Int. J. Hum. Capital Urban Manage., 10(1): 121-134, Winter 2025

International Journal of Human Capital in Urban Management
(IJHCUM)

Homepage: http://www.ijhcum.net/

ORIGINAL RESEARCH PAPER

Statistical analysis based on a two-staged SEM-SVM approach for environmental noise annoyance prediction and identification of influencing factors

M.U. Rehman[1], M. Kamal[2], S.K. Tiwari[3,*]

[1]*Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Dammam, 32256, Saudi Arabia*

[2]*Department of Information Technology, College of Computing and Informatics, Saudi Electronic University, Dammam 32256, Saudi Arabia*

[3]*Balaji Institute of Modern Management, Sri Balaji University, Pune- 411 033, India*

ARTICLE INFO

Article History:
Received 06 March 2024
Revised 27 May 2024
Accepted 03 July 2024

Keywords:
Noise annoyance
Support Vector Machine (SVM)
Structural Equation Modeling (SEM)
Noise Sensitivity
Socio-acoustics Survey

DOI: 10.22034/IJHCUM.2025.01.08

ABSTRACT

BACKGROUND AND OBJECTIVES: Noise pollution is an environmental stressor that is mainly caused due to heavy transportation in urban scenarios. Traffic noise is a growing concern in urban environments, impacting public health and well-being. As urbanization expands, understanding and mitigating traffic-induced noise annoyance becomes increasingly critical. This study aimed to develop a machine-learning model for predicting traffic-induced noise annoyance in Riyadh, Saudi Arabia. The research explored the influence of factors like demographics, noise characteristics, and traffic conditions on noise annoyance.

METHODS: A survey was conducted at 21 locations in Riyadh, collecting data from 928 participants. The survey included questions on demographics (gender, age, education, marital status, profession), traffic conditions (traffic flow), and noise perception (transportation noise, noise sensitivity, perceived noisiness). The sampling method employed was a combination of stratified and random sampling. Stratified sampling was used to ensure that various demographic segments (e.g., different age groups, genders, and education levels) were proportionately represented in the survey. Structural Equation Modeling was used to analyze the collected data and identify factors significantly affecting noise annoyance. These significant factors were then used as input variables for a Support Vector Machine model designed to predict noise annoyance. The performance of the Support Vector Machine model was evaluated using Root Mean Square Error, Mean Absolute Error, and R-squared.

FINDINGS: The Structural Equation Model analysis revealed that gender, age, education level, traffic flow, noise from traffic, and individual noise sensitivity were significant contributors to noise annoyance. The developed Support Vector Machine model achieved a high level of accuracy with a root mean square error of 1.416 and a coefficient of determination of 0.90. Noise sensitivity emerged as the most crucial factor influencing noise annoyance.

CONCLUSION: This study demonstrates the effectiveness of machine learning, specifically the Support Vector Machine, in predicting traffic-induced noise annoyance. The findings highlight the importance of both individual characteristics and environmental factors in noise perception and can be valuable for urban planning and noise mitigation strategies, promoting a more noise-resilient city environment. For the community, urban planners and policymakers can use these findings to design silent areas by implementing noise barriers, optimizing traffic flow, and enforcing stricter noise regulations.

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
51	2	8

*Corresponding Author:
Email: *shashikanttiwari61@gmail.com*
Phone: 8778533124
ORCID: 0000-0003-4305-3092

Note: Discussion period for this manuscript open until April 1, 2025 on IJHCUM website at the "Show Article.

INTRODUCTION

Noise pollution is one of the major environmental stressors that adversely impacts human health, well-being, and quality of life (Gupta *et al.*, 2018). It originates from various sources such as transportation (road, air, and rail), construction, industrial activities, and recreational pursuits (Tiwari *et al*., 2023; Zengin *et al*., 2020). In Saudi Arabia (KSA), environmental noise, particularly from traffic, has emerged as a persistent issue (Al-Ghonamy, 2010; Al-Ghonamy, 2009; Zytoon, 2016; AlSaleem *et al.,* 2023; PME, 2015). This can be attributed to the rapid development and expansion of cities, along with a substantial increase in the population, leading to higher demands for transportation facilities (Al-Ghonamy, 2009). As a result, major roads in cities like Jeddah frequently witness traffic volumes exceeding 300,000 vehicles per day (Elshorbagy, 1999). In response to this growing concern, the Presidency of Meteorology and Environment (PME) took action from 2012 onwards. They introduced new Saudi standards for environmental noise, which included guidelines for community noise levels during daytime, evening, and nighttime. These standards were categorized based on the sensitivity of the area (PME, 2015). Moreover, noise pollution can interfere with daily activities, disrupt communication, and contribute to social isolation (Singh and Davar, 2004). As urbanization increases and population expansion takes place, the prevalence of noise pollution escalates, necessitating concerted efforts to mitigate its adverse effects and safeguard public health and tranquility (Bronzaft and Hagler, 2010). Effective management strategies encompassing regulation, urban planning, technological innovations, and public awareness campaigns are imperative to address the multifaceted challenges posed by noise pollution in contemporary societies. Recent studies have indicated that excessive noise exposure can lead to a myriad of sleep disturbances, hypertension, cardiovascular disease, stress, hypertension, stroke, diabetes, feelings of annoyance, and cognitive impairment (Shamsipour *et al.,* 2022; Hemmat *et al.,* 2023; Garg, 2022). Noise annoyance is demarcated as a sensitive sign of discomfort, resentment, dissatisfaction, and offense (Hegewald *et al*., 2020; Tiwari *et al.,* 2023, Stansfeld *et al*., 2021). A lot of investigations were carried out to evaluate the feeling of annoyance caused by acoustical and non-*acoustical* factors (Kim *et al*., 2012; Bravo-Moncavo *et al.,* 2019; Das *et al.,* 2022; Gilani *et al.,* 2021; Morihara *et al.,* 2022). Sound level frequency, exposure period, Noise source, and the recipient's psychological and internal processes play a major role in the assessment of noise annoyance (Ouis, 2001; Jakovljevic *et al.,* 2009; Fyhri and Klæboe, 2009). As urbanization accelerates, particularly in rapidly growing cities like Riyadh, Saudi Arabia, understanding and mitigating noise pollution becomes increasingly imperative. Riyadh, the capital and largest city of Saudi Arabia, experiences considerable urban development and population growth, leading to elevated levels of noise pollution originating from various sources such as traffic, construction, and industrial activities. In the context of developing countries, the utilization of advanced technologies such as machine learning presents a promising approach for predicting and managing traffic-induced noise annoyance levels. Most of the researchers have used the Multiple Linear Regression model and Artificial Neural Network model for the determination of noise levels (Ranpise *et al.,* 2021; Nourani *et al.,* 2020) but rarely any studies have reported the use of a Support Vector Machine (SVM) in the prediction of noise annoyance. Machine learning models, leveraging complex datasets and complex algorithms, have demonstrated remarkable capabilities in discerning patterns and making predictions in diverse applications (Jordan and Mitchell, 2015). This study presents a machine learning-based framework known as an SVM tailored for the prediction of noise annoyance in Riyadh city, Saudi Arabia. The applicability of the SVM model is widely used in a lot of domain areas such as bankruptcy, bearing life, air pollution, and noise levels (Min and Lee, 2005; Yan *et al.,* 2020; Leong *et al.,* 2020; Tiwari *et al.,* 2022). The SVM is robust against overfitting, especially for high-dimensional data, and can effectively handle non-linear relationships using various kernel functions. SVMs are less sensitive to outliers and aim for a global optimization solution, aided by their regularization parameter for fine-tuning (Yan *et al.,* 2020; Leong *et al.,* 2020; Tiwari *et al.,* 2022). Applying the SVM model to forecast noise annoyance levels can offer valuable insights for city planners, policymakers, and residents alike, facilitating targeted interventions and strategies to mitigate and control the adverse effects of noise pollution on respondents. By integrating acoustical

and non-acoustical factors, with residents' subjective assessments of noise annoyance, the SVM model aims to accurately predict traffic-induced noise annoyance levels across different urban areas of Riyadh city.

Key objectives of this study include:

- Developing a robust machine learning algorithm (Support Vector Machine) for predicting noise annoyance levels in Riyadh city.
- Integrating diverse acoustical and non-acoustical datasets encompassing demographic characteristics, traffic conditions, and acoustical factors to enhance prediction accuracy.
- Assessing the performance of the proposed SVM model through a case study focusing on Riyadh, Saudi Arabia.
- Providing insights and recommendations for city planners and policymakers to control and mitigate noise pollution in Saudi Arabia.

Despite the existing body of research on noise pollution and its health impacts, there remains a gap in the utilization of advanced machine learning techniques like SVM for predicting noise annoyance specifically in rapidly urbanizing contexts such as Riyadh. By addressing these objectives, the present study endeavors to offer a novel approach to managing noise pollution in urban environments, in Riyadh city. Through the amalgamation of machine learning algorithms and traffic noise data, this research aims to foster sustainable urban development practices conducive to the well-being and better satisfaction of life.

MATERIALS AND METHODS

Based on the literature survey, a proposed methodology has been introduced in the study which is presented in Fig. 1. This proposed methodology includes various sections such as study area, socio-acoustics survey, data-preprocessing, support vector machine, and performance matrix.

Study area

Riyadh city, also known as the capital of Saudi Arabia, is characterized by a dynamic urban landscape and an increasing population in the last few years.

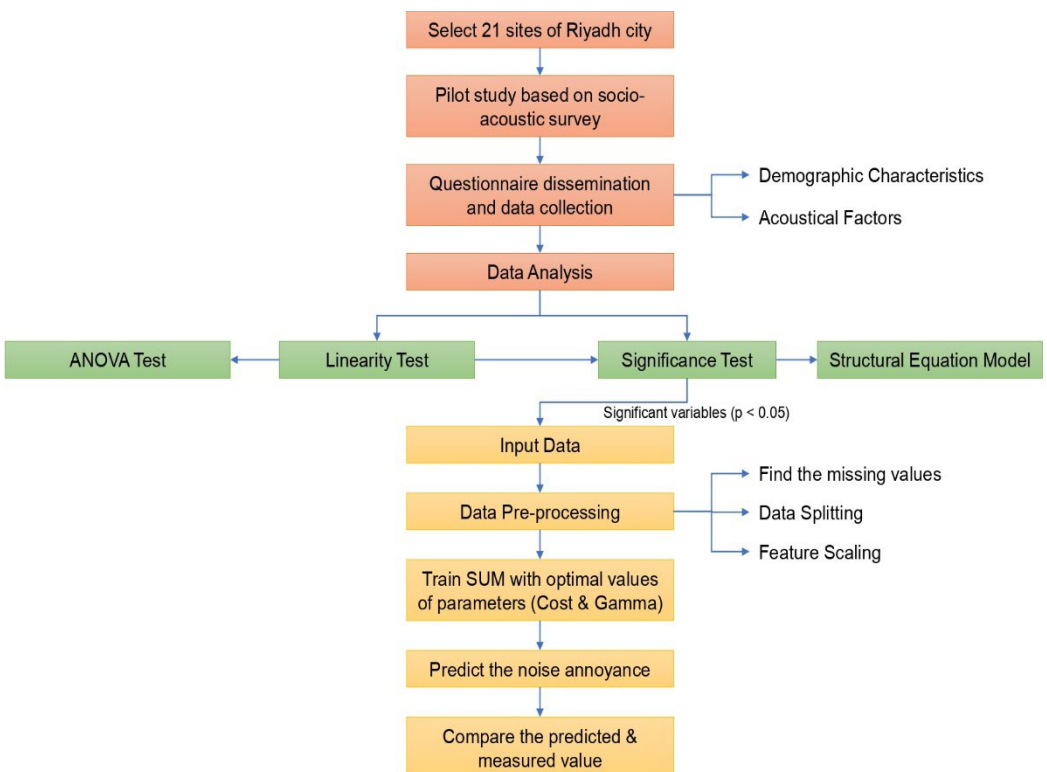

Fig. 1: Research methodology for prediction of noise annoyance

Riyadh city serves as a major economic and cultural hub throughout the Arab region. The traffic scenario in Riyadh is often influenced by the city's expansive road network system, industrial and economic development, increasing population density, and ongoing urban projects (Alghamdi *et al.,* 2021). The city's planning incorporates various locations with diverse metrological characteristics and traffic conditions, mainly from residential and commercial areas. The road characteristics of the city play a crucial role in accommodating the rising number of vehicles, contributing to traffic patterns that vary across different areas. Riyadh's traffic conditions are subject to factors such as population density, land-use planning, and the distribution of residential and commercial zones. The city is situated at a latitude of 24.774265 and a longitude of 46.738586. The study utilized 21 commercial and residential sites in the city has been considered for the study. Table 1 shows the characteristics of various locations such as zoning, latitude, and longitude.

Socio-acoustics Survey

A socio-acoustics survey was conducted in Riyadh city, involving 960 respondents through a random sampling approach at 21 different commercial and residential areas. Householders, shopkeepers, consumers, and working professionals were considered for the survey. Before the main survey, a pilot study was conducted with 38 participants to validate the reliability of the questionnaire. The survey teams visited selected respondents, explaining the study's purpose and benefits in English and Arabic to ensure understanding. Participation was voluntary, with 928 response sheets obtained. The questionnaire comprised two sections: demographic characteristics and acoustical factors. The demographic characteristics included gender, age, marital status, profession, and education, respectively. Gender and marital status were measured on a binary scale (0 and 1) (Gilani *et al.,* 2021; Das *et al.,* 2021; Nazneen *et al.,* 2020), whereas age and education were measured on a scale of 5 (Das *et al.,* 2021; Tiwari *et al.,* 2023). In the acoustical factors section, traffic flow was rated on a Likert scale of 3, while other factors such as transportation noise, noise sensitivity, and perceived noisiness were rated on a Likert scale of 1 to 5 (Tiwari *et al.,* 2023; Nazneen *et al.,* 2020; Fyhri and Klæboe, 2009).

Data pre-processing

Data pre-processing acts as a critical phase in the predictive modeling of noise annoyance. It consists of three steps i.e., data categorization, data splitting, and feature scaling (Zelaya 2019). The first step in data pre-processing is data categorization

Table 1: Characteristics of various locations of Riyadh city

Locations	Characteristics	Longitude	Latitude
Olaya District	Commercial	46°40'42.27"E	24°41'45.98"N
King Fahd Road	Mix	46°36'53.84"E	24°49'29.05"N
Tahlia Street	Commercial	46°41'26.08"E	24°41'57.65"N
Al-Ma'ather Street	Commercial	46°39'25.80"E	24°43'5.90"N
King Abdulaziz Road	Mix	46°38'29.62"E	24°49'22.73"N
Prince Turki bin Abdulaziz Road	Residential	46°36'44.71"E	24°45'9.56"N
Imam Saud bin Abdulaziz Road	Mix	46°41'24.24"E	24°45'32.45"N
Northern Ring Road	Mix	46°48'16.00"E	24°49'39.71"N
Eastern Ring Road	Mix	46°45'58.36"E	24°42'28.98"N
Southern Ring Road	Mix	46°42'36.70"E	24°35'18.31"N
Makkah Road	Mix	46°32'37.40"E	24°35'26.20"N
Salahuddin Ayyubi Road	Residential	46°45'1.99"E	24°40'15.64"N
King Abdullah Road	Mix	46°43'14.00"E	24°45'5.37"N
Makkah Mukarramah Street	Commercial	46°32'37.13"E	24°35'21.02"N
Thumamah Road	Residential	46°44'3.86"E	24°50'1.51"N
Abu Bakr Al Siddiq Road	Mix	47°18'59.16"E	24° 9'15.49"N
Al-Uraija Al-Gharbiyah Street	Residential	46°36'0.48"E	24°35'23.60"N
Khurais Road	Mix	46°53'10.11"E	24°49'30.72"N
Abu Bakr Al Siddiq Road	Commercial	47°19'0.25"E	24° 9'12.63"N
Al-Qassim Road	Mix	46°45'15.85"E	24°50'30.16"N
Al-Ahsa Street	Commercial	46°43'13.52"E	24°44'9.80"N

of variables which encode the categorical variables (Gender) into numerical format to facilitate their inclusion in the proposed model. This transformation ensures that qualitative information is appropriately integrated into the analysis without introducing bias (Leong *et al.,* 2020). Subsequently, the traffic-induced acoustical and non-acoustical dataset was partitioned into training and testing datasets, adhering to the conventional 80:20 split ratio (Bravo-Moncayo *et al.,* 2019). The data-splitting strategy ensures that the proposed model is trained on a sufficiently large subset of the data while retaining a separate portion for the model evaluation. It is used to prevent variables with larger scales from disproportionately influencing the model training process. The third step in data pre-processing is feature scaling. Min-max scaling was utilized in the present study to transform the features to a comparable scale. Mathematically, Min-Max scaling is represented as (Ambarwari *et al.,* 2020):

$$N_{scaled} = \frac{(N - N_{min})}{(N_{max} - N_{min})} \quad (1)$$

where N_{min} and N_{max} denote the minimum and maximum values of the feature, respectively.

Proposed model-SVM

Support Vector Machine (SVM) is a powerful supervised learning algorithm widely utilized for regression problems (Trafalis and Ince, 2000). For noise annoyance prediction, SVM will prove to be a robust tool due to its ability to handle high-dimensional data and nonlinear relationships. The fundamental principle of SVM is to find the hyperplane that best separates the data points into different classes or, in regression tasks, predicts the continuous target variable with maximum margin (Leong *et al.,* 2020). Mathematically, the hyperplane in SVM is represented as follows (Trafalis and Ince, 2000):

$$w \cdot n + t = 0 \quad (2)$$

where w is the weight vector, n is the input vector, and t is the bias term.

For regression problems, the objective is to find the hyperplane that minimizes the error between the predicted and actual noise annoyance values, which is formulated as follows (Xue *et al.,* 2009):

$$min_{w,b} \frac{1}{2} \| \mathbf{w} \|^2 + c \sum_{i=1}^{n} (a_i - (w.n_i + b))^2 \quad (3)$$

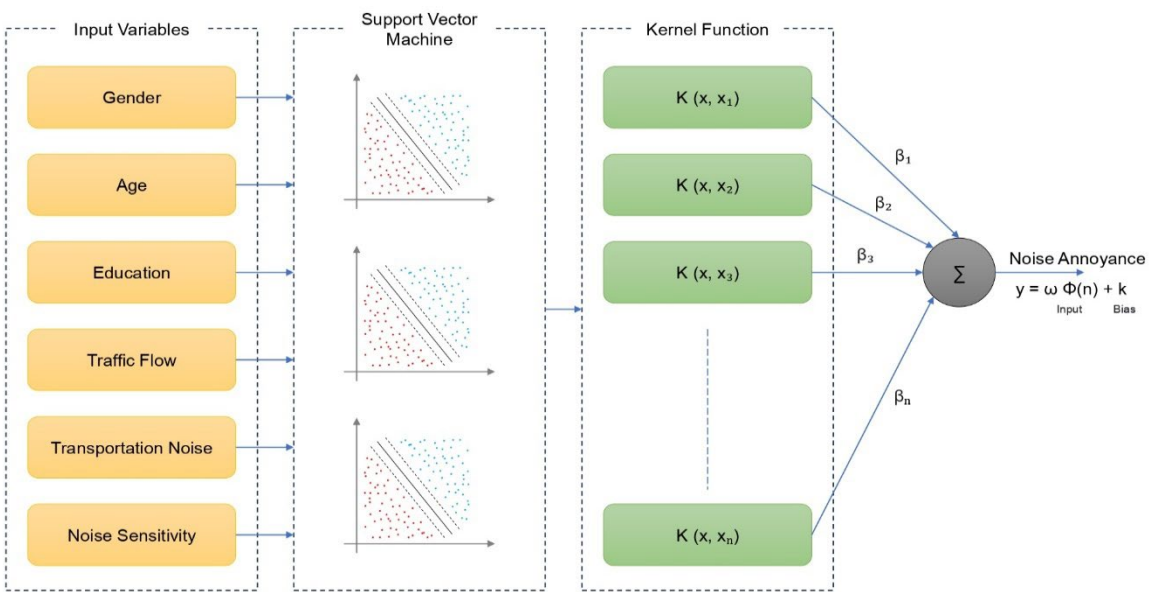

Fig. 2: Architecture of SVM model

Here, C is the regularization parameter that balances the trade-off between maximizing the margin and minimizing the error. SVM employs the kernel function to implicitly map the input data into a higher-dimensional space where the data becomes linearly separable. The most commonly used kernels include linear, polynomial, and radial basis function (RBF) kernels (Tiwari *et al.,* 2022). By optimizing the parameters and selecting the appropriate kernel function, SVM can effectively capture complex patterns in the data and produce accurate predictions for noise annoyance levels. Fig. 2 shows the architecture of the SVM model. It consists of input data, data pre-processing, kernel function, and out variable. In the present study, a radial basis function kernel was employed in SVM due to its flexibility in capturing nonlinear relationships, resulting in a robust predictive model for noise annoyance regression problems (Xue *et al.,* 2009).

Performance matrix

The performance of the SVM model in predicting noise annoyance is assessed using three key metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2).

RMSE quantifies the average magnitude of the errors between predicted and actual noise annoyance values, providing insight into the model's overall accuracy. MAE measures the average absolute difference between predicted and actual noise annoyance values, offering a robust evaluation of prediction capability. Additionally, R^2 indicates the proportion of variance in the target variable illustrated by the model, with higher values signifying a better fit. These metrics collectively offer a comprehensive evaluation of the SVM model's performance in noise annoyance prediction. The mathematical formulation for the following performance matrix is given as follows (Tiwari *et al.,* 2022; Patel *et al*., 2024):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(a_i - \hat{a}_i\right)^2} \quad (4)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|a_i - \hat{a}_i\right| \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(a_i - \hat{a}_i\right)^2}{\sum_{i=1}^{n}\left(a_i - \bar{a}_i\right)^2} \quad (6)$$

a_i, $\hat{a}_i$, and $\bar{a}_i$ represents the actual value, predicted value, and mean of the actual values of the target variable (noise annoyance). n represents the number of samples in the dataset.

RESULTS AND DISCUSSION

Demographic characteristics

Table 2 represents the demographic characteristics of key variables such as gender, age, education, marital status, and profession among the population. Gender distribution indicates a higher presence of males, comprising 89% of the respondents, while females represent 11%. The age variable highlights a range between 18 to above 60, with the majority of respondents falling within the 31-40 age group. The educational variable reveals a significant proportion of individuals with graduation-level qualifications, comprising 48.9% of the respondents, followed by those with a high school (16.5%), master's (13.5%), matric (12.7%), and higher education (8.4%). Marital status indicates a majority of married respondents (86.4%), with a smaller proportion being unmarried (13.6%). Profession distribution shows that a substantial portion of the respondents were self-employed (50.7%), followed by employed respondents (38.7%), and unemployed respondents (10.6%).

Socio-acoustics Survey

Table 3 illustrates the acoustical characteristics of the survey and explores respondent's perception of noise across five key variables: Traffic Flow, Transportation Noise, Noise Sensitivity, Perceived Noisiness, and Noise Annoyance. The respondents reported high traffic flow (45.1%) in the majority of the residential and commercial areas. Transportation noise was found to be a highly prominent factor for traffic-induced noise by 39.2%. Noise sensitivity is structured around the Weinstein Noise Sensitivity Scale (WNSS) (Kishikawa *et al.,* 2006). It consists of 21 questions on a Likert scale of 1 to 5. A substantial proportion reported high (32.9%) or extreme (22.3%) sensitivity to noise. The noisiness perceived by the respondents was reported as high by 31.6% and extremely high by 15.5%. The noise annoyance due to traffic was mostly considered to be moderate (31.7%) and extremely high by 14.3%.

Structural model

Table 4 indicates that the variables show a non-linear relationship, with p-values below 0.05. This analysis confirms that the dataset exhibits non-linear characteristics, making it suitable for input into the SVM model (Memon *et al.*, 2021). The SVM model can extract intricate non-linear features and utilize them for predicting the targeted noise annoyance variable. Table 5 presents the outcomes from the structural path model, wherein path coefficients gauge the strength and direction of relationships between the constructs. To establish a significant relation between constructs, a path coefficient must exceed 0.2 at a 95% confidence interval (Das *et al.*,

Table 2: Demographic characteristics of the respondents

S.No.	Variables	Level of Measurement	Total Count	Percentage of Count (in %)
1	Gender	Male	826	89.0
		Female	102	11.0
2	Age (in years)	18-30	228	24.5
		31-40	325	35.0
		41-50	181	19.5
		51-60	130	14.0
		>60	64	1.0
3	Education	Matric	118	12.7
		High School	135	16.5
		Graduation	454	48.9
		Masters	144	13.5
		Higher	77	8.4
4	Marital Status	Married	802	86.4
		Unmarried	126	13.6
5	Profession	Unemployed	98	10.6
		Employed	359	38.7
		Self-Employed	471	50.7

Table 3: Acoustical response of the respondents

S.No.	Variables	Level of Measurement	Total Count	Percentage of Count (in %)
1	Traffic Flow	Low	127	13.7
		Medium	382	41.2
		High	419	45.1
2	Transportation Noise	Not at all	21	2.3
		Low	88	9.5
		Medium	256	27.6
		High	364	39.2
		Extreme	199	21.4
3	Noise Sensitivity	Not at all	45	4.8
		Low	127	13.7
		Medium	244	26.3
		High	305	32.9
		Extreme	207	22.3
4	Perceived Noisiness	Not at all	81	8.8
		Low	135	14.5
		Medium	293	29.6
		High	275	31.6
		Extreme	144	15.5
5	Noise Annoyance	Not at all	48	5.2
		Low	182	19.6
		Medium	271	31.7
		High	294	29.2
		Extreme	133	14.3

Table 4: Linearity between the variables

Variables	Df	Sum of Square	Non-Linearity
Gender - Annoyance	1	1.228	Yes
Age - Annoyance	4	2.651	Yes
Education - Annoyance	4	2.905	Yes
Marital Status - Annoyance	1	1.641	Yes
Profession - Annoyance	2	2.544	Yes
Traffic Flow - Annoyance	2	2.218	Yes
Transportation Noise - Annoyance	4	3.936	Yes
Noise Sensitivity - Annoyance	4	2.814	Yes
Perceived Noisiness - Annoyance	4	2.472	Yes

Table 5: Structural model results

Variables	Mean	St.Dev	T-statistics	p-values	Supported
Gender - Annoyance	0.483	0.038	12.716	***	**Yes**
Age - Annoyance	0.292	0.046	6.358	***	**Yes**
Education - Annoyance	0.198	0.051	3.892	**	**Yes**
Marital Status - Annoyance	0.077	0.045	1.714	0.076	No
Profession - Annoyance	0.083	0.072	1.164	0.522	No
Traffic Flow - Annoyance	0.245	0.055	4.465	**	**Yes**
Transportation Noise - Annoyance	0.674	0.082	8.217	***	**Yes**
Noise Sensitivity - Annoyance	0.182	0.037	4.936	***	**Yes**
Perceived Noisiness - Annoyance	0.221	0.042	0.528	0.889	No

Table 6: SVM parametric values

S.No.	Variables	Range	Optimum value
1	Cost	$2^{(-5,5)}$	32
2	Gamma	$2^{(-5,5)}$	32

2022). Table 5 illustrates that all path coefficients were highly significant at a level of 0.001. Evaluation parameters including mean, standard deviation, t-statistics, and p-value were utilized to assess the relationship significance between the constructs. Notably, Gender (t: 12.716), Age (t: 6.358), Education (t: 3.892), Traffic Flow (t: 4.465), Transportation Noise (t: 8.217), and Noise Sensitivity (t: 936) significantly influenced Noise Annoyance. Moreover, Marital Status, Profession, and Perceived Noisiness do not affect Noise Annoyance.

Proposed model

The current study aims to assess the effectiveness of a Support Vector Machine (SVM) model using a multivariate dataset, as outlined in the "Materials and Methods" section, for predicting and conducting parametric sensitivity analysis of traffic-induced noise annoyance across Riyadh city. A 10-fold cross-validation was employed for each approach with the training dataset, wherein the training set was segmented into 10 subsets for iterative analysis. This cross-validation method was chosen to mitigate the risk of overfitting and facilitate easy parallelization due to the independence of hyper-parameters (Bergmeir *et al.*, 2012). Two hyper-parameters cost and gamma have been taken to architect the SVM model. Table 6 presents parameter values for an SVM model. It includes Cost and Gamma parameters, with ranges from 2^{-5} to 2^{5}. The optimum values for both parameters are 32, indicating their effectiveness in the model. These parameters play a crucial role in tuning SVM models for optimal performance. Table 7 illustrates the training and testing prediction

Table 7: Performance matrix of the proposed model

Matrix	Training data (80%)	Testing data (20%)
RMSE	1.162	1.416
MAE	1.307	1.538
R^2	0.93	0.90

Table 8: Sensitivity analysis

S.No.	I/P Variables	Percentage (%)
1	Gender	14.7
2	Age	16.2
3	Education	5.8
4	Traffic Flow	12.9
5	Transportation Noise	21.6
6	Noise Sensitivity	28.8

performance of the proposed SVM model for noise annoyance prediction, obtaining RMSE values of [1.162, 1.416] for the training and testing datasets, respectively. Notably, compared to other machine techniques cited in previous literature (Khorshid *et al.,* 2015, Chand *et al.,* 2016, Tiwari *et al.,* 2022), the Support Vector Machine model demonstrates superior predictive capabilities, because SVM model is less sensitive to outliers and aim for a global optimization solution, aided by their regularization parameter for fine-tuning within the survey dataset acquired from Riyadh city.

Additionally, sensitivity analysis as shown in Table 8 highlights the significance of input variables, such as gender, age, education, traffic flow, transportation noise, and noise sensitivity for their relevance in predicting noise annoyance levels.

Discussions related with hypothesis

Saudi Arabia, as a developing nation, has witnessed a substantial rise in the volume of vehicle density on its roads. However, this rapid expansion of road networks and ongoing infrastructure development, has become a challenge for countries like Saudi Arabia grappling with resource constraints. The sudden boost in vehicular traffic, particularly in industrial or metropolitan cities of Saudi Arabia, has led to a decline in overall quality of life, primarily attributable to heightened levels of traffic-generated noise. Previous research by Tiwari *et al.* (2022) has developed a direct correlation between traffic-induced noise annoyance levels and increased road traffic, as opposed to air or rail transportation.

A few observations have been acquired from the sensitivity analysis:

- Noise sensitivity is observed to have a sensitivity of 28.8% for noise annoyance. The finding aligns with similar observations by some researchers (Tiwari *et al.,* 2023; Bravo-Moncayo *et al.,* 2019; Park *et al.,* 2017) highlighting noise sensitivity as a primary predictor of noise annoyance. Sensitive respondents tend to experience an adverse effect from traffic-induced noise pollution and more are attuned to environmental stressors (Garg *et al.,* 2022).
- Transportation noise is observed as the second most influencing predictor with a sensitivity of 21.6% of noise annoyance. Nevertheless, researchers such as Sung *et al.* (2017) and Cerletti *et al.,* (2022) have underscored increased transportation noise as a significant contributor to increased noise annoyance. Increased transportation noise is generated due to an increase in traffic flow scenario (Tiwari *et al.,* 2023; Méline *et al.*, 2013; Eze *et al.*, 2018). This is also shown in the sensitivity analysis that traffic flow has a significant effect on noise annoyance (Das *et al.,* 2021; Nazneen *et al.,* 2020).
- Gender, age, and education also emerged as an important variable, with sensitivities of 14.7%, 16.2%, and 5.8% respectively. Multiple studies (Das *et al.,* 2022; Gilani *et al.,* 2021; Tiwari *et al.,* 2023) have attributed to the fact that Gender, age, and education affect noise annoyance.

Limitation and scope for future study

Based on the present study, certain limitation was depicted:

1. The applicability of the SVM model may be limited to Riyadh city, Saudi Arabia, and may not be generalized to other locations of the country due to variations in environmental aspects, metrological factors, urban infrastructure, traffic scenarios, and cultural contexts.
2. The accuracy and reliability of the noise annoyance prediction model heavily rely on the observational labeled data. Limited or inaccurate data on traffic scenarios, demographic characteristics, and other acoustical variables may affect the model's performance.
3. The machine learning model's (Support Vector Machine) complexity may have a deficiency in terms of socio-acoustics data interpretability, scalability, and computational resources required for training the model.
4. Temporal changes, including fluctuations in population density, urban growth, and transportation trends, could impact the model's predictive performance. It's crucial to continuously monitor and adjust the model to accommodate these temporal variations.

Based on the limitation certain future work can be done:

1. Future work will explore the applicability of the deep learning model to predict noise annoyance in other cities within Saudi Arabia considering variations in geographical, cultural, and socio-economic factors.
2. Incorporating real-time data from sensors and IoT devices could enhance the model's accuracy and enable timely interventions to mitigate noise annoyance.
3. Additional variables such as land use patterns, green spaces, and community characteristics to capture a more comprehensive understanding of noise annoyance determinants.
4. Improving the interpretability of the machine learning model could enhance its utility for stakeholders and policymakers.

Policy, practical, and theoretical implication of the study

The findings of this study have significant implications for urban planning and policy-making in Riyadh, Saudi Arabia, and other rapidly urbanizing cities. The identification of key factors influencing noise annoyance, such as traffic flow, noise sensitivity, and demographic characteristics, highlights the need for targeted noise mitigation strategies (Bröer *et al.*, 2010). Policymakers can leverage these insights to develop and enforce stricter noise regulations, particularly in areas with high traffic volumes and sensitive populations. The implementation of noise barriers, optimized traffic management systems, and designated quiet zones could substantially reduce the impact of traffic-induced noise. Additionally, the study underscores the importance of integrating noise pollution considerations into broader urban development plans to foster sustainable and liveable urban environments. For urban planners and city officials, this study offers practical tools and strategies for mitigating traffic-induced noise annoyance. The developed SVM model, with its high predictive accuracy, can be employed to identify noise hotspots and predict future noise annoyance levels under various traffic scenarios. This predictive capability allows for proactive measures, such as rerouting traffic, adjusting road usage patterns, and planning construction activities during less sensitive times. Moreover, public health officials can utilize these findings to design and implement community awareness programs that educate residents about noise pollution and promote behaviors to minimize noise exposure. The integration of machine learning models into urban planning processes can significantly enhance the effectiveness of noise mitigation efforts. Theoretically, this study contributes to the body of knowledge on environmental noise pollution and its impact on human health and well-being. By employing a two-staged SEM-SVM approach, the research demonstrates the value of combining structural equation modeling with machine learning techniques to analyze complex environmental issues. This methodological innovation provides a robust framework for future studies investigating the multifaceted nature of noise annoyance and other environmental stressors. The study also highlights the importance of considering both acoustical and non-acoustical factors in noise annoyance research, paving the way for more comprehensive and nuanced models in the field of socio-acoustics. Future research can build on this foundation by exploring additional variables and contexts, further

refining the predictive capabilities of machine learning models in environmental health studies.

CONCLUSIONS

The research presents a comprehensive investigation into the prediction of noise annoyance in Riyadh city, Saudi Arabia, utilizing a machine learning framework. The SEM analysis revealed that gender, age, education, traffic flow, transportation noise, and noise sensitivity significantly contribute to noise annoyance. The SVM model, with a RMSE of 1.416 and an R-squared (R^2) value of 0.90, demonstrates strong predictive performance. This underscores its utility in informing urban planning and noise pollution management strategies, highlighting the potential for machine learning models to accurately predict and mitigate environmental stressors. The sensitivity analysis further elucidated the factors contributing to noise annoyance, revealing that noise sensitivity (28.8%) is the most critical factor, followed by transportation noise, age, gender, traffic flow, and education. These findings indicate that individual differences in noise sensitivity must be a central consideration in noise mitigation strategies. Moreover, the significant influence of transportation noise on annoyance levels underscores the necessity for targeted interventions in traffic management and urban design to alleviate noise pollution. The results indicate that older individuals and those with higher education levels tend to report higher noise annoyance, suggesting that demographic factors play a crucial role in how noise is perceived and experienced. Additionally, gender differences in noise annoyance emphasize the need for tailored approaches in addressing noise issues, considering the varying sensitivities among different population segments. The strong performance of the SVM model, with its high predictive accuracy, signifies the model's robustness and reliability in real-world applications. This model can serve as a valuable tool for urban planners and policymakers to identify noise hotspots, predict future annoyance levels, and develop effective noise mitigation measures. By integrating machine learning models into urban planning, cities can proactively manage noise pollution and improve the quality of life for their residents.

AUTHOR CONTRIBUTIONS

M.U. Rehman and M. Kamal compiled the collected data and reviewed the manuscript preparation. S.K. Tiwari performed the literature review, analyzed and interpreted the data, prepared the manuscript.

ACKNOWLEDGEMENT

The authors thank to the Deanship of Scientific Research at Saudi Electronic University for funding this research work through the project number (8290).

CONFLICT OF INTEREST

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication, falsification, double publication, submission, and redundancy, have been entirely witnessed by the authors.

OPEN ACCESS

©2025 The author(s). This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution, and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third-party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit: http://creativecommons.org/licenses/by/4.0/

PUBLISHER'S NOTE

Tehran Urban Research & Planning Centre (TURPC) takes a neutral position about claims on disputed territories, place names, international boundaries, jurisdiction in published maps, and institutional affiliations. TURPC is committed to retracting a paper after its publication if it becomes

apparent that there are serious problems in its content, in terms of research and publication ethics (https://www.ijhcum.net/journal/process?ethics).

ABBREVIATIONS (NOMENCLATURE)

C	Regularization Parameter
MAE	Mean Absolute Error
N	Input Vector
R^2	Coefficient of Determination
RBF	Radial Basis Function
RMSE	Root Mean Square Error
SEM	Structural Equation Model
SVM	Support Vector Machine
T	Bias Term
W	Weight
IoT	Internet of Things

REFERENCES

Agarwal, S.; Swami, B. L., (2011). Comprehensive approach for the development of traffic noise prediction model for Jaipur city. Environmental monitoring and assessment., 172: 113-120 **(8 pages)**.

Alghamdi, A. S.; Alzhrani, A. I.; Alanazi, H. H., (2021). Local climate zones and thermal characteristics in Riyadh city, Saudi Arabia. Remote Sensing., 13(22): 4526-4543 **(18 pages)**.

Al-Ghonamy, A. I., (2009). Assessment of traffic noise pollution in Al-Khobar, a typical city in the Kingdom of Saudi Arabia. Sci. J. King Faisal Univ. Basic Appl. Sci., 10: 165-180 **(16 pages)**.

Al-Ghonamy, A. I., (2010). Analysis and evaluation of road traffic noise in Al-Dammam: A business city of the eastern province of KSA. Journal of Environmental Science and Technology., 3(1): 47-55 **(9 pages)**.

AlSaleem,S.S.; Almhafdy,A.; Berardi, U.; Al-Shargabi, A. A.; Ali, A. A. M., (2023). Field Measurements and Human Perception to Remediate Noise Pollution in the Urban Public Parks in Saudi Arabia. Sustainability., 15(13): 9977-9992 **(16 pages)**.

Ambarwari, A.; Adrian, Q. J.; Herdiyeni, Y., (2020). Analysis of the effect of data scaling on the performance of the machine learning algorithm for plant identification. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)., 4(1): 117-122 **(6 pages)**.

Bergmeir, C.; Benítez, J. M., (2012). On the use of cross-validation for time series predictor evaluation. Information Sciences., 191: 192-213 **(22 pages)**.

Bravo-Moncayo, L.; Lucio-Naranjo, J.; Chávez, M.; Pavón-García, I.; Garzón, C., (2019). A machine learning approach for traffic-noise annoyance assessment. Applied Acoustics., 156: 262-270 **(9 pages)**.

Bröer, C., (2008). Private trouble, policy issue people's noise annoyance and policy discourse. Critical Policy Analysis., 2(2): 93-117 **(25 pages)**.

Bronzaft, A. L.; Hagler, L., (2010). Noise: The invisible pollutant that cannot be ignored. Emerging Environmental Technologies., Volume II: 75-96 **(22 pages)**.

Cerletti, P.; Eze, I. C.; Schaffner, E.; Foraster, M.; Viennau, D.; Cajochen, C.; ... Probst-Hensch, N., (2020). The independent association of source-specific transportation noise exposure, noise annoyance and noise sensitivity with health-related quality of life. Environment international., 143: 105960-105969 **(10 pages)**.

Chand, N.; Mishra, P.; Krishna, C. R.; Pilli, E. S.; Govil, M. C., (2016, April). A comparative analysis of SVM and its stacking with other classification algorithm for intrusion detection. In 2016 International Conference on Advances in Computing, Communication, & Automation (ICACCA)(Spring) (pp. 1-6) **(6 pages)**.

Das, C. P.; Goswami, S.; Das, M., (2022). Prediction of traffic noise induced annoyance of vendors through noise indices using structural equation modeling: Artificial neural network model. Environmental Quality Management., 32(2): 29-44 **(16 pages)**.

Elshorbagy, K. A., (1999). Noise and community response in Jeddah and Riyadh cities. J. King Abdulaziz Univ., 10: 87-99 **(13 pages)**.

Eze, I. C.; Foraster, M.; Schaffner, E.; Vienneau, D.; Héritier, H.; Pieren, R.; ... Probst-Hensch, N., (2018). Transportation noise exposure, noise annoyance and respiratory health in adults: A repeated-measures study. Environment international., 121: 741-750 **(10 pages)**.

Fyhri, A.; Klæboe, R., (2009). Road traffic noise, sensitivity, annoyance and self-reported health, A structural equation model exercise. Environment International., 35(1): 91-97 **(7 pages)**.

Garg, N., (2022). Environmental Noise Environmental noise Control Noise control: A Review. Environmental Noise Control: The Indian Perspective in an International Context., 1-25 **(25 pages)**.

Gilani, T. A.; Mir, M. S., (2021). A study on the assessment of traffic noise induced annoyance and awareness levels about the potential health effects among residents living around a noise-sensitive area. Environmental Science and Pollution Research., 28(44): 63045-63064 **(20 pages)**.

Gupta, A.; Gupta, A.; Jain, K.; Gupta, S., (2018). Noise pollution and impact on children health. The Indian Journal of Pediatrics., 85(4): 300-306 **(7 pages)**.

Hegewald, J.; Schubert, M.; Freiberg, A.; Romero Starke, K.; Augustin, F.; Riedel-Heller, S. G.; ... Seidler, A., (2020). Traffic noise and mental health: a systematic review and meta-analysis. International journal of environmental research and public health., 17(17): 6175-6200 **(26 pages)**.

Hemmat, W.; Hesam, A. M.; Atifnigar, H., (2023). Exploring Noise Pollution, Causes, Effects, and Mitigation Strategies: A Review Paper. European Journal of Theoretical and Applied Sciences., 1(5): 995-1005 **(11 pages)**.

Jakovljevic, B.; Paunovic, K.; Belojevic, G., (2009). Road-traffic noise and factors influencing noise annoyance in an urban population. Environment international., 35(3): 552-556 **(5 pages)**.

Jordan, M. I.; Mitchell, T. M., (2015). Machine learning: Trends, perspectives, and prospects. Science., 349(6245): 255-260 **(6 pages)**.

Khorshid, M.; Abou-El-Enien, T. H.; Soliman, G., (2015). A

comparison among support vector machine and other machine learning classification algorithms. IPASJ International Journal of Computer Science (IIJCS)., 3(5): 25-35 **(11 pages)**.

Kim, M.; Chang, S. I.; Seong, J. C.; Holt, J. B.; Park, T. H.; Ko, J. H.; Croft, J. B., (2012). Road traffic noise: annoyance, sleep disturbance, and public health implications. American journal of preventive medicine., 43(4): 353-360 **(8 pages)**.

Kishikawa, H.; Matsui, T.; Uchiyama, I.; Miyakawa, M.; Hiramatsu, K.; Stansfeld, S. A., (2006). The development of Weinstein' s noise sensitivity scale. Noise and Health., 8(33): 154-160 **(7 pages)**.

Leong, W. C.; Kelani, R. O.; Ahmad, Z., (2020). Prediction of air pollution index (API) using support vector machine (SVM). Journal of Environmental Chemical Engineering., 8(3): 103208 **(10 pages)**.

Méline, J., Van Hulst, A.; Thomas, F.; Karusisi, N.; Chaix, B., (2013). Transportation noise and annoyance related to road traffic in the French RECORD study. International journal of health geographics., 12: 1-13 **(13 pages)**.

Min, J. H.; Lee, Y. C., (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Expert systems with applications., 28(4): 603-614 **(12 pages)**.

Morihara, T.; Sato, T.; Yano, T., (2022). Re-analysis of socio-acoustic survey data in Tomakomai: seasonal effects of a tree belt on road traffic noise annoyance. Applied Acoustics., 196, 108861-108874 **(14 pages)**.

Nazneen, S.; Raza, A.; Khan, S., (2020). Assessment of noise pollution and associated subjective health complaints and psychological symptoms: analysis through structure equation model. Environmental Science and Pollution Research., 27: 21570-21580 **(11 pages)**.

Nourani, V.; Gökçekuş, H.; Umar, I. K., (2020). Artificial intelligence based ensemble model for prediction of vehicular traffic noise. Environmental research., 180: 108852-108866 **(15 pages)**.

Ouis, D., (2001). Annoyance from road traffic noise: a review. Journal of environmental psychology., 21(1): 101-120-139 **(20 pages)**.

Park, J.; Chung, S.; Lee, J.; Sung, J. H.; Cho, S. W.; Sim, C. S., (2017). Noise sensitivity, rather than noise level, predicts the non-auditory effects of noise in community samples: a population-based survey. BMC public health., 17: 1-9 **(9 pages)**.

Patel, R.; Singh, P. K.; Saw, S., (2024). Traffic Noise Modeling under Mixed Traffic Condition in Mid-Sized Indian City: A Linear Regression and Neural Network-Based Approach. International Journal of Mathematical, Engineering and Management Sciences., 9(3): 411-434 **(24 pages)**.

Presidency of Meteorology & Environment (PME). Kingdom of Saudi Arabia National Environmental Standard: Environmental Noise.

Ranpise, R. B.; Tandel, B. N.; Darjee, C., (2021). Assessment and MLR modeling of traffic noise at major urban roads of residential and commercial areas of Surat City. In Sustainability in Environmental Engineering and Science: Select Proceedings of SEES 2019 (pp. 181-191). Springer Singapore **(11 pages)**.

Shamsipour, M.; Zaredar, N.; Monazzam, M. R.; Namvar, Z.; Mohammadpour, S., (2022). Burden of diseases attributed to traffic noise in the metropolis of Tehran in 2017. Environmental Pollution., 301: 119042.

Stansfeld, S.; Clark, C.; Smuk, M.; Gallacher, J.; & Babisch, W., (2021). Road traffic noise, noise sensitivity, noise annoyance, psychological and physical health and mortality. Environmental Health., 20: 1-15 **(15 pages)**.

Sung, J. H.; Lee, J.; Jeong, K. S.; Lee, S.; Lee, C.; Jo, M. W.; Sim, C. S., (2017). Influence of transportation noise and noise sensitivity on annoyance: a cross-sectional study in South Korea. *Int. J. Environ. Res. Public Health*., 14(3): 322-331 **(10 pages).**

Tiwari, S. K.; Kumaraswamidhas, L. A.; Garg, N., (2022). Comparison of SVM and ARIMA model in time-series forecasting of ambient noise levels. In Advances in Energy Technology: Select Proceedings of EMSME 2020 (pp. 777-786). Springer Singapore **(10 pages)**.

Tiwari, S. K.; Kumaraswamidhas, L. A.; Garg, N., (2022). Time-series prediction and forecasting of ambient noise levels using deep learning and machine learning techniques. Noise Control Engineering Journal., 70(5): 456-471 **(16 pages)**.

Tiwari, S. K.; Kumaraswamidhas, L. A.; Garg, N., (2023). Assessment of noise pollution and associated subjective health complaints in Jharia Coalfield, India: A structural equation model analysis. Noise Mapping., 10(1): 20220172 **(17 pages)**.

Tiwari, S. K.; Kumaraswamidhas, L. A.; Prince; Kamal, M.; Rehman, M. U., (2023). A hybrid deep leaning model for prediction and parametric sensitivity analysis of noise annoyance. Environmental Science and Pollution Research., 30(17): 49666-49684 **(19 pages)**.

Tiwari, S. K.; Kumaraswamidhas, L. A.; Patel, R.; Garg, N.; Vallisree, S., (2024). Traffic noise measurement, mapping, and modeling using soft computing techniques for mid-sized smart Indian city. Measurement: Sensors., 33: 101203-101214 **(12 Pages)**.

Trafalis, T. B.; Ince, H., (2000, July). Support vector machine for regression and applications to financial forecasting. In Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium (Vol. 6, pp. 348-353). IEEE **(6 pages)**.

Xue, H.; Yang, Q.; Chen, S., (2009). SVM: Support vector machines. The top ten algorithms in data mining., 6(3): 37-60 **(24 pages)**.

Yan, M.; Wang, X.; Wang, B.; Chang, M.; Muhammad, I., (2020). Bearing remaining useful life prediction using support vector machine and hybrid degradation tracking model. ISA transactions., 98: 471-482 **(12 pages)**.

Zelaya, C. V. G., (2019, April). Towards explaining the effects of data preprocessing on machine learning. In 2019 IEEE 35th international conference on data engineering (ICDE)., (pp. 2086-2090) **(5 pages)**.

Zengin, N.; Ören, B.; Üstündag, H., (2020). The relationship between stressors and intensive care unit experiences. Nursing in critical care., 25(2): 109-116 **(8 pages)**.

Zytoon, M. A., (2016). Opportunities for environmental noise mapping in Saudi Arabia: A case of traffic noise annoyance in an urban area in Jeddah city. *Int. J. Environ. Res. Public Health*., 13(5): 496-515 **(20 pages)**.

COPYRIGHTS

©2025 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

HOW TO CITE THIS ARTICLE

Rehman, M.U.; Kamal, M.; Tiwari, S.K., (2025). Statistical analysis based on a two-staged SEM-SVM approach for environmental noise annoyance prediction and identification of influencing factors. Int. J. Hum. Capital Urban Manage., 10(1): 121-134.

DOI: 10.22034/IJHCUM.2025.01.08

URL: https://www.ijhcum.net/article_714213.html