STATISTICS IN TRANSITION new series, December 2024
Vol. 25, No. 4 pp. 51–77, https://doi.org/10.59139/stattrans-2024-003
Received – 26.03.2024; accepted – 17.07.2024

Modelling Tinnitus Functional Index reduction using supervised machine learning algorithms

Edmund Fosu Agyemang[1]

Abstract

This study aims to model the reduction in the Tinnitus Functional Index (TFI) utilizing supervised machine learning algorithms, focusing primarily on Ordinary Least Squares (OLS), K-Nearest Neighbor (KNN), Ridge, and Lasso regressions. Our analysis highlighted Group, ISI, and SWLS as significant predictors of TFI reduction, identified through the best subset selection and confirmed by both forward and backward selection criteria in the OLS regression. Notably, the shrinkage methods, Ridge and Lasso regressions, demonstrated superior performance compared to OLS and KNN, with the Ridge regression presenting the smallest test mean square error (MSE) of 318.30. This finding establishes the Ridge regression as the best model for analyzing our Tinnitus dataset relative to the other methods, which exhibited test MSEs of 319.28 (Lasso), 330.76 (OLS), and 584.92 (KNN), respectively. This research highlights the potential of supervised machine learning algorithms in advancing personalized Tinnitus treatment, reflecting broader trends in the field as evidenced by studies in the literature. By leveraging these algorithms, we can enhance treatment precision and outcomes, contributing significantly to improved quality of life for individuals with Tinnitus. Future research should explore the integration of multimodal data and longitudinal applications of these algorithms to further refine predictive capabilities and treatment effectiveness.

Key words: Tinnitus, K-Nearest Neighbor regression, Ridge regression, Lasso regression, multiple linear regression.

1. Introduction

Tinnitus, characterized by the perception of noise or ringing in the ears in the absence of any external sound, is a prevalent and often distressing auditory phenomenon (De Ridder et al., 2021). Affecting a significant portion of the population,

[1] Corresponding author. School of Mathematical and Statistical Science, College of Sciences, University of Texas Rio Grande Valley, USA &Department of Statistics and Actuarial Science, College of Basic and Applied Sciences, University of Ghana, Ghana &Department of Computer Science, Ashesi University, No. 1 University Avenue, Berekuso-Ghana. E-mail: edmundfosu6@gmail.com, ORCID: https://orcid.org/0000-0001-8124-4493.

Tinnitus can vary in severity from a mild annoyance to a debilitating condition, impacting quality of life, emotional well-being, and cognitive function (Gasparre et al., 2023). Despite its common occurrence and prevalence, the etiology of Tinnitus remains complex and the precise mechanisms underlying Tinnitus remain poorly understood, and as such, treatment can be challenging, with potential contributions from auditory and non-auditory structures, suggesting a multidimensional pathology (Mohan et al., 2022). The Tinnitus Functional Index (TFI) has emerged as a crucial measure for evaluating the impact of Tinnitus on daily life, encompassing emotional distress, auditory difficulties, and interference with mental concentration among other factors. However, given the multifarious nature of Tinnitus, individual responses to treatment vary significantly, highlighting the need for personalized intervention strategies. Recent advances in supervised machine learning algorithms offer promising new avenues for understanding and managing complex health conditions like Tinnitus (Manta et al., 2023). These algorithms, capable of discerning patterns and associations in large datasets, present an innovative approach to predicting patient outcomes and optimizing treatment protocols. By integrating patient data, including demographic, audiometric, and psychological factors, supervised machine learning models can potentially identify the most influential predictors of TFI reduction, thereby aiding clinicians in tailoring interventions more effectively to individual needs. The integration of machine learning in the medical field is not without challenges. The accuracy of predictive models depends on the quality and quantity of the available data, and in the context of Tinnitus, this includes accurately captured TFI scores, detailed patient histories, and comprehensive treatment records.

Despite these challenges, the potential benefits of applying supervised machine learning to Tinnitus management are considerable. By enabling more precise and personalized treatment approaches, these technologies have the potential to significantly improve quality of life for Tinnitus sufferers. Furthermore, the insights gleaned from machine learning models could contribute to a deeper understanding of Tinnitus pathology and the development of more effective therapeutic interventions. The most common cause of Tinnitus is prolonged exposure to loud noise, which can damage the tiny sensory hair cells in the ear that transmit sound to the brain (Runciman and Johnson, 2023). However, Tinnitus can also result from other factors such as age-related hearing loss, earwax blockage, ear bone changes (otosclerosis), and conditions such as Meniere's disease, which affects the inner ear. Additionally, certain medications, including some antibiotics, cancer medications, and even high doses of aspirin, can induce Tinnitus as a side effect. Other potential causes include head or neck injuries (Biswas et al., 2023), which can affect the auditory nerves or brain function linked to hearing, and various cardiovascular diseases like hypertension, which can interfere with blood flow and cause Tinnitus. Machine learning is rapidly transforming the field of healthcare, offering new ways to enhance diagnostic accuracy and patient treatment outcomes (Agbota et al., 2024). Understanding the specific cause is crucial for effective treatment and management of the condition.

A systematic review by McCormack et al. (2016) highlighted that Tinnitus prevalence increases with age and exposure to noisy environments. The study also noted that about 10-15% of the adult population experiences Tinnitus, but only a small fraction, around 1-2%, finds it severely debilitating. The exact pathophysiological mechanisms of Tinnitus remain partially understood, but recent research suggests it involves both peripheral and central auditory pathways. Research by Eggermont and Roberts (2004) proposed that Tinnitus results from neural plasticity in response to auditory system damage, leading to altered neural activity perceived as sound. This theory is supported by findings of increased activity in the auditory cortex and changes in the brain's neural network connectivity. Tinnitus is not merely a sensory condition but also has profound psychological impacts. Studies have shown a strong correlation between Tinnitus severity and psychological distress, including anxiety, depression, and reduced quality of life. A study by Langguth et al. (2013) reviewed the impact of Tinnitus on mental health, underscoring the need for holistic approaches in treatment to address both auditory and psychological components. Smith et al. (2021) applied various supervised machine learning techniques to predict the outcomes of cognitive-behavioral therapy in individuals suffering from Tinnitus, as measured by changes in Tinnitus Functional Index (TFI) scores. They found that the random forest algorithm outperformed other models, such as decision trees and support vector machines, in predicting treatment success. The study emphasizes the value of incorporating diverse patient data to enhance the predictive accuracy of treatment outcomes, suggesting a move towards more personalized treatment approaches.

Doborjeh et al. (2023) proposed an Artificial Intelligence algorithm to predict patients' responses to Tinnitus therapies using EEG data. By employing deep learning techniques, the study achieved prediction accuracies ranging from 98% to 100%. The study identified the most informative EEG sensors and demonstrated how EEG frequency and functional connectivity could classify patients into therapy respondents and non-respondent groups, thereby suggesting a potential for real-time monitoring of therapy outcomes. Rodrigo et al. (2021) utilized decision tree models, specifically CART and gradient boosting, to predict the outcomes of Internet-Based Cognitive Behavioral Therapy (ICBT) for Tinnitus. The research highlighted that higher education levels were significantly influential for ICBT outcomes, and the CART decision tree model was able to identify participant groups with an 85% success probability following ICBT. In Cardon et al. (2022), machine learning algorithms, particularly random forest classifiers, were employed to predict outcomes of Tinnitus treatment modalities. The study noted that Tinnitus Functional Index (TFI) scores varied among participants, with some showing no improvement. The study contributes to understanding how machine learning can assist in predicting the response to specific

Tinnitus treatments, indicating a significant advancement toward personalized medicine in audiology. Prominent statistical tools such as cluster analysis (Van den Berge et al., 2017) and factor analysis (Wakabayashi et al., 2020) have been utilized to analyze Tinnitus in the literature.

However, this study aims to explore and identify the significant predictors of Tinnitus reduction, encompassing a range of possible factors including auditory exposure, demographic variables, psychological factors, and underlying health conditions. By employing a comprehensive and integrative approach, this study seeks to unravel the complex web of contributors to TFI reduction. By analyzing data from a diverse cohort of patients, we aim to identify key factors influencing treatment outcomes and to establish a predictive framework that can guide clinical decision-making. This research not only has the potential to enhance individual patient care but also contributes to the broader field of Tinnitus research by incorporating novel data-driven methodologies. The decision to employ four variants of supervised learning algorithms in the analysis of significant predictors of Tinnitus reduction represents a robust approach, leveraging the strengths of both parametric and non-parametric statistical techniques. This diverse approach enhances the ability to accurately model the predictors of Tinnitus reduction, catering for both linear and non-linear dynamics within the data.

2. Data and Methods

The dataset used in this study is of secondary nature and comprises pre and post Tinnitus Functional Index (TFI) scores, along with clinical and demographic information, from a pre-post intervention research involving 142 individuals affected by Tinnitus. The dependent variable for the study is TFI Reduction, which was computed as the difference between the TFI score at the beginning of the study and the TFI score after completion of the study. The predictors for the study include **HHI**: Hearing survey score; **GAD**: Generalized Anxiety Disorder, **PHQ**: Patient Health Questionnaire, **ISI**: Insomnia Severity Index, **SWLS**: Satisfaction with Life Scales, **HYP**: Hyperacusis, **CFQ**: Cognitive Failures, **Duration**: Duration of Tinnitus (in years), **PRE**: TFI score at the beginning of the study, **POST**: TFI score after the completion of the study. The TFI score at the end of the study had 17 missing observations. In this study, the MICE (Multiple Imputation by Chained Equations) algorithm was employed to address missing values in the dataset. Utilizing the Iterative Imputer from Scikit-learn with a Random Forest Regressor, the algorithm iteratively predicted missing numerical data over ten cycles, starting with mean values as place holders. This approach facilitated the estimation of missing data by exploiting inter-variable relationships, ensuring a more accurate and robust imputation compared to simpler methods. 10-fold cross-validation (i.e. setting $k = 10$) was used in the study. This process is repeated for

10 iterations. In each iteration, a different fold is kept for testing, and the remaining 9 folds are used for training. OLS, KNN, Ridge and Lasso regression were the main statistical tools adopted for the study. The data and python codes used for the study are freely available on Github and can be assessed at the repository at https://github.com/Agyemang1z/Tinnitus-Case-Study-1. The complete description of the Tinnitus dataset is given in Table 1.

Table 1: Description of the Tinnitus Dataset

Attribute	Description	Data Type
Group	Treatment and Control	Binary
Gender	1: Male 2: Female	Binary
HHI _Score	Hearing survey- Overall score- 0-40 (higher score more severe)	Numeric
Generalized Anxiety Disorder (GAD)	Anxiety sum: 0-21 (higher score more severe)	Numeric
Patient Health Questionnaire (PHQ)	Depression sum: 0-28 (higher score more severe)	Numeric
Insomnia Severity Index (ISI)	Insomnia total: 0-28 (higher score more severe)	Numeric
Satisfaction with Life Scales (SWLS)	Overall score, satisfaction with life, like Quality of Life (QOF) Higher scores better QOL (opposite to all other scales)	Numeric
Hyperacusis	0-42 (higher score more severe)	Numeric
Cognitive Failures (CFQ)	0-100 (higher score more severe)	Numeric
Age	In years	Numeric
Duration of tinnitus	In years	Numeric
Pre TFI Score	TFI score at the beginning of the study: Tinnitus score out of 100, higher more severe	Numeric
Post TFI Score	TFI score after the completion of the study: Tinnitus scores out of 100, higher more sever	Numeric

2.1. Ordinary Least Squares and Multiple Regression

Multiple regression allows us to determine the overall fit of the model and the relative contribution of each of the predictors to the total variance. Linearly, we consider the general model given in (1) as:

$$Y_i = \sum_{j=0}^{p} \beta_j X_{i,j} \; ; \quad i = 1,2,\dots,N \tag{1}$$

The least squares estimates, β_j in (1) that minimizes residual sums of squares (RSS) is given in (2) by:

$$RSS(\beta) = \sum_{i=1}^{N}\left[Y_i - \beta_0 - \sum_{j=1}^{p}\beta_j X_{ij}\right]^2 \tag{2}$$

However, it is possible that not all predictors are significantly associated with the outcome variable. Some predictors might not contribute meaningfully to the model, potentially leading to overfitting. To address this, variable selection techniques such as Best Subset Selection, Forward Selection, and Backward Elimination are employed. These methods aim to find the most relevant subset of predictors that offer the best prediction accuracy, balancing the model's complexity with its predictive power. Best Subset Selection was adopted for the study, where all possible combinations of predictors are evaluated, which can be computationally intensive for a large number of predictors. To evaluate the performance and select the optimal model among different candidate models, various statistical metrics and validation techniques are utilized. Commonly used criteria include the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and adjusted R^2, which help to compare models based on their explanatory power and complexity. Cross-validation techniques, such as k-fold cross-validation, are also widely used to estimate a model's prediction error on new data, ensuring that the selected model generalizes well beyond the observed dataset. Ultimately, the chosen MLR model should strike a balance between complexity (number of predictors) and prediction accuracy, avoiding overfitting while effectively capturing the relationships within the data.

2.2. Ridge and Lasso Regression

The Ordinary Least Squares (OLS) method is commonly used for estimation in regression models, offering adequate predictions under certain conditions. However, its effectiveness diminishes when dealing with nonorthogonal explanatory variables, leading to inaccurate weighting of these predictors (Saleh and Norouzirad, 2018). This issue is particularly pronounced in datasets not derived from controlled experiments, where the assumption of non-orthogonality among variables does not hold, and multicollinearity is present. OLS tends to produce unstable and overfitted models in such scenarios, emphasizing the need for strategies that mitigate model overfitting by reducing variance. To address these limitations, alternative methods like ridge and lasso regression have been developed. These techniques are designed to provide biased estimates of regression coefficients, particularly beneficial in cases of correlated predictors or multicollinearity (Omer, 2022). By introducing a bias into the regression results, they effectively lower the variances of the estimates, which enhances the interpretability of the regression coefficients and the overall model reliability. These strategies aim to utilize all p predictors while reducing their coefficients towards zero, effectively creating a subset when some coefficients become zero. It is important to note that the intercept is exempt from this reduction process. Using the Lagrangian equation in (3), the estimation of $\hat{\beta}$ is made possible by:

$$\hat{\beta}_{\text{shrink}} = \arg\min_{\beta}\left[\sum_{i=1}^{N}\left(Y_i - \beta_0 - \sum_{k=1}^{p}\beta_j X_{i,j}\right)^2 + \lambda\sum_{k=1}^{p}\psi(\beta_j)\right] \tag{3}$$

Here, $\lambda\sum_{k=1}^{p}\psi(\beta_j)$ is the shrinkage penalty and $\lambda \geq 0$ is the regularization parameter.

To estimate the ridge (also known as the l_2 regularization), we use the shrinkage penalty $\lambda\sum_{k=1}^{p}\psi(\beta_j)$ so that (3) is modified as in (4) as:

$$\hat{\beta}_{\text{shrink}}^{\boldsymbol{Ridge}} = \arg\min_{\beta}\left[\sum_{i=1}^{N}\left(Y_i - \beta_0 - \sum_{k=1}^{p}\beta_j X_{i,j}\right)^2 + \lambda\sum_{k=1}^{p}\beta_j^2\right] \tag{4}$$

In similar fashion, to estimate the lasso (also known as the l_1 regularization), we use the shrinkage penalty $\psi(\beta_j) = |\beta_j|$ so that (3) is modified as in (5) as:

$$\hat{\beta}_{\text{shrink}}^{\boldsymbol{Lasso}} = \arg\min_{\beta}\left[\sum_{i=1}^{N}\left(Y_i - \beta_0 - \sum_{k=1}^{p}\beta_j X_{i,j}\right)^2 + \lambda\sum_{k=1}^{p}|\beta_j|\right] \tag{5}$$

Y_i denotes the ith observation of the response variable (TFI Reduction), β_0 is a constant term, $X_{i,j}$ represents the ith observation of the jth explanatory variables (predictors of TFI Reduction), β_j is the associated jth coefficient. As λ increases, the absolute value amount of the estimated coefficient shrinks towards zero. Under LASSO, the coefficient of unimportant variables is reduced completely to zero. Ridge does not perform variable selection but minimizes the impact of irrelevant predictors in the model shrinking the estimated coefficient near zero but not completely zero. Whenever $\lambda = 0$ it produces the results of the normal Ordinary Least Squares regression. These techniques effectively address collinearity by design, balancing the coefficients of correlated variables—reducing them to small and negative values when one is significantly large. They also adjust for instances where **X** lacks full rank.

2.3. KNN Regression

Let $X = \{x_1, x_2, \ldots, x_n\}$ be the set of training data and $y = \{y_1, y_2, \ldots, y_n\}$ be the set of corresponding TFI Reduction scores. For a new patient data point x', the KNN algorithm searches the training set to find the k nearest neighbors based on a distance metric (Euclidean distance was adopted in this study) given in (6) by:

$$D(x, x') = \sqrt{(x_1 - x_1')^2 + (x_2 - x_2')^2 + \cdots + (x_n - x_n')^2} \tag{6}$$

The KNN algorithm was then used to calculate the average severity score of these k neighbors to predict the TFI Reduction score for the new patient:

$$\hat{y} = \frac{1}{k}\sum_{i=1}^{k} y_i$$

where y_i are the severity scores of the k nearest neighbors to x'.

3. Results and Findings

Table 2: Descriptive statistics of quantitative variables

Specification	HHI	GAD	PHQ	ISI	SWLS	HYP	CFQ	Age	Duration	PRE	POST
count	142.00	142.00	142.00	142.00	142.00	142.00	142.00	142.00	142.00	142.00	125.00
mean	17.79	7.48	8.03	12.96	20.32	19.04	40.59	55.45	11.99	59.37	50.52
std	11.37	5.58	5.67	7.04	7.36	8.50	15.96	12.88	12.50	18.25	21.88
min	0.00	0.00	0.00	0.00	5.00	1.00	7.00	22.00	0.30	24.40	4.00
25%	8.00	3.00	4.00	8.00	14.00	13.00	29.25	46.25	3.00	46.80	32.00
50%	18.00	6.00	7.00	13.00	20.00	18.50	41.00	58.00	10.00	58.60	57.10
75%	26.00	11.00	11.00	18.00	26.00	25.00	50.00	65.00	15.00	73.60	66.00
max	40.00	21.00	27.00	27.00	35.00	42.00	86.00	83.00	55.00	97.20	88.40

Table 2 offers a statistical summary for a selection of variables collected from 142 participants, including TFI score after the completion of the study which has data for 125 participants. The 'HHI' has an average of 17.79, with a standard deviation of 11.37, and ranges from 0 to 40. The 'GAD' and 'PHQ' scores, which assess anxiety and depression, have similar counts and display averages of 7.48 and 8.03, respectively. Both PHQ and ISI have a maximum score of 27. 'ISI' and 'SWLS' scores indicate sleep disturbance and satisfaction with life, averaging 12.96 and 20.32 with a standard deviation of 7.04 and 7.36, respectively. Hyperacusis, or increased sensitivity to sound, has an average score of 19.04, while 'CFQ' averages at 40.59, possibly measuring cognitive failures. Participants' age averages at 55 years with a standard deviation of approximately 13 years. The duration of Tinnitus averages nearly 12 years, and TFI score at the beginning of the study averages at 59.37, with TFI score after the completion of the study having a lower average of 50.52, indicating a potential improvement after an intervention.

3.1. Correlation Analysis of Predictors

The correlogram provides a visual and quantitative depiction of the correlation coefficients between pairs of study variables. Correlation coefficients range from -1 to 1, where values close to 1 indicate a strong positive correlation, values close to -1 indicate a strong negative correlation, and values around 0 indicate no linear relationship. The color scale enhances interpretability: red shades signify positive correlations, blue shades indicate negative correlations, and the intensity of the color corresponds to the strength of the relationship.

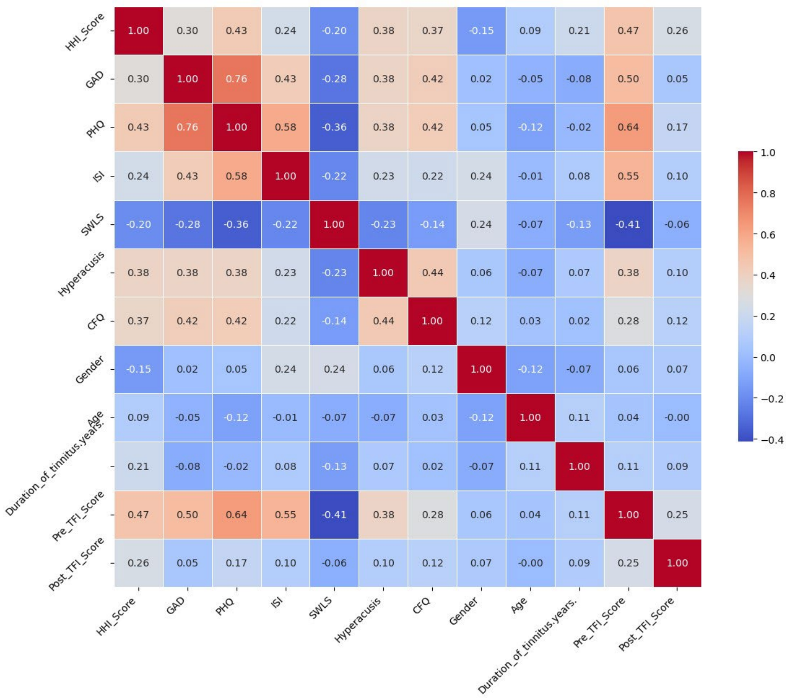

Figure 1: Correlogram of Variables

From Figure 1, notably strong positive correlations are observed between GAD and PHQ scores (0.76), illustrating that higher anxiety levels are associated with higher depression symptomatology. Similarly, PHQ scores and *TFI score at the beginning of the study* (0.64) are strongly correlated, suggesting a significant relationship between depression symptoms and the impact of Tinnitus before treatment. SWLS (Satisfaction with Life Scale) shows negative correlations with GAD, PHQ, and *TFI score at the beginning of the study*, indicating that higher life satisfaction is inversely related to anxiety, depression, and the perceived impact of Tinnitus. This underpins the psychological impact of Tinnitus and its comorbidities on life satisfaction. Moderate positive correlations were observed between GAD and *TFI score at the beginning of the* *study* (0.50). Additionally, no linear relationship was observed between Age and *TFI score at the end of the study* (-0.00). The relationship between Hyperacusis and both PHQ and CFQ suggests that sensitivity to sound is linked with hearing challenges and cognitive failures.

3.2. Distribution of quantitative variables

From Figure 2, the 'HHI Score' histogram shows a slightly left-skewed distribution, while 'GAD' (Generalized Anxiety Disorder) and 'PHQ' (Patient Health Questionnaire) show a moderate right skew. The 'ISI' (Insomnia Severity Index), 'SWLS' (Satisfaction with Life Scale), and 'Hyperacusis' measures appear to be somewhat normally distributed, with 'SWLS' showing a slight left skew. The 'CFQ' (Cognitive Failures Questionnaire) histogram suggests a normal distribution with a peak around the 40 score mark. The 'Age' distribution is slightly right-skewed, with a concentration of participants in the middle-age range. The histogram for Duration of Tinnitus in years is highly right-skewed, with most participants having a shorter duration of Tinnitus. Lastly, the 'Pre TFI Score' and 'Post TFI Score' histograms, assessing the impact of Tinnitus before and after treatment, show that most participants' scores are moderately high with a right skew, and the 'Post TFI Score' appears to have a slightly more pronounced peak, indicating a possible concentration of scores post-treatment.

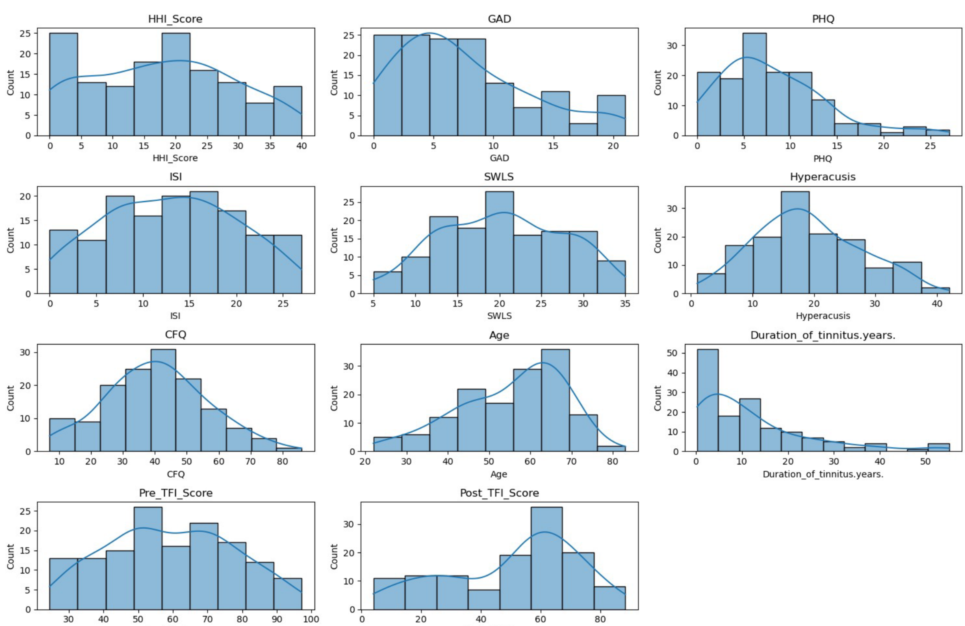

Figure 2: Histogram of quantitative variables

3.3. Descriptive Statistics of qualitative variables

Figure 3 presents two charts displaying the distribution of the study's participants across different groups and genders. The pie chart on the left shows the split between participants in the control group (73 individuals corresponding to 51.4%) and those in the treatment group (69 individuals corresponding to 48.6%). The bar chart on the right indicates the gender distribution, with a count of 80 male participants, making up 56.3% of the total, and 62 female participants, accounting for 43.7%. Both charts include the actual counts along with the corresponding percentages, providing a clear visual representation of the composition of the study's sample.

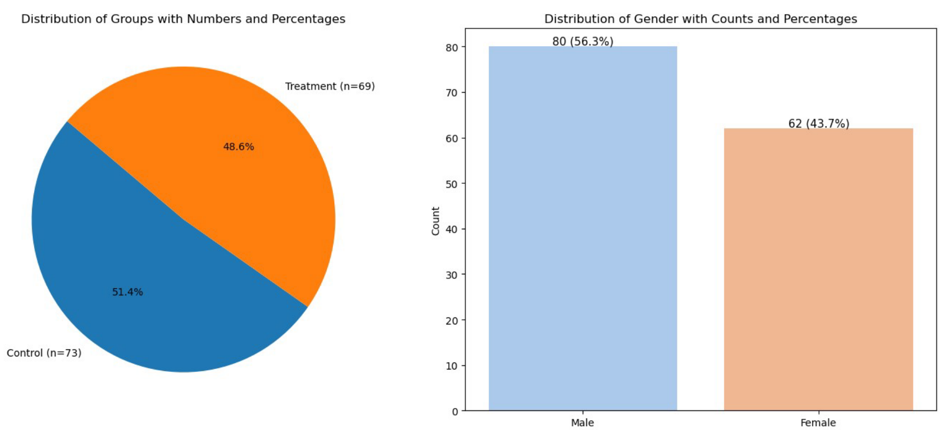

Figure 3: Distribution of Groups and Gender

3.4. Ordinary Least Squares Regression Model

In this section, we utilized the multiple linear regression with best subset selection to predict the TFI Reduction.

Table 3: Full Regression Model Coefficients

Variable	Estimate	Std. Error	$t-value$	$P>\|t\|$	[0.025	0.975]
Constant	-1.0735	13.109	-0.082	0.935	-27.078	24.931
HHI Score	-0.0960	0.196	-0.490	0.625	-0.485	0.293
GAD	0.1591	0.504	0.316	0.753	-0.841	1.159
PHQ	0.2950	0.602	0.490	0.625	-0.899	1.489
ISI	0.8482	0.333	2.551	0.012	0.189	1.508
SWLS	-0.5744	0.284	-2.022	0.046	-1.138	-0.011
Hyperacusis	0.2910	0.244	1.194	0.235	-0.193	0.775
CFQ	-0.0635	0.136	-0.468	0.641	-0.333	0.206

Table 3: Full Regression Model Coefficients (cont.)

Variable	Estimate	Std. Error	*t – value*	*P* > |*t*|	[0.025	0.975]
Gender	-2.1539	4.072	-0.529	0.598	-10.232	5.924
Age	-0.0281	0.146	-0.193	0.848	-0.318	0.262
Duration of Tinnitus (in years)	-0.0886	0.143	-0.620	0.537	-0.372	0.195
Group	25.3661	3.610	7.027	0.000	18.206	32.527

The full regression model in 3 presents the effects of various predictors on TFI reduction. The intercept, while estimated at -1.0735, is not statistically significant (p = 0.935), indicating the threshold TFI Reduction, in the absence of all the predictors. On the other hand, the coefficient of GAD implies that the TFI Reduction would increase by 0.1591 per unit change in SWLS when all other factors are kept constant. The variables with positive coefficients, i.e. PHQ, ISI and Hyperacusis can be interpreted in the same manner as GAD. This means that GAD, PHQ, ISI, Hyperacusis and TFI Reduction move in the same directions, i.e. TFI Reduction increases as they increase and also TFI Reduction decreases as they decrease. Additionally, when all other variables remain constant, the coefficient of HHI Score predicts that TFI Reduction would decrease by 0.0960 per unit change in HHI Score. It also suggests that TFI Reduction and HHI Score have a negative or inverse relationship. SWLS, CFQ, Age and Duration of Tinnitus (in years) all have a negative coefficient and hence can be explained in the same way as HHI _Score. Considering the categorical variables, we observed that the coefficient for gender (male set as the reference category) is -2.1539, which is not statistically significant at the 5% significance level (p = 0.598). A negative coefficient indicates that females, on average, have a lower TFI Reduction compared to males. Likewise, the coefficient for Group (treatment set as the reference category) is 25.3661, which is statistically significant at the 5% significance level (p = 0.000). A positive coefficient indicates that the control group has a higher TFI reduction compared to the treatment group. At the 5% significance level, we observed that ISS (p = 0.012), SWLS (0.046) and group (0.000) were the only significant predictors of TFI Reduction. The other predictors do not statistically contribute significantly to the prediction of TFI Reduction.

Table 4: Summary Statistics for Full Regression Model

***R*2**	**Adjusted-R^2**	**F-statistic**	**Prob (F-statistic)**
0.448	0.388	7.451	0.0000

From Table 4, the R^2 value of 0.448 implies that approximately 44.80% of the variations in TFI Reduction was explained by all the predictors (HHI Score, GAD, PHQ, ISI, SWLS, Hyperacusis, CFQ, Gender, Age, Duration of Tinnitus (in years) and Group). Likewise, the Adjusted R^2 value of 0.388 implies that the percentage of variation explained by the independent variables that actually affect TFI Reduction is approximately 38.80%. The p-value of the F-statistic (p=.0000) which is less than 0.05 indicates that the full regression model is statistically significant. This means that at least one of the predictors contribute significantly to predicting TFI Reduction.

3.4.1. Best Subset Selection Criteria

In our pursuit to refine the predictive accuracy of the comprehensive model, the selection of predictors underwent a rigorous statistical evaluation. The criterion for this accurate selection was their statistical significance in predicting the reduction in TFI (TFI Reduction). To achieve this, we employed five distinct measures namely Akaike Information Criterion (AIC), Akaike Information Criterion corrected (AICc), Bayesian Information Criterion (BIC), Mallow's Criterion (C_p) and Final Prediction Error (FPE) as our guiding metrics. These criteria are instrumental in identifying models that strike an optimal balance between complexity and fit, thereby ensuring that only the most relevant predictors are included in the final model. This strategic approach aimed at both minimizing overfitting and enhancing the model's overall predictive capacity.

Now, we let X_1 =Group, X_2 = ISI, X_3 = SWLS, X_4 = Hyperacusis, X_5 =Duration of Tinnitus (in years), X_6 = Gender, X_7 = HHI Score, X_8 =PHQ, X_9 = CFQ, X_{10} =GAD and X_{11} = Age.

Table 5: Best Subsets Regression Summary

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	AIC	AICc	BIC	FPE	Cp
1 ✓											1005.44	1005.47	1008.16	428.28	40.44
2 ✓	✓										1004.55	1004.66	1010.01	424.95	38.62
3 ✓	**✓**	**✓**									**980.64**	**980.86**	**988.82**	**343.91**	**10.37**
4 ✓	✓	✓	✓								981.57	981.95	992.49	346.77	11.27
5 ✓	✓	✓	✓	✓							982.30	982.86	995.94	349.01	11.96
6 ✓	✓	✓	✓	✓	✓						983.85	984.64	1000.21	353.85	13.51
7 ✓	✓	✓	✓	✓	✓	✓					985.59	986.65	1004.67	359.35	15.24
8 ✓	✓	✓	✓	✓	✓	✓	✓				986.83	988.22	1008.65	363.37	16.48
9 ✓	✓	✓	✓	✓	✓	✓	✓	✓			988.57	990.32	1013.12	369.03	18.22
10 ✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		990.47	992.63	1017.75	375.34	20.12
11 ✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	992.36	994.97	1022.36	381.72	22.01

NB: A ✓ indicates that a given variable is included in the corresponding model. For instance, the output in the first row in Table 5 above indicates that the model contains only Group as the predictor. Here, the intercept is included in all models.

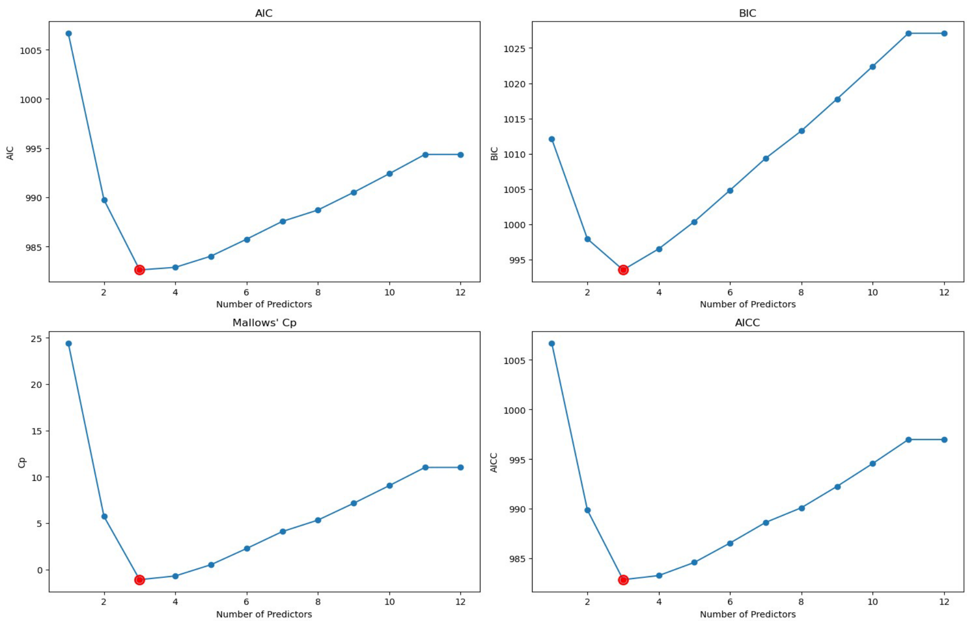

Figure 4: Plots of AIC, AICc, BIC and C(p) for p Subsets

From Table 5 and Figure 4, the best subsets of predictors with the least AIC, AICc, BIC and C(p) values include Group, ISI and SWLS. The FPE value in Table 5 also confirms this claim. It is also worth knowing that the results of both the forward and backward selection criteria (not included in this because of brevity) were also in conformity with the one obtained by best subset selection criteria.

The best subset model can be expressed mathematically in (7) as:

$$\textbf{TFI Reduction} = \beta_0 + \beta_1 * \textbf{Group} + \beta_2 * \textbf{ISI} + \beta_3 * \textbf{SWLS}. \quad (7)$$

Table 6 gives the coefficients and summary of the best subsets regression model.

Table 6: Model Coefficients for Best Subset Model from OLS Regression

Variable	Estimate	Std. Error	$t - value$	$P > \|t\|$	[0.025	0.975]
Constant	-0.9332	6.998	-0.132	0.895	-14.793	12.947
Group	25.6116	3.490	7.338	0.000***	18.694	32.529
ISI	0.9787	0.255	3.840	0.000***	0.474	1.484
SWLS	-0.7194	0.238	-3.028	0.003***	-1.190	-0.249

Based on the regression coefficients in Table 6, the linear regression equation of the best subset model is given in (8) by

$$\textbf{TFI Reduction} = -0.9232 - 25.6116 * \textbf{Group} + 0.9787 * \textbf{ISI} - 0.7194 * \textbf{SWLS}. \quad (8)$$

This relationship shows that TFI Reduction decreases with a corresponding change in SWLS when all other variables are kept constant. This means that a one unit increase in SWLS is expected to decrease TFI reduction by 0.7194. Also, TFI Reduction increases with a corresponding increase in ISI when all factors are kept constant. In general, a one unit increase in ISI is expected to increase TFI reduction by 0.9787. Both ISI (p=0.000) and SWLS (p=0.003) are statistically significant at the 5% significance. It was also observed that the coefficient for Group (treatment category as the reference category) is 25.6116, which is statistically significant at the 5% significance level (p = 0.000). A positive coefficient indicates that the control group has a higher TFI reduction compared to the treatment group.

Table 7: Summary Statistics for Best Subset Regression Model

***R*2**	**Adjusted-R^2**	**F-statistic**	**Prob (F-statistic)**
0.427	0.411	27.04	0.0000

From Table 7, the R^2 value of 0.427 implies that approximately 42.70% of the variations in TFI Reduction was explained by Group, ISI and SWLS. Likewise, the Adjusted R^2 value of 0.411 implies that the percentage of variation explained by Group, ISI and SWLS that actually affect TFI Reduction is approximately 41.10%. The p-value of the F-statistic (p=.0000) which is less than 0.05 indicates that the best subset regression model is statistically significant. This means that at least one of the predictors contributes significantly to predicting TFI Reduction but in this case all the three predictors (Group, ISI and SWLS) statistically contribute significantly to the prediction of TFI Reduction. The increment in the value of the adjusted R^2 from 0.388 to 0.411 is due to the the removal of the non-significant predictors and thus improving the overall model fit.

3.5. Tests of Model Assumptions

Kutner et al. (2005) defines model validity as the stability and reasonableness of the regression coefficients, the plausibility and usability of the regression function and the ability to generalize inferences drawn from the regression analysis. Validation is a valuable and an essential part of the model building process. Torres-Reyna (2007) established that how good a model is depends on how well it predicts the dependent variable. The tests of the model assumptions are performed on the normality of the error terms, presence of homoscedasticity, independence of the residuals, multicollinearity and linearity of the regression function.

3.5.1. Test for Normality of Residuals

An informal test of normality was conducted to check the distribution of the data using the Normal Q-Q plot. From Figure 5, it could be seen that most of the data values fall within the confidence bands with some few deviations showing that normality is a suspect. This informal test is further augmented by the Sharpiro Wilk formal test for normality.

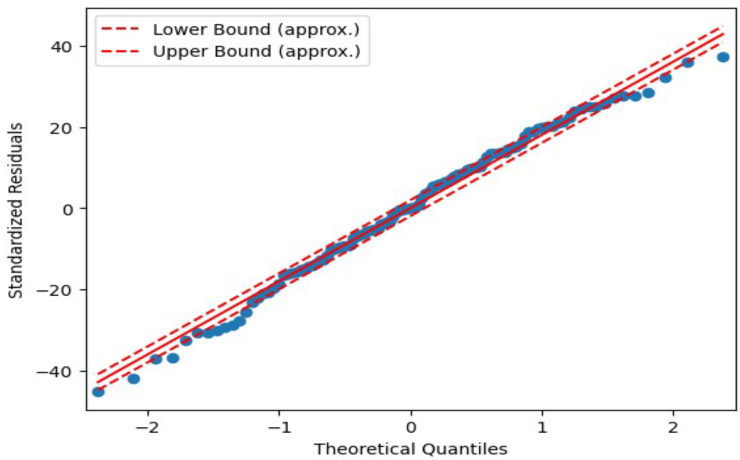

Figure 5: Normal Q-Q with Approximate Bounds

3.5.2. Sharpiro Wilk test for Normality

From Table 8, at the 5% significance level, since the p-value is greater than the alpha level (0.2195 > 0.05) we fail to reject the null hypothesis of normality satisfied and conclude that normality assumption is indeed satisfied. This confirms the initial claim **using the Q-Q plot.**

Table 8: Sharpiro Wilk test statistic and sig-value

Test	Test Statistic	P-value
Shapiro-Wilk	0.9845	0.2195

3.5.3. Test for homoscedasticity of the Residuals

As a prerequisite, the outcome of the test for normality of the error term above satisfies the condition to go on with the test for homoscedasticity. As an informal test, we resort to the scale-location plot.

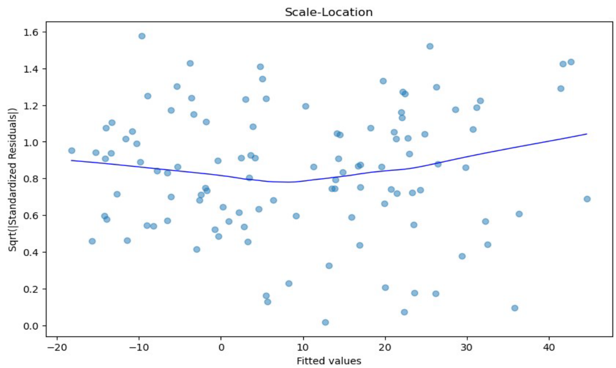

Figure 6: Plot of Scale Location

From Figure 6, we observe that at every fitted value, the spread of the residuals is roughly the same. We thus conclude that homoscedasticity seems plausible. Likewise in Figure 7, the spread of residuals remains constant across different fitted values indicative of homoscedasticity. The Breusch-Pagan test was employed as a formal test to verify our claim.

Table 9: Breusch-Pagan test statistic and sig-value

Test	Test Statistic	P-value
Breusch-Pagan	3.3205	0.3448

Using the Breusch-Pagan test, from Table 9 since the significance probability value = 0.3448 > 0.05, we fail to reject the null hypothesis of homoscedasticity, indicating no strong evidence of heteroscedasticity. This implies that heteroscedasticity is not present thus homoscedasticity assumption is satisfied.

3.5.4. Test for Multicollinearity

The collinearity test was conducted using the Variance Inflation Factor (VIF) approach. The resulting Collinearity statistics are presented in Table 10:

Table 10: Collinearity Statistics

Variable	VIF	VIF> 5
Group	1.0162	False
ISI	1.0654	False
SWLS	1.0551	False

From Table 10, since all the VIFs are all quite moderate ($VIF < 5$), there is no evidence of serious collinearity. Thus, the predictor variables (Group, ISI and SWLS) are not highly correlated with each other.

3.5.5. Test for Independence of Residuals

The Durbin-Watson test of autocorrelation was adopted to test for the independence of residuals of the best subset fitted regression model.

Table 11: Durbin-Watson Test Statistics and sig-value

Test	Lag	Test Statistic
Durbin-Watson	1	2.1254

From Table 11, since the Durbin-Watson Statistic is approximately 2, we do not reject the null hypothesis of no autocorrelation, and therefore, conclude that there is little to no auto or serial correlation in the residuals, indicating that the residuals are random (thus, the residuals are independent of each other).

3.5.6. Testing for linearity of the regression function

Here, we test to see if the regression function is linear using the Residual versus Fitted plot as an informal test.

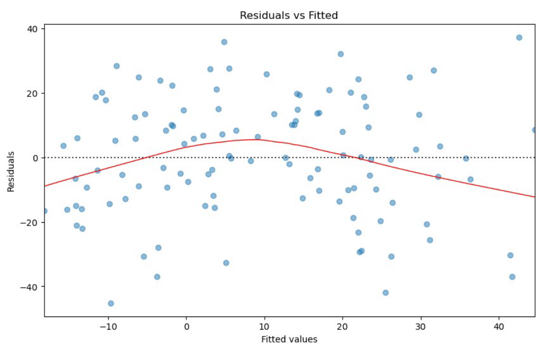

Figure 7: Residuals versus Fitted Plot

From the residuals versus fitted plot in Figure 7, we observe that at every fitted value, the spread of the residuals is roughly the same. We thus conclude that linearity seems plausible. From Figure 7, we also observe that there is no pattern in the residual plot. This suggests that we can assume a linear relationship between the predictors and the outcome variables. Since this is the case, linearity is a suspect. The Harvey-Collier test is then employed as a formal test to validate our suspicion.

Table 12: Harvey-Collier test statistic and sig-value

Test	Test Statistic	P-value
Harvey-Collier	-0.3094	0.7576

From Table 12, we clearly see that since the p-value is greater than the indicated alpha level ($0.7576 > 0.05$), we do not reject the null hypothesis of linearity, indicating no strong evidence of non-linearity. Hence, the regression function is linear.

In summary, since the best subset regression model satisfies all the assumptions it can be used for predictions based on the test set. The mean square error (MSE) using the test set was found to be **330.76**.

3.6. KNN Regression Results and Data Analysis

In addition to employing the conventional least squares estimation technique, our analysis extends to incorporating KNN regression to assess the impact of the independent variables on the target variable, TFI Reduction in this section. Unlike the traditional regression methods, KNN regression does not presuppose a specific functional form for the outcome variable. This deviation provides enhanced flexibility, allowing the model to adapt more freely to the underlying data structure. The primary objective of integrating this method is to compare and ascertain which algorithm, between least squares and KNN regression, exhibits superior predictive accuracy for the Tinnitus dataset in question. By adopting this dual approach, we aim to gain deeper insights into the dynamics influencing TFI Reduction and identify the most effective predictive framework based on the characteristics of the available data.

In this section, we present 5 different KNN regression models fitted for K = 2, 4, 6, 8, 10 and their test mean squared errors are presented in Figure 8 and Table 13.

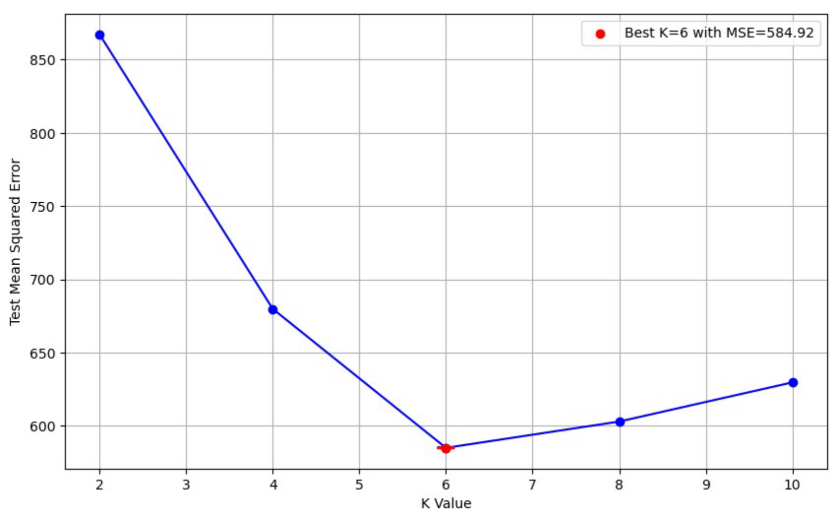

Figure 8: KNN Regression: Plot of Test MSE for Different K values

The outcomes of the cross-validation process from both Figure 8 and Table 13 indicate that the choice of $K = 6$ results in the minimal mean squared error on the test data. Consequently, this particular value was selected for our predictive analysis concerning TFI Reduction using the predictor variables. Given the complexities involved in formulating a direct equation to express TFI Reduction as a function of these predictors, we opted to employ a K-Nearest Neighbors (KNN) regression model. By applying this model with K set to 6 for our test dataset, we were able to compute the predictions. The performance of these predictions was quantified using the Mean Squared Error (MSE), which amounted to **584.92**, providing a quantitative measure of the prediction accuracy for the TFI Reduction based on the selected predictors.

Table 13: Test MSE for Different K values

K-value	2	4	6	8	10
Test MSE	867.26	679.83	584.92	602.88	629.60

3.7. Lasso and Ridge Regression Analysis

In this section of our analysis, we explore the efficacy of shrinkage methodologies—specifically Ridge and Lasso regression—in predicting the impact of various predictors on the TFI Reduction. The underpinning rationale for employing these techniques lies in their capacity to impose a shrinkage penalty on the coefficients, thereby mitigating the risk of overfitting and enhancing model generalizability compared to the traditional Ordinary Least Squares (OLS) approach. Employing 10-fold cross-validation, we embarked on a rigorous search for the optimal λ (tuning parameter), which minimizes the Mean Squared Error (MSE). This endeavor is not merely a quest for minimal error but a strategic move to discern the most robust model that harmonizes complexity and prediction accuracy. Figure 9 offers a visual exposition of this optimization process, showcasing the relationship between various λ values and their MSE outcomes. This visual analysis is important as it illuminates the trade-offs inherent in model regularization and aids in the selection of a model that judiciously balances bias and variance.

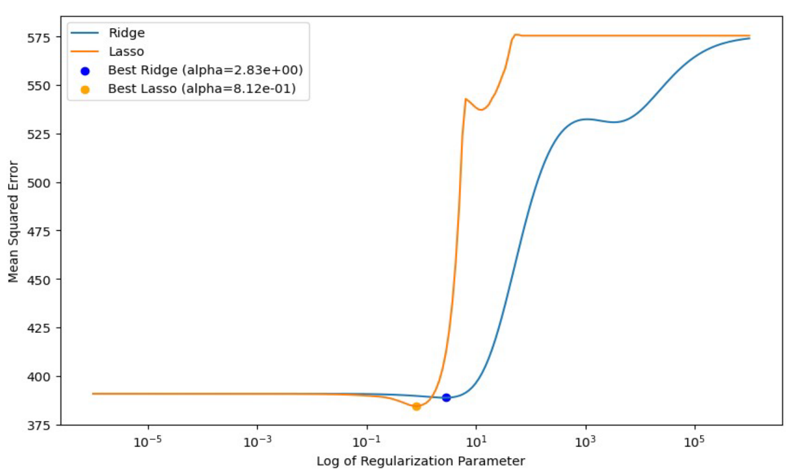

Figure 9: Cross-Validated λ Values for Ridge and Lasso Regression

From the cross-validation plot, the optimal λ-value with the least MSE for each method (Ridge and Lasso) is 2.83 for Ridge and 0.812 for Lasso. We made use of these values for the tuning parameter, λ in each method, to build the model and also to perform the prediction for the testing data.

3.8. Coefficients Estimate and Testing Errors Using the Optimal λ-Values

Table 14 presents the coefficient estimates for each model and their respective test errors of the final model produced by the optimal λ-values.

Table 14: Model Coefficients and MSEs for Ridge and Lasso Shrinkage Methods

Predictors	Ridge (Coefficients)	Lasso (Coefficients)
(Intercept)	0.35578	-1.11778
HHI Score	-0.10206	-0.0633
GAD	0.18036	0.18863
PHQ	0.28242	0.24065
ISI	0.82577	0.77362
SWLS	-0.58996	-0.62518
Hyperacusis	0.28859	0.26886
CFQ	-0.06558	-0.07340
Gender	-2.07104	.
Age	-0.01991	-0.01068
Duration of Tinnitus	-0.09540	-0.09591
Group	22.97202	22.12796
Testing Error	**318.30088**	**319.28195**

The coefficients indicate the relationship between each predictor and the response variable (TFI Reduction). Positive coefficients suggest a direct relationship, while negative coefficients suggest an inverse relationship. Common trends observed in both models include the negative impact of the HHI Score, SWLS, CFQ, Age and Duration of Tinnitus (in years); and the positive impact of measures like PHQ, ISI, and Hyperacusis on TFI Reduction. Differences between the models are evident in the magnitude of the coefficients, with Lasso regression driving some coefficients to zero (e.g., Gender in Lasso has no estimate), indicating it performs variable selection. The MSE for Ridge regression is slightly lower **(318.30)** than for Lasso regression **(319.28)**, suggesting that Ridge regression might be slightly more effective in predicting TFI Reduction with the given dataset and chosen λ values. Both Age and duration of Tinnitus have a slight negative association with TFI Reduction in both models, suggesting that older individuals and those who have had Tinnitus for longer period might experience less reduction. A positive coefficient for Group in both models suggests that the control group has a higher TFI reduction compared to the treatment group.

4. Discussion and Study Alignment with Health Economics

This study introduces several novel aspects in terms of statistical analysis and methodology that contribute to the existing body of literature on Tinnitus management and treatment effectiveness. It also enriches the statistical analysis of Tinnitus treatment effectiveness through the application of advanced machine learning techniques, rigorous model validation, and comprehensive residual analysis. The study employs a variety of supervised machine learning algorithms, including Ordinary Least Squares (OLS), K-Nearest Neighbor (KNN), Ridge regression, and Lasso regression. Each method is analyzed for its predictive accuracy in modelling Tinnitus Functional Index (TFI) reduction. The inclusion of these diverse algorithms allows for a comprehensive comparison of their performance, highlighting the strengths and weaknesses of each approach in the context of Tinnitus treatment data. Secondly, the application of Ridge and Lasso regressions introduces regularization techniques that address multicollinearity and improve model stability. These methods add a shrinkage penalty to the regression coefficients, which helps in reducing the variance of the estimates. Ridge regression demonstrated superior performance with the smallest test mean square error (MSE) of 318.30, showcasing its effectiveness in handling multicollinear predictors and providing more reliable predictions. Also, the study employs the best subset selection criteria to identify the most significant predictors of TFI reduction. This involves evaluating all possible combinations of predictors and selecting the model that optimally balances complexity and predictive power. The criteria used include Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Mallow's Cp, which are instrumental in ensuring the selected model is both parsimonious and highly predictive.

Moreover, the research emphasizes rigorous model validation techniques, including 10-fold cross-validation, to assess the robustness and generalizability of the models. This approach helps in mitigating overfitting and ensures that the models perform well on unseen data. Various statistical metrics, such as R-squared, adjusted R-squared, and F-statistics, are used to evaluate model fit and significance, providing a detailed understanding of model performance. Additionally, the study conducts extensive residual analysis to check the assumptions underlying the regression models. This includes tests for normality (Shapiro-Wilk test), homoscedasticity (Breusch-Pagan test), independence of residuals (Durbin-Watson test), and multicollinearity (Variance Inflation Factor). Such thorough diagnostic checks ensure the validity and reliability of the regression models, enhancing the credibility of the findings. To add to the above, the use of correlograms and histograms provides visual insights into the relationships between variables and their distributions. These visual tools help in understanding the data structure and the nature of the predictors. The study also presents visualizations of the cross-validation process for selecting the optimal lambda values in Ridge and Lasso regressions, aiding in the interpretation of the regularization effects. The statistical findings highlight the importance of a holistic approach to Tinnitus management, considering factors beyond just the auditory symptoms. Aside from the statistical implications, the study aligns closely with the field of health economics in several significant ways such as cost-effectiveness analysis, resource allocation, quality of life and economic productivity and long-term economic benefits.

By identifying significant predictors of TFI reduction, the study facilitates the development of more targeted treatment strategies. This can lead to more efficient use of healthcare resources, reducing unnecessary treatments and associated costs. Machine learning models, such as Ridge and Lasso regressions, enable the prediction of treatment outcomes with higher accuracy, allowing healthcare providers to allocate resources more effectively and prioritize interventions that are likely to yield the best results. The findings of this study can inform policy decisions regarding the allocation of resources for Tinnitus treatment. By understanding which factors most significantly influence treatment outcomes, policymakers can prioritize funding and resources towards interventions that address these key areas. Efficient resource allocation based on predictive modeling can help reduce the economic burden of Tinnitus on both patients and the healthcare system as indicated by Tuepker et al. (2018). Tinnitus, particularly when severe, can significantly impair an individual's quality of life and economic productivity. By improving the precision of treatment outcomes, the research contributes to enhancing patients' quality of life, which in turn can have positive economic implications. Better management of Tinnitus can lead to reduced absenteeism and increased productivity among individuals affected by the condition, contributing to broader economic benefits.

5. Conclusion and Recommendations

The main objective of the study was to model Tinnitus Functional Index (TFI) Reduction via supervised machine learning algorithms. Notably, Ordinary Least Squares (OLS), Nearest Neighbor (KNN), Ridge and Lasso regressions were the main statistical tools adopted for the study. The OLS regression revealed that Group, ISI and SWLS were the main significant predictors of TFI Reduction using the best subset selection criteria which was also confirmed by both the forward and backward selection criteria. It was found out that the two shrinkage methods (Ridge and Lasso) outperformed the OLS and the KNN regression. Based on the test mean square error (MSE), Ridge regression was chosen as the best supervised machine learning algorithm for the analysis of the Tinnitus data under consideration, with the smallest test MSE of **318.30** compared to that of Lasso, OLS, KNN regression's respective test MSE of **319.28**, **330.76** and **584.92**. Hence, by the metric considered, Ridge regression was ranked more accurate in predicting TFI Reduction than the other three methods. The exploration of supervised machine learning algorithms for modeling Tinnitus Functional Index (TFI) reduction presents a promising frontier in the personalized treatment of Tinnitus. Studies such as those by Rodrigo et al. (2021), and investigations into the predictive capabilities of EEG sensors and deep learning algorithms by Doborjeh et al. (2023), demonstrate significant advancements in predicting patient responses to therapies and classifying Tinnitus severity. These machine learning approaches, including decision tree models, random forests, and neural networks, offer new insights into the complex nature of Tinnitus and its management. By effectively predicting treatment outcomes, these tools can facilitate more targeted and effective interventions, ultimately improving the quality of life for individuals suffering from Tinnitus.

As research continues to evolve, the integration of AI and machine learning into Tinnitus treatment protocols holds the potential to revolutionize the field, offering hope for effective management strategies for this challenging condition. The study revealed a significant difference between the treatment and control groups, with the control group showing a higher TFI reduction. This finding may suggest that factors outside the treatment regimen, such as individual coping mechanisms or placebo effects, could influence TFI outcomes. This warrants further investigation to understand the underlying reasons for the observed group differences. Future studies could investigate the integration of multimodal data, such as combining EEG patterns with behavioral and clinical metrics, to develop more comprehensive predictive models. Additionally, exploring the longitudinal application of machine learning algorithms could offer insights into the progression of Tinnitus over time and the long-term effectiveness of different treatment modalities. There is also a significant opportunity to refine machine learning models by incorporating patient feedback loops, enabling the models to learn from each treatment outcome and continuously improve predictions. The study's recommendation to explore the integration of multimodal data and longitudinal applications of machine learning algorithms has long-term economic benefits. Continuous improvement in predictive capabilities can lead to sustained enhancements in treatment strategies, reducing long-term healthcare costs. Investing in such modernistic research aligns with health economics principles of achieving long-term cost savings and health improvements through advanced technological solutions.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability statement

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Additional Information

No additional information is available for this paper.

Acknowledgement

The author acknowledges the enormous support of the University of Texas Rio Grande Valley (UTRGV) Presidential Research Fellowship.

References

Agbota, L., Agyemang, E., Kissi-Appiah, P., Moshood, L., Osei-Nkwantabisa, A., Agbenyeavu, V., Nsiah, A., and Adjei, A. (2024). Enhancing tumor classification through machine learning algorithms for breast cancer diagnosis. *Computer Engineering and Intelligent Systems*, 15(1), p. 71.

Biswas, R., Genitsaridi, E., Trpchevska, N., Lugo, A., Schlee, W., Cederroth, C. R., Gallus, S., and Hall, D. A., (2023). Low evidence for tinnitus risk factors: A systematic review and meta-analysis. *Journal of the Association for Research in Otolaryngology*, 24(1), pp. 81–94.

Cardon, E., Jacquemin, L., Schecklmann, M., Langguth, B., Mertens, G., Vanderveken, O. M., Lammers, M., Van de Heyning, P., Van Rompaey, V., and Gilles, A., (2022). Random forest classification to predict response to high-definition transcranial direct current stimulation for tinnitus relief: A preliminary feasibility study. *Ear and Hearing*, 43(6), p. 1816.

De Ridder, D., Schlee, W., Vanneste, S., Londero, A., Weisz, N., Kleinjung, T., Shekhawat, G. S., Elgoyhen, A. B., Song, J.-J., Andersson, G., et al., (2021). Tinnitus and tinnitus disorder: Theoretical and operational definitions (an international multidisciplinary proposal). *Progress in brain research*, 260, pp. 1–25.

Doborjeh, M., Liu, X., Doborjeh, Z., Shen, Y., Searchfield, G., Sanders, P., Wang, G. Y., Sumich, A., and Yan, W. Q., (2023). Prediction of tinnitus treatment outcomes based on eeg sensors and tfi score using deep learning. *Sensors*, 23(2), p. 902.

Eggermont, J. J., Roberts, L. E., (2004). The neuroscience of tinnitus. *Trends in neurosciences*, 27(11), pp. 676–682.

Gasparre, D., Pepe, I., Laera, D., Abbatantuono, C., De Caro, M. F., Taurino, A., D'Erasmo, D., Fanizzi, P., Antonucci, L. A., Pantaleo, A., et al., (2023). Cognitive functioning and psychosomatic syndromes in a subjective tinnitus sample. *Frontiers in Psychology*, 14.

Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W., (2005). *Applied linear statistical models*. McGraw-hill.

Langguth, B., Kreuzer, P. M., Kleinjung, T., and De Ridder, D., (2013). Tinnitus: causes and clinical management. *The Lancet Neurology*, 12(9), pp. 920–930.

Manta, O., Sarafidis, M., Schlee, W., Mazurek, B., Matsopoulos, G. K., and Koutsouris, D. D., (2023). Development of machine-learning models for tinnitus-related distress classification using wavelet-transformed auditory evoked potential signals and clinical data. *Journal of Clinical Medicine*, 12(11), p. 3843.

McCormack, A., Edmondson-Jones, M., Somerset, S., and Hall, D., (2016). A systematic review of the reporting of tinnitus prevalence and severity. *Hearing research*, 337, pp. 70–79.

Mohan, A., Leong, S. L., De Ridder, D., and Vanneste, S., (2022). Symptom dimensions to address heterogeneity in tinnitus. *Neuroscience & Biobehavioral Reviews*, 134, p 104542.

Omer, P. A., (2022). Improving prediction accuracy of lasso and ridge regression as an alternative to ls regression to identify variable selection problems. *Zanco Journal of Pure and Applied Sciences*, 34(s6), pp. 33–45.

Rodrigo, H., Beukes, E. W., Andersson, G., and Manchaiah, V., (2021). Exploratory data mining techniques (decision tree models) for examining the impact of internet-based cognitive behavioral therapy for tinnitus: Machine learning approach. *Journal of Medical Internet Research*, 23(11), e28999.

Runciman, L., Johnson, C., (2023). Young adults' knowledge and perceptions of permanent noise-induced tinnitus and its influence on behavioural intentions. *Noise and Health*, 25(119), pp. 236–246.

Saleh, A. M. E., Norouzirad, M., (2018). On shrinkage estimation: Non-orthogonal case. *Statistics, Optimization & Information Computing*, 6(3), pp. 427–451.

Smith, S. S., Kitterick, P. T., Scutt, P., Baguley, D. M., and Pierzycki, R. H., (2021). An exploration of psychological symptom-based phenotyping of adult cochlear implant users with and without tinnitus using a machine learning approach. *Progress in Brain Research*, 260, pp. 283–300.

Torres-Reyna, O., (2007). Linear regression using stata. *Princeton University. New Jersey: Princeton University*.

Tuepker, A., Elnitsky, C., Newell, S., Zaugg, T., and Henry, J. A., (2018). A qualitative study of implementation and adaptations to progressive tinnitus management (ptm) delivery. *PLoS One*, 13(5), e0196105.

Van den Berge, M. J., Free, R. H., Arnold, R., De Kleine, E., Hofman, R., Van Dijk, J. M. C., and Van Dijk, P., (2017). Cluster analysis to identify possible subgroups in tinnitus patients. *Frontiers in neurology*, 8, p. 115.

Wakabayashi, S., Oishi, N., Shinden, S., and Ogawa, K., (2020). Factor analysis and evaluation of each item of the tinnitus handicap inventory. *Head & face medicine*, 16, pp. 1–9.