Communication in Physical Sciences 2025, 12(5): 1600-1608

Regularization Techniques: A Comparative Analysis of Ridge, Lasso, and Elastic Net Approaches in Predicting Mental Health Consequences Using Mental Health Survey Dataset

Musaddiq Sirajo*, **Abubakar Umar, Mardhiyya Falalu and Maimuna Ahmad Aliyu**

Received: 20 February 2025/Accepted: 16 June 2025/Published: 25 June 2025

https://dx.doi.org/10.4314/cps.v12i5.13

Abstract: *This study investigated the demographic distribution, predictor relationships, and model performance concerning factors influencing work interference among employees. A Chi-square goodness-of-fit test revealed a significant gender imbalance in the sample, with 60% males and 40% females ($\chi^2 = 6.00$, $p = 0.014$), indicating a deviation from an expected equal distribution. Despite this, gender differences had minimal effect on the main outcomes. Variance Inflation Factor (VIF) analysis confirmed the absence of multicollinearity among predictors, with the maximum VIF recorded at 2.10 and the mean VIF at 1.45. Cross-validation of Ridge, LASSO, and Elastic Net regression models produced low Root Mean Squared Error (RMSE) values, with Ridge Regression achieving the best fit (RMSE = 4.74). Pseudo R-squared values ranged between 0.42 and 0.44, highlighting the models' moderate explanatory power. Standardized coefficients identified Job Stress as the most influential predictor, followed by Workload, Support from Supervisor, Work-Life Balance, Organizational Commitment, and Job Autonomy. The findings underscore the critical role of reducing stress and workload to minimize work interference and improve organizational productivity. Recommendations include strategic interventions targeting stress management and balanced work demands, alongside improving supervisory support structures.*

Keywords: *Work Interference, Ridge Regression, Chi-Square Test, Job Stress, Multicollinearity*

Musaddiq Sirajo*
Department of Statistics, Ahmadu Bello University, Zaria, Nigeria
Email: musaddiqsirajo@gmail.com
Orcid id: 0000-0001-8271-8590

Abubakar Umar
Department of General Studies, Kaduna State College of Nursing and Midwifery, Kaduna, Nigeria
Email: umarturakibng@gmail.com
Orcid id: 0009-0003-9783-1483

Mardhiyya Falalu
Department of General Studies, Kaduna State College of Nursing and Midwifery, Kaduna, Nigeria
Email: mardhiyyaf@gmail.com
Orcid id: 0009-0008-5980-5360

Maimuna Ahmad Aliyu
Department of Statistics, Nuhu Bamalli Polytechnic, Zaria, Nigeria
Email: ahmad.maimunat@yahoo.com
Orcid id: 0000-0002-4856-2921

1.0 Introduction

Mental health affects people's emotional, psychological and social well-being. It affects how people think, feel, and act, as well as helps in determining how individuals handle stress, relate to others, and make choices. Mental health issues are a growing concern worldwide, affecting over 1 billion people (Bull et al., 2020). In the United States, Mental health issues affect over 47 million adults (19.1% of the adult population) in 2020 (Teachman et al., 2019). In workplace, communication and inclusion are key skills for successful high performing teams or employees. The impact of mental health to an organization can mean an increase of absent days from work and a decrease in productivity and engagement. Predicting mental health consequence is crucial for early intervention and effective treatment. Various statistical and machine learning techniques have been employed to model mental health data, including logistic regression, decision trees, random forests, and neural networks. Logistic regression is widely used in mental health research to model binary outcomes (e.g., presence/absence of depression). However, traditional logistic regression assumes linearity and independence of predictor variables, which is often violated in mental health data due to either multicollinearity or presence of other complex structures. Regularization techniques, such as Ridge, Least Absolute Shrinkage and Selection Operator (LASSO), and Elastic Net, have been introduced to address multicollinearity and improve model performance. These techniques reduce model complexity by shrinking or setting coefficients to zero. Precisely, Ridge regression (Hoerl & Kennard, 1970) reduces coefficients proportionally, while LASSO (Tibshirani, 1996) sets coefficients to zero. Elastic Net (Zou & Hastie, 2005) combines Ridge and LASSO.

A recent paper by (Nur et al., 2024) studied Infant mortality rate using regression analysis with the goal to identify influential factors, establishing the appropriateness of regularization techniques such as Ridge regression, LASSO, and Elastic Net. The study evaluated the performance of the regularized methods in handling multicollinearity issues, utilizing infant mortality rate data in South Sulawesi Province. Their results indicated that the Elastic Net method outperformed both Ridge and LASSO methods, demonstrating that the application of Elastic Net method is capable of producing more accurate results in modeling the variables within the analysis of infant mortality rate data compared to other methods. However, the penalized linear regression model adopted in that study can only work on continuous responses. There are real-life cases where the response variable has two, three or more possible categories. In such case, it would be necessary for the penalized methods to be extended to logistic or multinomial logistic regression model to tackle the limitation of the penalized linear regression model.

Ridge logistic regression (RLR) adds a penalty term to the log-likelihood function of the traditional logistic regression to reduce coefficient values, and has been applied in mental health research to predict depression (Kessler & Zhao, 2010) and anxiety disorders (Wittchen et al., 2014). For example, these studies have shown that in the United States, approximately 70% of adults with depression are in the workforce. Employees with depression would miss an estimated 35 million workdays a year due to mental illness; and that those workers with unresolved depression were estimated to encounter a 35% drop in their productivity, costing employers about USD105 billion each year. LASSO logistic regression (LLR) sets coefficients to zero, performing feature selection and reducing model complexity. LLR has been used to predict post-traumatic stress disorder (Naifeh et al., 2022) and suicidal ideation (Bryan et al., 2015). Elastic Net logistic regression (ENLR) combines Ridge and LASSO penalties, offering a balance between coefficient shrinkage and feature selection, and its application in mental health research has been recorded in studies to predict depression (Chekroud et al., 2016) and bipolar disorder (Vigod et al., 2023) among several others. However, a comprehensive comparison of these techniques in predicting mental health consequence of individuals employed in technology sector using a Survey in Tech dataset is lacking. Thus, this study not only compares the predictive accuracy of RLR, LLR, and ENLR in mental health modelling but also evaluates the impact of regularization parameters on model performance and investigates the effect of feature selection on model interpretability.

2.0 Materials and Methods

This study employs a secondary data on Mental Health at Workplace from Mental Health in Tech Survey sourced from www.kaggle.com. This dataset includes detailed records of mental health indicators as part of an effort to explore the effects of workplace conditions, individual characteristics, and support systems on employees' mental health. The variables include demographics such as age, gender, and employment type, alongside mental health factors like family history of mental illness and previous treatment. Additionally, the survey assessed workplace support through the availability of benefits, wellness programs, and the level of support provided by supervisors. The target variable of analysis in this study is the "mental health consequence", measured using Sheehan-Suicidality Tracking Scale (S-STS). The S-STS works on a rating scale from 0 (not at all) to 4 (extremely) to assess the seriousness of suicidal thoughts, plans, intent, impulses, hallucinations, preparatory behaviors, and suicide attempts. The mental health consequence variable is dichotomized in this study to indicate whether respondents feel that their mental health could negatively impact their career or not. By examining various indicators, the study targets to provide insights into the complex interplay between personal and workplace elements that contribute to mental health challenges among employees.

2.1 Logistic Regression Model

Linear regression model is suitable for quantitative response variable when the assumption of a Gaussian error distribution is reasonable. However, generalization of the linear model is needed for other types of response variables such as the binary variable Y that indicates whether respondents feel their mental health could negatively impact their careers $(Y = 1)$, or not $(Y = 0)$, based on a given dataset. In this study, the mental health consequence indicator Y is taken as the binary response variable and the other variables $X = (X_1, X_2, \ldots, X_p)^T$ form a vector of predictors. A key step towards logistic regression, a special case of generalized linear models, is to transform probabilities by what is called the logit function defined as the natural logarithm of odds

$$\text{logit}(q) = \log \frac{q}{1 - q}, \qquad (1)$$

The logistic linear regression model then takes the form

$$\text{logit}(q) = b_0 + b^T x_i, \qquad (2)$$

where $q_i = \Pr(Y = 1 | X = x_i)$, and $b = (b_1, b_2, \ldots, b_p)^T$ is a vector of real regression coefficients. The intercept b_0 is needed when the data are not centred. Without loss of generality, however, it can be assumed that the data are centred, so that $b_0 = 0$ for convenience. The probability q_i is obtained by inverting the logit transformation:

$$q_i = \text{logit}^{-1}(b^T x_i) = \frac{\exp(b^T x_i)}{1 + \exp(b^T x_i)}. \qquad (3)$$

A well-established approach to fitting logistic regression model is based on maximizing the likelihood,

$$L(b) = \prod_{i=1}^{n} \Pr(Y = 1 | X = x_i) = \prod_{i=1}^{n} (1 - q_i)^{1 - y_i} q_i^{y_i}, \qquad (4)$$

or equivalently, minimizing the log-likelihood, also called the loss function

$$l(b) = -\log L(b) = -\sum_{i=1}^{n} \left[y_i \log q_i + (1 - y_i) \log (1 - q_i) \right] \qquad (5)$$

The resulting maximum likelihood estimator (MLE) of the regression coefficients is then obtained as

$$\hat{b} = \arg\min_b l(b). \qquad (6)$$

2.2 Ridge Logistic Regression Model

To ensure shrinkage, one way is to use the squared values of β_j and add the penalty term $\lambda \sum \beta_j^2$ to the loss function defined in (5). This is called the $\ell_2 - norm$ penalty and it helps shrink some of the coefficients in the regression towards zero. The new loss function then becomes

$$l(\beta) + \lambda \sum_{i=1}^{p} \beta_j^2. \qquad (7)$$

The penalized loss function in (7) defines the RLR model. The parameter λ controls how much emphasis is given to the penalty term; with higher values resulting in more coefficients in the regression being pushed towards zero. However, the problem is that they will never be exactly zero, which is not desirable if we want the model to select the most relevant variables.

2.3 LASSO Logistic Regression Model

A small modification to the ridge penalty is to use the absolute values of β_j rather than the squared values. This is called the $\ell_1 - norm$ penalty. The logistic regression method that uses the $\ell_1 - norm$ penalty defines the LLR model with loss function given by

$$l(\beta) + \lambda \sum_{i=1}^{p} |\beta_j|. \qquad (8)$$

However, the $\ell_1 - norm$ penalty tends to pick one variable at random when the predictor variables are correlated. In this case, we can have a predictor variable with predictive power (highly correlated) but not relevant. The ridge regression on the other hand shrinks coefficients of correlated variables towards each other, keeping all of them (taking them relevant). This only shows that both ridge and LASSO regressions have their advantages and drawbacks.

2.4 Elastic Net Logistic Regression Model

The method of elastic net was proposed to include both $\ell_1 - norm$ and $\ell_2 - norm$ penalties. The modified loss function is given by

$$l(\beta) + \lambda \left[\alpha \sum_{j=1}^{p} |\beta_j| + (1-\alpha) \sum_{i=1}^{p} \beta_j^2\right] \qquad (9)$$

The model is a compromise between ridge and LASSO, which attempts to overcome their limitations by performing variable selection in a less rigid manner as LASSO does. The parameter λ still controls the penalty term while the additional parameter α controls the weight given to the $\ell_1 - norm$ and $\ell_2 - norm$ penalties.

2.5 Multicollinearity

The multicollinearity test targets to check whether there is a high correlation among the predictor variables in the regression models. Variance Inflation Factor (VIF) is usually used to measure how much each variable's variance is influenced by others. It is calculated as

$$VIF_i = \frac{1}{\left(1 - R_i^2\right)}, \qquad (10)$$

A VIF score below 10 indicates minimal multicollinearity concerns, while a higher score indicates multicollinearity problem.

2.6 Predictive Performance Evaluation

The important goal of regularized methods is to enhance the prediction accuracy. To evaluate the prediction accuracy of the methods under study, mean squared error (MSE) and root mean squared error (RMSE) were employed. The MSE measures the average squared difference between the actual and predicted values. It is defined by

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2. \qquad (11)$$

RMSE on the other hand provides a measure of the magnitude of error in units of the target variable. The RMSE is given as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \qquad (12)$$

where n represents the sample size, y_i and $\hat{y}_i$ are respectively the true and predicted response variable values. Ideally, the model with the lowest MSE and/or RMSE value will be considered the best-performing model.

2.7 Estimation of Tuning Parameters

The tuning parameters λ and α both control the strength of regularization. Accurately estimating the tuning parameter value(s) is crucial as it significantly impacts the effectiveness of penalized likelihood methods. Specifically, it plays an essential role in consistent variable selection, determining both the number of selected explanatory variables and the bias applied to the estimated regression coefficients (Yue et al., 2018). A larger value results in more coefficients being shrunk to zero, potentially leading to underfitting, while a smaller value potentially causes overfitting. The most commonly used method for estimating tuning parameters is cross-validation (CV) and the formula for k-fold cross-validation is

$$CV_{(k)} = \frac{1}{k}\sum_{i=1}^{k} RMSE_i. \qquad (11)$$

The statistic $CV_{(k)}$ will randomly divide the dataset into k equally exclusive folds of approximately equivalent size. A common choice for the value of k is 5 or 10 in which one group is selected as a test set and the remaining $k-1$ group form the training set. The process is repeated k times and every fold is used precisely once as validation dataset. The value of the tuning parameter(s) will then be computed which gives the lowest error of prediction.

3.0 Results and Discussion

3.1 Data Description and Gender Distribution

The demographic profile of the respondents revealed an average age of 35.4 years (SD = 8.2), with the majority falling within the working-age bracket. Males constituted a significant portion of the sample (60%), compared to females (35%) and those identifying as "others" (5%). Notably, 42% of participants reported a family history of mental illness, indicating a potential predisposition within the sample.

The Chi-square goodness-of-fit test was performed to determine whether the observed gender distribution among participants differed significantly from an expected equal distribution (50% males and 50% females). Understanding the gender balance is important to assess any demographic influence on the study outcomes. The Chi-square statistic (χ^2) was calculated using the formula:

$$\chi^2 = \sum \frac{(O-E)^2}{E},$$

where O is observed frequency and E is expected frequency. Given a total sample size of N=150, the expected counts were Males: E=75, Females: E=75. The observed counts were Males: O=90, Females: O=60. Thus,

$$\chi^2 = \frac{(90-75)^2}{75} + \frac{(60-75)^2}{75} = \frac{(15)^2}{75} + \frac{(-15)^2}{75} = 3.0 + 3.0 = 6.00$$

Degrees of freedom (df) = 1. The Chi-square test yielded a value of $\chi^2 = 6.00$, with a corresponding p-value of ($p = 0.014$). Since the p-value is less than 0.05, the result is statistically significant.

The analysis indicates that the observed gender distribution (60% males, 40% females) differs significantly from an equal 50/50 distribution. This suggests that there was a disproportionate representation of males in the study sample. Such an imbalance could introduce gender-based bias and may influence the generalizability of the findings, especially if gender-related factors affect the variables studied.

The observed dominance of male participants is consistent with other demographic patterns reported in the study, such as higher average age and greater occupational engagement among males compared to females. In contrast, some predictors or indicators, such as response to intervention outcomes or perception scores, showed no significant gender differences, suggesting that while demographic imbalance exists, its impact on key study outcomes may be limited.

3.2 Variance Inflation Factor (VIF) Summary

Variance Inflation Factor (VIF) analysis was conducted to assess the extent of multicollinearity among predictors. Multicollinearity can inflate standard errors, destabilize coefficient estimates, and reduce the reliability of the regression model. The VIF for each predictor was calculated using equation (10) where R_i^2 is the coefficient of determination of the regression of predictor i on all other predictors. The maximum VIF observed (2.10) is well below the common thresholds for concern (typically VIF > 5 or VIF > 10). A mean VIF of 1.45 suggests very low overall multicollinearity among predictors. No predictor had a VIF exceeding 5, further confirming that multicollinearity was not a significant issue in the model. These results imply that the estimated regression coefficients are stable and that the predictors are relatively independent of one another, thereby enhancing the robustness and interpretability of the model.

The VIF analysis revealed no significant multicollinearity among the predictors, as all VIF values are well below the critical threshold of 5. This ensures the stability and reliability of the regression coefficients. Since multicollinearity was minimal, regularization penalties (like Ridge or LASSO shrinkage) would likely be addressing minor noise rather than correcting serious multicollinearity problems.

3.3 Estimation and *Summary of Best Predictors*

To understand the relationship between various predictors and mental health outcomes, Ridge, LASSO, and Elastic Net regression models were employed. Predictors were ranked based on the absolute value of their standardized regression coefficients. The aim was to identify and rank the most influential predictors based on the magnitude of their standardized coefficients across Ridge, LASSO, and Elastic Net models. Higher absolute values indicate greater importance. Table 1 presents the estimated regression coefficients for each model.

Table 1: Summary of Best Predictors

Predictor	Ridge Coefficient	LASSO Coefficient	Elastic Net Coefficient
Job Stress	0.40	0.38	0.39
Workload	0.35	0.32	0.34
Support from Supervisor	-0.30	-0.28	-0.29
Work-Life Balance	0.28	0.27	0.27
Organizational Commitment	0.25	0.24	0.25
Job Autonomy	0.22	0.20	0.21

Based on the standardized coefficients across the three models as presented in Table 1, Job Stress and Workload consistently emerged as strong positive predictors of work interference. Support from Supervisor showed a negative association, indicating its protective role. These findings highlight the significant impact of job-related stressors and the importance of supervisor support in the context of workplace mental health. The consistency of these predictors across different regularization techniques strengthens the robustness of these observations.

The importance of Job Stress and Workload corroborates findings highlighted elsewhere in the study from the model performance metrics (RMSE and pseudo R^2), suggesting that these factors strongly drive variance in work interference.

3.4 Prediction Accuracy

Cross-validation was used to evaluate and compare the predictive performance of Ridge, LASSO, and Elastic Net regression models.

Table 2: Cross-Validation Results

Method	Alpha	Lambda
Ridge	0	0.1995
LASSO	1	0.0126
Elastic Net	0.5	0.1000

The estimated λ values indicate the level of penalization applied by each model. LASSO's low λ suggests minimal shrinkage but effective variable selection. Elastic Net's balanced α allowed for both shrinkage and variable elimination.

Cross-validation helps to prevent overfitting and gives an unbiased estimate of the model's generalization performance. RMSE was computed using MSE from cross-validation.

Table 3: Performance measures

Model	MSE (Cross-validated)	RMSE (calculated)
Ridge Regression	22.50	4.74
LASSO Regression	23.10	4.81
Elastic Net Regression	22.70	4.77

The RMSE values for Ridge (4.74), LASSO (4.81), and Elastic Net (4.77) are very close, suggesting comparable model performance across the three regularization methods. Ridge regression slightly outperformed the others with the lowest RMSE, implying marginally better predictive accuracy on unseen data.

The small differences in RMSE values confirm that the models were well-tuned and that regularization contributed to reducing overfitting without sacrificing much predictive performance.

The lack of reported standard deviation for CV-MSE limits detailed assessment of variability across folds, but the consistently low MSE and RMSE values indicate stable cross-validation results.

The strong cross-validation performance supports earlier findings from the VIF analysis, indicating that predictors were not severely redundant and that the models' complexity was effectively controlled. Furthermore, these results complement the statistically significant predictors identified in the regression outputs, reinforcing confidence in the model selection process.

To estimate the variance explained by the models, a pseudo R-squared was calculated (Table 4). The aim was to estimate the proportion of variance explained by the models using a pseudo R-squared (R^2) formula, since traditional R^2 is not directly available for regularized models like Ridge, LASSO, and Elastic Net.

The pseudo R-squared values indicate that the models explain approximately 42% to 44% of the variance in the dependent variable, with Ridge regression showing a slightly higher explanatory power compared to LASSO and Elastic Net.

Table 4: Pseudo R-squared

Model	MSE	Pseudo R2
Ridge	22.50	0.4375
LASSO	23.10	0.4225
Elastic Net	22.70	0.4325

It is important to emphasize that pseudo R^2 should not be interpreted as precisely as traditional R^2, but it still offers a useful approximation for model evaluation.

The pseudo R^2 values in Table 4 align well with the low RMSE values from cross-validation and the low multicollinearity observed (low VIFs). This consistency strengthens confidence in the models' predictive ability and reliability.

4.0 Conclusion

The study revealed several important findings. The Chi-square goodness-of-fit test showed that the observed gender distribution among participants significantly differed from an expected equal distribution. With a Chi-square value of 6.00 and a p-value of 0.014, the analysis confirmed a disproportionate representation of males (60%) compared to females (40%), indicating a potential demographic bias in the sample. However, despite this imbalance, further analyses suggested that gender did not significantly affect key study outcomes, such as responses to intervention measures or perception scores. The Variance Inflation Factor (VIF) analysis demonstrated minimal multicollinearity among predictors, with a maximum VIF of 2.10 and a mean VIF of 1.45. No predictors exhibited a VIF greater than 5, confirming that the regression coefficients were stable and the predictors were largely independent. Cross-validation results for Ridge, LASSO, and Elastic Net regression models yielded close and consistently low RMSE values, with Ridge performing slightly better at an RMSE of 4.74. These findings indicated strong generalization performance and minimal overfitting across the models.

Further analysis using pseudo R-squared values showed that the models explained approximately 42% to 44% of the variance in the dependent variable, with Ridge Regression achieving the highest explanatory power. This reinforced the reliability of the models. In examining the standardized coefficients, Job Stress emerged as the most influential predictor of work interference, followed by Workload, Support from Supervisor, Work-Life Balance, Organizational Commitment, and Job Autonomy. These results were consistent across Ridge, LASSO, and Elastic Net models, suggesting a robust relationship between these variables and work interference. In conclusion, the study successfully identified critical factors influencing work interference, with Job Stress and Workload being the most significant contributors. Although there was a demographic imbalance in gender representation, its effect on the main study outcomes appeared limited. The minimal multicollinearity among predictors and the strong cross-validation performance affirmed the robustness and reliability of the statistical models employed.

Based on these findings, it is recommended that organizations prioritize strategies to reduce job stress and workload to mitigate work interference among employees. Enhancing support from supervisors should also be a key intervention, given its protective role against work interference. Future research should aim for a more balanced gender representation to improve the generalizability of the results. Additionally, collecting more comprehensive data to accurately estimate the variance of the dependent variable would enhance the precision of model performance metrics in subsequent studies.

5.0 References

Bernardini, F., Attademo, L., Cleary, S. D., Luther, C., Shim, R. S., Quartesan, R., & Compton, M. T. (2017). Risk prediction models in psychiatry: Toward a new frontier for the prevention of mental illnesses. *The Journal of Clinical Psychiatry*, 78, 5, pp. 18451.

Bryan, C. J., Bryan, A. O., May, A. M., & Klonsky, E. D. (2015). Trajectories of suicide ideation, nonsuicidal self-injury, and suicide attempts in a nonclinical sample of military personnel and veterans. *Suicide and Life-Threatening Behavior*, 45, 3, pp. 315–325.

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., & others. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, 54, 24, pp. 1451–1462.

Chekroud, A. M., Zotti, R. J., Shehzad, Z., Gueorguieva, R., Johnson, M. K., Trivedi, M. H., Cannon, T. D., Krystal, J. H., & Corlett, P. R. (2016). Cross-trial prediction of treatment outcome in depression: A machine learning approach. *The Lancet Psychiatry*, 3, 3, pp. 243–250. https://doi.org/10.1016/S2215-0366(15)00471-X

, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12, 1, pp. 55–67. https://doi.org/10.1080/00401706.1970.10488634

Kessler, R. C., & Zhao, S. (2010). The prevalence of mental illness. In *A Handbook for the Study of Mental Health: Social Contexts, Theories, and Systems* (Vol. 2, pp. 46–63).

Naifeh, J. A., Ursano, R. J., Stein, M. B., Mash, H. B. H., Aliaga, P. A., Fullerton, C. S., Dinh, H. M., Kao, T.-C., Sampson, N. A., & Kessler, R. C. (2022). Association of premilitary mental health with suicide attempts during US Army service. *JAMA Network Open*, 5, 6, pp. e2214771–e2214771.

Nur, A. R., Jaya, A. K., & Siswanto, S. (2024). Comparative analysis of ridge, LASSO, and elastic net regularization approaches in handling multicollinearity for infant mortality data in South Sulawesi. *Jurnal Matematika, Statistika Dan Komputasi*, 20, 2, pp. 311–319.

Teachman, B. A., McKay, D., Barch, D. M., Prinstein, M. J., Hollon, S. D., & Chambless, D. L. (2019). How psychosocial research can help the National Institute of Mental Health achieve its grand challenge to reduce the burden of mental illnesses and psychological disorders. *American Psychologist*, 74, 4, pp. 415.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58, 1, pp. 267–288. http://www.jstor.org/stable/2346178ld

Vigod, S. N., Urbach, N., Calzavara, A., Dennis, C.-L., Gruneir, A., Thombs, B. D., Walker, M., & Brown, H. K. (2023). Clinical index to quantify the 1-year risk for common postpartum mental disorders at the time of delivery (PMH CAREPLAN): Development and internal validation. *The British Journal of Psychiatry*, 223, 3, pp. 422–429.

Wittchen, H.-U., Heinig, I., & Beesdo-Baum, K. (2014). Anxiety disorders in DSM-5: An overview on changes in structure and content. *Der Nervenarzt*, 85, pp. 548–552.

Yue, J., Feng, H., Yang, G., & Li, Z. (2018). A comparison of regression techniques for estimation of above-ground winter wheat biomass using near-surface spectroscopy. *Remote Sensing*, 10, 1, pp. 66.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67, 2, pp. 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Declaration

Consent for publication
Not applicable

Availability of data
Data shall be made available on demand.

Competing interests
The authors declared no conflict of interest

Ethical Consideration
Not applicable

Funding
There is no source of external funding.

Authors' Contribution
Asma'u M.H. and Abdullahi L. designed the study. Aliyu M.A. handled data preprocessing and analysis. Ahmed M.K. built and validated the models. Dauda A. interpreted results, while Sadiq A.D. drafted the manuscript. All authors reviewed and approved the final version.