
EN	93.30%	85.50%	178.80%	
AHP	70.10%	70.20%	140.30%	Lowest



Source: Authors, 2023

Figure 20 AUC-ROC for the ensemble model a) Training Set b) Test Set

Source: Authors, 2023

Figure 21 Tourist Spot Identification by Non-Participant Observation Technique and Google Earth Imagery

3.7 Detailed Discussion and Implication of the Research

The research has several implications via following significant aspects at first, theoretically this research provides an engaging real-world application of Leiper's Tourism System Theory (Leiper, 1990) and Michael Porter's Diamond Model (Porter, 1998). The tourist, the generating region, the transit route, the destination region, and the tourism industry are the five main components that make up Leiper's model, which describes tourism as a dynamic system. The study successfully maps and quantifies these interrelated components throughout Madhya Pradesh utilising Multi-Criteria Decision Making (MCDM) approaches and machine learning algorithms, offering a methodical examination of how tourism operates as an integrated whole. Further this research implicates the Diamond Model of Tourism by evaluating and enhancing the regional competitiveness of tourism destinations through a structured, data-driven approach. The Diamond Model, adapted from Michael Porter's original framework (Porter, 1998), comprises four key determinants—factor conditions, demand conditions, related and supporting industries, —that collectively influence a destination's tourism competitiveness.

Through the application of Multi-Criteria Decision Making (MCDM) methods and machine learning algorithms, this study quantifies and spatially analyzes these determinants across the diverse regions of Madhya Pradesh. Factor conditions, including geology, relief, aspect, distance from river, cultural such as density of tourist spots, sex ratio, literacy rate, total population and infrastructural such as road and railway density, distance from road and railway are analyzed through geospatial and socio-economic information, determining areas with high tourism value. Demand conditions are indirectly considered by taking into account accessibility and connectivity, affecting tourist movements from domestic and international markets. Through the integration of these factors into a complete map of tourism potential zones, the research not only strengthens the theoretical model of the Diamond Model but also makes it more applicable to regional tourism planning. The inclusion of machine learning provides predictive and adaptive functions to the model, allowing stakeholders to foresee shifts in tourism demand and infrastructure requirements. Therefore, this study closes the gap between theory and practice by converting the conceptual dimensions of the Diamond Model into a working tool for destination development, competitive positioning, and strategic investment in Madhya Pradesh's tourism industry.

Further, there are many studies on tourism in Madhya Pradesh, including resource development (Pandey et al., 2014), social media's impact on tourism (Gohil, 2015), art and craft tourism (Kumar et al., 2023), and tourism's economic effects (Sharma, 2019), role of mass tourism (Chandravanshi & Jain, 2023; Gohil, 2015) development of sustainable tourism sector (Kishnani, 2022), eco-tourism (Ahmad & Pandey, 2016) for the state of Madhya Pradesh, but hardly any research work is available on tourism potential zone identification on this tract. Therefore, the tourism potential zone identification for the state of Madhya Pradesh is a noble attempt. In case of AHP based TPZ; the high tourism potentiality dominated at Chambal (34.37% area); Gwalior (15.46% area), Indoor (13.02% area), Rewa (11.84% area) and Sagar (16.77% area) subdivisions. For the LM based TPZ; the high TPZ was trounced at the Chambal (81.65% area), Gwalior (27.94% area), Indore (37.56% area), Narmadapuram (17.81% area). Rewa (32.11% area), and Sagar (32.75% area) subdivisions.

Similar feature exists for the KNN and EN models also; but the percentage area differs slightly. For the ensemble model, 18% districts were marked under the high TPZ category. For the Bhopal, Gwalior and Rewa each, 2 districts were categorized under the high TPZ. Higher tourism potential in any region is boosted by the low aspect (Vijay et al., 2016), moderate to low relief (Li et al., 2024), and closer proximity (small distance) to river water (Maaiah et al., 2023). Further, this higher tourism potential is amplified by moderate to low road and railway density (Sang et al., 2022) and the closest proximity (small distance) to road and railway (Rolando & Scandiffio,2022). These portions also marked with comparatively low total population (Chen et al., 2019), population density (Raha et al., 2021) but higher literacy rate (Chen & Li, 2023; Raha et al., 2021). Chambal, Gwalior, Indoor, Rewa and Sagar are identified with a popular tourist circuit (Olivelle, 2006).

The starting point for this travel circuit is Indore. The Holkar kings' seat was this thriving trading town. From Delhi and Mumbai, it has excellent air, rail, and road connections. The Rajwada, the Palace, and the cenotaphs of the Holkar kings are some of its intriguing features. It is known as the "Mini Mumbai" due to the significance of its business activity (The Market Research Division, Department of Tourism, 2003). Chambal is well-known for the Canoeing Safari or White-Water Rafting (Kohli, 2002). Gwalior features a tall citadel with 14th-century AD Rajput residences and mediaeval monuments (Sijatha, 2017). An historic fort and cave provide a touch of heritage to Rewa's National Park which is known as the Bandhavgarh National Park (Lahiri et al., 2022). It is the greatest location to witness tigers in their native environment (Lahiri et al., 2023). Sagar becomes extremely attractive in the monsoon because of scenic beauty (Rakhra, 2023). Those above-mentioned regions bear the state's cultural heritage, are dominated by handlooms. The Madhya Pradesh tourism Development Corporation also took necessary steps to popularise these portions such as, development of hotels, lodges and dormitories for providing accommodation to tourists, facilities of accessibility and connectivity through efficient transport networks, developing the tourist information centres, and advertising and marketing of tourist places (Ministry of Tourism Govt. of India, 2023).

Additionally, the study used an integrated 9-step technique to identify tourist potential zones in the state of Madhya Pradesh. To the best of our knowledge, this study is a cutting-edge effort in the use of decision-making and machine learning models for the accurate prediction of tourist potential zones. The research gives a detailed and analytical view on the on-the dynamics of tourism in the region by examining the spatial distribution of tourist potential levels, classifying them as high, moderate, and low. We performed a thorough evaluation and comparison of both traditional Multiple Criteria Decision Making (MCDM) methods, such as the Analytic Hierarchy Process (AHP), and modern machine learning models such as Elastic Net (EN), Linear Regression (LM), and k-Nearest Neighbours (KNN). It's important to highlight that although the Analytic Hierarchy Process (AHP) is well-acknowledged, the adoption of machine learning models EN, LM, and KNN for prediction of tourism potential zone identification has been widely appreciated by scholars worldwide.

This research not only brings to the forefront the effectiveness of these various methodologies but also offers valuable perspectives on their real-world usefulness in forecasting tourism potential zones in the state of Madhya Pradesh. Furthermore, the intrinsic variety in ensemble model in this research acts as a buffer against the vagaries of uncertainty, protecting decision-making frameworks from the negative influence of outliers and noise. The ensemble model provides decision-makers with a more comprehensive and nuanced understanding of the underlying dynamics governing a given domain by leveraging the collective wisdom distilled from a variety of algorithmic perspectives, empowering them to make informed and judicious decisions amidst the tumult of uncertainty. Moreover, the synergistic interplay of constituent algorithms (i.e., AHP, LM, EN and KNN models) improve ensemble models' ability to infer complicated correlations contained within datasets, allowing for a more in-depth knowledge of underlying phenomena. This collective intelligence, created by the harmonic merger of many algorithms, generates a greater range of ideas, hence increasing the effectiveness of decision-making processes.

3. Conclusion

Madhya Pradesh is a very prominent and well-known tourist destination in India. The TPZ of Madhya Pradesh was explored in this research with the help of one MCDM (i.e., AHP), three machine learning (EN, LM and KNN models) and one ensemble models. The methodology was implemented here through 9 steps. First of all, total 11 layers (i.e., GL, RL, AS, DR, DRRD, RRD, SR, LR, TP, GR, and TS) was collected and prepared the raster layer in the GIS platform. Next, the multicollinearity was checked which is one of the basic prerequisites. Further residuals vs. fitted plot, normal Q-Q plot, residuals vs, leverage plot and scale location plot were checked. In each case, the trend line (red) was fitted with the dashed line, parallel with x axis. The residuals vs. leverage plot shows that there exist no influential significant extreme points exist in the dataset. At the third step, the data was normalized to 0 to 1. The data was partitioned into training and test set in a 70:30 ratio at the fourth step. The TS was set as the target variable and the others were set as input variables in the fifth step. Next, the AHP, LM, EN and KNN models were applied in demarcation of TPZs. The ensemble model was prepared at seventh step by combining the AHP, LM, EN and KNN models. The proposed tourism potential maps were validated through the AUC-ROC curve and RMSE value.

The ensemble model appears as the best model as it was noticed with a low RMSE and higher AUC value. The northern, south-western, and middle regions emerge as high-potential areas, whilst the south-western edges were appeared with less potential. Meanwhile, the north-western expanse offers a scene of moderate potential. However, by adding more variables, the tourism potential map forecast accuracy may be further ameliorated. Furthermore, any alteration to the natural environment brought about by human activity or any changes in the natural phenomena such as, the relief, aspect, or distance from river, may alter the area's current status of tourism potentiality. Therefore, the tourism potential zone map should be updated annually by incorporating all the changes. The TPZ map can be used as the base data to support the planning and developmental activities in the state. To the best of our knowledge, this research is the first to highlight the tourism potentiality of the Madhya Pradesh using the decision making and machine learning models. This intricates the novelty of the research.

This research holds significant value in advancing sustainable tourism development by integrating Multi-Criteria Decision-Making (MCDM) techniques and machine learning to identify and prioritize regions with high tourism potential. This innovative approach enhances spatial planning and resource allocation by combining expert-driven criteria assessment with data-driven predictive capabilities, offering a more accurate and dynamic mapping of tourism zones. The study not only aids policymakers and tourism stakeholders in making informed decisions but also contributes to regional economic growth, heritage conservation, and balanced tourism distribution across the state, aligning with broader goals of sustainable development and digital governance.Through the integration of machine learning and Multi-Criteria Decision-Making (MCDM) methodologies, the research effectively advances sustainable tourism development by identifying and prioritising regions with high tourism potential. By fusing data-driven prediction capabilities with expert-driven criterion assessment, this novel method improves spatial planning and resource allocation while providing a more dynamic and accurate mapping of tourism zones.

The study supports balanced tourism distribution throughout the state, historical preservation, and regional economic growth in addition to helping policymakers and tourism stakeholders make well-informed decisions. These outcomes are in line with the larger objectives of sustainable development and digital governance. The TPZ map can be used as the base data to support the planning and developmental activities in the state. To the best of our knowledge, this research is the first to highlight the tourism potentiality of the Madhya Pradesh using the decision making and machine learning models. This intricates the novelty of the research. Future researchers should consider expanding the spatial and thematic resolution of datasets to improve model precision and scalability. They should integrate dynamic datasets such as real-time tourist footfall, social media sentiment, environmental change indicators, and transportation network updates to better capture evolving tourism patterns. Incorporating participatory GIS and crowd-sourced local knowledge can further enhance model accuracy and stakeholder relevance. Comparative analyses between different MCDM techniques (e.g., TOPSIS, PROMETHEE) and ML algorithms (e.g., Random Forest, Gradient Boosting, Neural Networks) should be systematically conducted to identify the most robust and context-sensitive combinations. Researchers are encouraged to explore ensemble modeling approaches to minimize uncertainty and enhance predictive validity. Future studies should also address the interpretability and explainability of ML outputs to facilitate practical implementation by policymakers and tourism planners. Furthermore, integrating sustainability indicators—ecological, cultural, and socioeconomic—into the TPZ framework can ensure that development strategies align with long-term conservation goals. Finally, establishing a temporal component in TPZ models to analyze seasonal variations and long-term trends could significantly improve planning effectiveness and the adaptability of tourism strategies in Madhya Pradesh.

4. References

Aburomman, A. A., & Ibne Reaz, M. B. (2016). A novel SVM-kNN-PSO ensemble method for intrusion detection system. *Applied Soft Computing*, *38*, 360–372. [Crossref]

Acharya, A., Mondal, B. K., Bhadra, T., Abdelrahman, K., Mishra, P. K., Tiwari, A., & Das, R. (2022). Geospatial Analysis of Geo-Ecotourism Site Suitability Using AHP and GIS for Sustainable and Resilient Tourism Planning in West Bengal, India. *Sustainability*, *14*(4), Article 4. [Crossref]

Ahmad, S. Z., & Pandey, D. N. (2016). Development and assessment of ecotourism in Panna National Park (Madhya Pradesh), India. *International Journal of Applied and Universal Research*, *3*(1).1-5.

Aijaz, D. A. (2022). Tourism and resource development: A study of Madhya Pradesh. *International Journal of Economic Perspectives*, *16*(9), Article 9.

Al-Areeq, A. M., Abba, S. I., Halder, B., Ahmadianfar, I., Heddam, S., Demir, V., Kilinc, H. C., Farooque, A. A., Tan, M. L., & Yaseen, Z. M. (2023). Flood Subsidence Susceptibility Mapping using Elastic-net Classifier: New Approach. *Water Resources Management*, *37*(13), 4985–5006. [Crossref]

Atun, R. A., Nafa, H., & Türker, Ö. O. (2019). Envisaging sustainable rural development through 'context-dependent tourism': Case of Northern Cyprus. *Environment, Development and Sustainability*, *21*(4), 1715–1744. [Crossref]

Ayesha, S., Hanif, M. K., & Talib, R. (2020). Overview and comparative study of dimensionality reduction techniques for high dimensional data. *Information Fusion*, *59*, 44–58. [Crossref]

Banerjee, A. (2014). Human Resource Development in Tourism Industry in India: A Case Study of Jet Airways India Ltd. 1.

Banik, S., & Mukhopadhyay, M. (2022). Model-based strategic planning for the development of community-based tourism: A case study of Ayodhya Hills in West Bengal, India. *GeoJournal*, *87*(2), 1349–1365. [Crossref]

Bansal, S., Garg, I., & Sharma, G. D. (2019). Social Entrepreneurship as a Path for Social Change and Driver of Sustainable Development: A Systematic Review and Research Agenda. *Sustainability*, *11*(4), Article 4. [Crossref]

Boateng, E. Y., Otoo, J., & Abaye, D. A. (2020). Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review. *Journal of Data Analysis and Information Processing*, *8*(4), Article 4. [Crossref]

Brida, J. G., Lanzilotta, B., Moreno, L., & Santiñaque, F. (2018). A non-linear approximation to the distribution of total expenditure distribution of cruise tourists in Uruguay. *Tourism Management*, *69*, 62–68. [Crossref]

Cankurt, S., & Subasi, A. (2022). Tourism demand forecasting using stacking ensemble model with adaptive fuzzy combiner. *Soft Computing*, *26*(7), 3455-3467. [Crossref]

Chandravanshi, J., & Jain, N. (2023). Indigenous Art of Bastar District of Chhattisgarh: A Study of the Government Initiations for Artisans. *Splint International Journal of Professionals*, *10*(1), 25–34. [Crossref]

Chavan, G., & Momin, B. (2017). An integrated approach for weather forecasting over Internet of Things: A brief review. 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 83–88. [Crossref]

Chen, H.-L., Yang, B., Wang, G., Liu, J., Xu, X., Wang, S.-J., & Liu, D.-Y. (2011). A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. *Knowledge-Based Systems*, *24*(8), 1348–1359. [Crossref]

Chen, H., Wang, S., Guo, H., Lin, H., Zhang, Y., Long, Z., & Huang, H. (2019). Study of marine debris around a tourist city in East China: Implication for waste management. *Science of the Total Environment*, *676*, 278-289. [Crossref]

Chen, Y., Li, Y., Gu, X., Chen, N., Yuan, Q., & Yan, M. (2021). Evaluation of tourism development potential on provinces along the Belt and Road in China: Generation of a comprehensive index system. *Land*, *10*(9), 905. [Crossref]

Dedík, M., Šperka, A., Čamaj, J., & Zábovská, K. (2022). Traffic Potential Evaluation of the Regions in the Context of Rail Passenger Transport – A Case for Slovak Republic. *Promet – Traffic&Transportation*, *34*(2), 309–321. [Crossref]

Dong, J.-J., Tung, Y.-H., Chen, C.-C., Liao, J.-J., & Pan, Y.-W. (2011). Logistic regression model for predicting the failure probability of a landslide dam. *Engineering Geology*, *117*(1), 52–61. [Crossref]

Gohil, D. N. (2015). Role and Impact of Social Media in Tourism: A Case Study on the Initiatives of Madhya Pradesh State Tourism. *International Journal of Research in Economics and Social Sciences*, *5*(4).8-15.

Harianto, S. P., Masruri, N. W., Winarno, G. D., Tsani, M. K., & Santoso, T. (2020). Development strategy for ecotourism management based on feasibility analysis of tourist attraction objects and perception of visitors and local communities. Biodiversitas *Journal of Biological Diversity*, *21*(2), Article 2. [Crossref]

Huang, P., Hou, M., Sun, T., Xu, H., Ma, C., & Zhou, A. (2024). Sustainable groundwater management in coastal cities: Insights from groundwater potential and vulnerability using ensemble learning and knowledge-driven models. *Journal of Cleaner Production*, *442*, 141152. [Crossref]

Kohli, M. S. (2002). *Mountains of India: Tourism, adventure and pilgrimage.* Indus Publishing.

Li, Z., Ye, S., & Xi, J. (2024). Assessing Risks in Cross-Regional Tourism Corridors: A Case Study of Tibetan Plateau Tourism. *ISPRS International Journal of Geo-Information*, *13*(6), 171. [Crossref]

Kishnani, N. (2019). Sustainable Development of Ecotourism in Madhya Pradesh: Prospects and Problems. *Journal of Tourism*, *20*(1), 55-70. https://jothnbgu.in/article/JOT%20june%202019.pdf#page=61.

Kumar, P., Madaan, S., & Bhargava, G. (2023). An Insight About Bell Metal Craft—Scopes of Improvement and Promotion. In A. Chakrabarti & V. Singh (Eds.), Design in the Era of Industry 4.0, Volume 2 (pp. 965–975). Springer Nature. [Crossref]

Lahiri, N., Rajani, M. B., Sanyal, D., Banerjee, S., & Tiwari, S. (2023). Tracing ancient itinerants and Early Medieval rulers in the forests of Bandhavgarh. *South Asian Studies*, *39*(1), 76-99. [Crossref]

Lahiri, N., Rajani, M. B., Sanyal, D., & Banerjee, S. (2022). Exploring the forest and mapping its archaeology: Bandhavgarh National Park and Tiger Reserve, India. *Current Science*, *123*(6), 772-780. [Crossref]

Leiper, N. (1990). Tourist attraction systems. *Annals of Tourism Research*, *17*(3), 367-384. [Crossref]

Maaiah, B., Al-Badarneh, M., & Al-Shorman, A. (2023). Mapping potential nature-based tourism in Jordan using AHP, GIS and remote sensing. *Journal of Ecotourism*, *22*(2), 260-280. [Crossref]

Mamun, A. A., & Mitra, S. (2012). A Methodology for Assessing Tourism Potential: Case Study Murshidabad District, West Bengal, India. 2(9). *International Journal of Scientific and Research Publications*. 1-8.

Marrocu, E., & Paci, R. (2013). Different tourists to different destinations. Evidence from spatial interaction models. *Tourism Management*, *39*, 71-83. [Crossref]

Ministry of Tourism Govt. of India (2023). *Assessment of Tourist Destinations in Areas of Infrastructure and Cleanliness.* https://tourism.gov.in/sites/default/files/202402/Final%20Report_Assessment%20of%20Tourist%20Destination_%20English.pdf.

Mir, A., & Nasiri, J. A. (2018). KNN-based least squares twin support vector machine for pattern classification. *Applied Intelligence*, *48*(12), 4551–4564. [Crossref]

Mirsanjari, M. M., & Mirsanjari, M. O. (2012). *Study of Strategic Eco-Tourism Potential Based on Sustainable Development and Management* (SSRN Scholarly Paper 2054154). https://papers.ssrn.com/abstract=2054154

Mitra, R., Saha, P., & Das, J. (2022). Assessment of the performance of GIS-based analytical hierarchical process (AHP) approach for flood modelling in Uttar Dinajpur district of West Bengal, India. *Geomatics, Natural Hazards and Risk*, *13*(1), 2183–2226. [Crossref]

Mittal, K., Aggarwal, G., & Mahajan, P. (2019). Performance study of K-nearest neighbor classifier and K-means clustering for predicting the diagnostic accuracy. *International Journal of Information Technology*, *11*(3), 535–540. [Crossref]

Mohanapriya, M., & Lekha, J. (2018). Comparative study between decision tree and knn of data mining classification technique. *Journal of Physics: Conference Series*, *1142*(1), 012011. [Crossref]

Naghibi, S. A., Pourghasemi, H. R., & Dixon, B. (2015). GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. *Environmental Monitoring and Assessment*, *188*(1), 44. [Crossref]

Natalia, P., Clara, R. A., Simon, D., Noelia, G., & Barbara, A. (2019). Critical elements in accessible tourism for destination competitiveness and comparison: Principal component analysis from Oceania and South America. *Tourism Management*, *75*, 169–185. [Crossref]

Nugrahaeni, R. A., & Mutijarsa, K. (2016). Comparative analysis of machine learning KNN, SVM, and random forests algorithm for facial expression classification. *2016 International Seminar on Application for Technology of Information and Communication (ISemantic)*, 163–168. [Crossref]

Okfalisa, Gazalba, I., Mustakim, & Reza, N. G. I. (2017). Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification. *2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 294–298. [Crossref]

Olivelle, P. (Ed.). (2006). *Between the empires: society in India 300 BCE to 400 CE*. Oxford University Press.

Ovreiu, A. B., Bărsoianu, I. A., Comănescu, L., & Nedelea, A. (2018). *Assessing the Accessibility of Relief for Tourism Activities. Case Study-Cozia Massif (Southern Carpathians, Romania). GeoJournal of Tourism & Geosites, 22 (2), 509–523.* [Crossref]

Pandey, D., Mishra, S., & Shukla, S. (2014). Tourism and resource development: A case study of Madhya Pradesh. *International Journal of Biological Research*, *2(*2), 149. [Crossref]

Porter, M.E. (1998), The Competitive Advantage of Nations, The Free Press, New York, NY.

Pouya, S., & Başkaya, F. A. T. (2018). Residents' Perceptions of Riverine Landscape Changes; Case Study of Beykoz Stream/ Istanbul. *Anadolu University Journal of Science and Technology A - Applied Sciences and Engineering*, *19*(2), Article 2. [Crossref]

Raha, S., & Gayen, S. K. (2022a). Application of Analytic Hierarchy Process and weighted sum techniques for green tourism potential mapping in the Gangetic West Bengal, India. *GeoJournal*. [Crossref]

Raha, S., Mondal, M., & Gayen, S. K. (2021). Ecotourism potential zone mapping by using analytic hierarchy process (AHP) and weighted linear algorithm: A study on West Bengal, India. *Journal of Geographical studies*, *5*(2), 44-64. [Crossref]

Raha, S., & Gayen, S. K. (2022b). Tourism Potentiality Zone Mapping by Using the AHP Technique: A Study on Bankura District, West Bengal, India. *Journal of Geographical Studies*, *6*(2), 58–85. [Crossref]

Raha, S., Gayen, S. K., & Deb, S. (2024). Harnessing Machine Learning and Ensemble Models for Tourism Potential Zone Prediction for the Assam State of India. *Journal of Advanced Geospatial Science & Technology*, *4*(2), 29–78. [Crossref]

Rahman, M. M. (2021). Inbound tourism in Bangladesh: The trend of competitiveness. *Bangladesh Journal of Public Administration*, *29*(2), Article 2. [Crossref]

Rajan, S. I. (Ed.). (2018). *India migration report 2018: Migrants in Europe*. Taylor & Francis.

Rakhra, D. I. K. (2020). *Marketing Innovations in Tourism Industry: An Empirical Study of Madhya Pradesh*. Red'shine Publication. Pvt. Ltd.

Ramírez-Guerrero, G., García-Onetti, J., Arcila-Garrido, M., & Chica-Ruiz, J. A. (2021). A Tourism Potential Index for Cultural Heritage Management through the Ecosystem Services Approach. *Sustainability*, *13*(11), Article 11. [Crossref]

Rolando, A., & Scandiffio, A. (2022). Multimodal access to minor places in heritage-rich landscapes: GIS mapping to define slow-tourism routes from the stations in the railway networks in-between Turin and Milan. *Sustainability*, *14*(23), 15723. [Crossref]

Ruda, A. (2016). Exploring Tourism Possibilities using GIS-based Spatial Association Methods. *Geographia Technica*, *11*(2), 87–101. [Crossref]

Rutherford, J., Kobryn, H., & Newsome, D. (2015). A Case Study in the Evaluation of Geotourism Potential through Geographic Information Systems: Application in a Geology-Rich Island Tourism Hotspot. *Current Issues in Tourism*, *18*(3), 267-285. [Crossref]

Saaty, T. L. (1980). The Analytic Hierarchy Process. Mcgraw Hill, New York, 70, 34.

Sahabi Abed, S., & Matzarakis, A. (2018). Quantification of the Tourism Climate of Algeria Based on the Climate-Tourism-Information-Scheme. *Atmosphere*, *9*(7), Article 7. [Crossref]

Sahani, N. (2019). Assessment of ecotourism potentiality in GHNPCA, Himachal Pradesh, India, using remote sensing, GIS and MCDA techniques. *Asia-Pacific Journal of Regional Science*, *3*(2), 623-646. [Crossref]

Saner, R., Yiu, L., & Filadoro, M. (2019). Tourism Development in Least Developed Countries: Challenges and Opportunities. In Sustainable Tourism: Breakthroughs in Research and Practice (pp. 94–120). IGI Global. [Crossref]

Sang, K., Fontana, G. L., & Piovan, S. E. (2022). Assessing railway landscape by ahp process with gis: A study of the yunnan-vietnam railway. *Remote Sensing*, *14*(3), 603. [Crossref]

Sardar, S. (2021). Human resources development in tourism and the role of government: The case of Indian tourism. *Journal of the Maharaja Sayajirao University of Baroda*, *55*(1), 40-50. https://www.researchgate.net/profile/Sanjib-Sardar-2/publication/350383345_Human_Resources_Development_in_Tourism_and_the_Role_of_Government_The_Case_of_Indian_Tourism/links/605cab4f299bf173676b9093/Human-Resources-Development-in-Tourism-and-the-Role-of-Government-The-Case-of-Indian-Tourism.pdf.

Sarker, S. (2018). Resident's Awareness Towards Sustainable Tourism for Ecotourism Destination in Sundarban Forest, Bangladesh. *Pacific International Journal*, *1*(1), Article 1. [Crossref]

Sharma, B. (2019). Economic impacts of rural tourism in Madhya Pradesh. *International Journal of Advance Research, Ideas and Innovations in Technology*, *5*(1). 462-466.

Smith, M. K. (2015). Issues in Cultural Tourism Studies. Routledge. [Crossref]

Sofaer, H. R., Hoeting, J. A., & Jarnevich, C. S. (2019). The area under the precision-recall curve as a performance metric for rare binary events. *Methods in Ecology and Evolution*, *10*(4), 565–577. [Crossref]

Soomro, B. N., Xiao, L., Huang, L., Soomro, S. H., & Molaei, M. (2016). Bilayer Elastic Net Regression Model for Supervised Spectral-Spatial Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(9), 4102–4116. [Crossref]

Stępniak, M., Pritchard, J. P., Geurs, K. T., & Goliszek, S. (2019). The impact of temporal resolution on public transport accessibility measurement: Review and case study in Poland. *Journal of Transport Geography*, *75*, 8–24. [Crossref]

Suchting, R., Hébert, E. T., Ma, P., Kendzor, D. E., & Businelle, M. S. (2019). Using Elastic Net Penalized Cox Proportional Hazards Regression to Identify Predictors of Imminent Smoking Lapse. *Nicotine & Tobacco Research*, *21*(2), 173–179. [Crossref]

Sugiharti, E., Putra, A. T., & Subhan. (2020). Facial recognition using two-dimensional principal component analysis and k-nearest neighbor: A case analysis of facial images. *Journal of Physics: Conference Series*, *1567*(3), 032028. [Crossref]

Telfer, D. J., & Sharpley, R. (2015). *Tourism and development in the developing world*. Routledge.

The Market Research Division, Department of Tourism, (2003). *Twenty Years Perspective Plan of Tourism for the State of Madhya Pradesh (Final Report, Vol-I)*. *https://tourism.gov.in/sites/default/files/2020-04/MP.pdf*

Tien Bui, D., Le, K.-T. T., Nguyen, V. C., Le, H. D., & Revhaug, I. (2016). Tropical Forest Fire Susceptibility Mapping at the Cat Ba National Park Area, Hai Phong City, Vietnam, Using GIS-Based Kernel Logistic Regression. *Remote Sensing*, *8*(4), Article 4. [Crossref]

Trukhachev, A. (2015). Methodology for Evaluating the Rural Tourism Potentials: A Tool to Ensure Sustainable Development of Rural Settlements. *Sustainability*, *7*(3), Article 3. [Crossref]

Vairetti, C., Aránguiz, I., Maldonado, S., Karmy, J. P., & Leal, A. (2024). Analytics-driven complaint prioritisation via deep learning and multicriteria decision-making. *European Journal of Operational Research*, *312*(3), 1108–1118. [Crossref]

Vijay, R., Kushwaha, V. K., Chaudhury, A. S., Naik, K., Gupta, I., Kumar, R., & Wate, S. R. (2016). Assessment of tourism impact on land use/land cover and natural slope in Manali, India: a geospatial analysis. *Environmental Earth Sciences*, *75*, 1-9. [Crossref]

Wang, D., Niu, Y., & Qian, J. (2018). Evolution and optimization of China's urban tourism spatial structure: A highspeed rail perspective. *Tourism Management*, *64*, 218–232. [Crossref]

Woźniak, E., Kulczyk, S., & Derek, M. (2018). From intrinsic to service potential: An approach to assess tourism landscape potential. *Landscape and Urban Planning*, *170*, 209–220. [Crossref]

Yamin, M., Darmawan, A., & Rosyadi, S. (2021). Analysis of Indonesian tourism potentials through the sustainable tourism perspective in the new normal era. *Jurnal Hubungan Internasional*, *10*(1), 44-58. [Crossref]

Zhang, S., Li, X., Zong, M., Zhu, X., & Cheng, D. (2017). Learning k for kNN Classification. *ACM Transactions on Intelligent Systems and Technology*, *8*(3), 43:1-43:19. [Crossref]

Zhang, S., Zhang, Z., Yu, H., & Zhang, T. (2024). Assessment and Empirical Research on the Suitability of Eco-Tourism Development in Nature Reserves of China: A Multi-Type Comparative Perspective. *Land*, *13*(4), 438. [Crossref]

Zhao, H., Wang, Y., Duan, J., Huang, C., Cao, D., Tong, Y., Xu, B., Bai, J., Tong, J., & Zhang, Q. (2020). Multivariate Time-Series Anomaly *Detection via Graph Attention Network. 2020 IEEE International Conference on Data Mining (ICDM)*, 841–850. [Crossref]

Zhu, Q., Arabameri, A., Egbueri, J., & Agbasi, J. (2023). Integrated assessment of landslide susceptibility in the Kalaleh Basin, Golestan Province, Iran using novel SVR-GOA ensemble validated with BRT, ANN, and elastic net models. *ResearchSquare*. 1-27. [Crossref]