PARAMETER
JURNAL MATEMATIKA, STATISTIKA DAN TERAPANNYA

e-ISSN. 2829-6303
Vol. 04 No. 03, December 2025
Page 459-470

Check for updates

Evaluating the Performance of Ordinal Logistic Regression and XGBoost on Ordinal Classification Datasets

Jasmin Nur Hanifa[1], Rizka Annisa Mingka[2*], Indahwati[3], Pika Silvianti[4]

[1,2,3,4]Department of Statistics and Data Science, SSMI, IPB University
Dramaga Main Road St, Bogor Regency, 16680, West Java, Indonesia

***E-mail Correspondence Author:** *rizkaamingka@gmail.com*

Abstract

Choosing the appropriate classification model is crucial, especially when dealing with data featuring an ordinal dependent variable. This study explores and compares the performance of Ordinal Logistic Regression (OLR) and Ordinal XGBoost in classifying ordinal data using ten datasets obtained from the UCI Machine Learning Repository and Kaggle, which vary in the number of observations and features. Each dataset undergoes multicollinearity detection, an 80% training and 20% testing data split, and class balancing using SMOTE. Model performance is evaluated using metrics such as accuracy, F1-score, AUC, MSE, precision, and recall. The results show that ordinal XGBoost outperforms on datasets with complex structures and a higher number of features, achieving a maximum accuracy of 0.953. In contrast, Ordinal Logistic Regression demonstrates more stable performance on datasets with fewer features or balanced class distributions.

Keywords: *Classification, ordinal logistic regression, ordinal XGBoost.*

: https://doi.org/10.30598/parameterv4i3pp459-470

Research Article** · **Open Acces
Email: jurnalparameter@gmail.com
Homepage: https://ojs3.unpatti.ac.id/index.php/parameter

Submitted: May 2025
Accepted: December 2025

1. INTRODUCTION

In today's rapidly developing digital era, information and data have become strategic assets in various aspects of life. Human activities recorded digitally—from economic transactions and social interactions to educational and healthcare activities—produce massive volumes of data that continue to grow each second. This phenomenon is known as the big data era, which is characterized not only by volume but also by the velocity and variety of the data generated. Alongside advances in computing technology, data processing and analysis have become increasingly crucial, enabling more informative and evidence-based decision-making[1].

One widely used form of data analysis is classification analysis, a process of mapping observations into specific categories based on a number of predictor variables. Classification plays an essential role in many fields, such as disease detection, academic performance prediction, market segmentation, and recommendation systems. A long-established method used in classification is logistic regression, which offers a probabilistic statistical approach for modeling the relationship between independent variables and a categorical dependent variable. Logistic regression is widely favored for its interpretability and efficiency when dealing with data of relatively small to moderate dimensions [1]. However, logistic regression has limitations, particularly when facing complex, large-scale, and non-linear data. To address these challenges, various machine learning methods have been developed, including Extreme Gradient Boosting (XGBoost). XGBoost is a tree-based ensemble algorithm that efficiently implements the boosting technique with regularization to prevent overfitting and features parallelization capabilities, making it exceptionally fast [2]. Numerous studies have shown that XGBoost excels in various data science competitions and real-world applications such as fraud detection, genomic analysis, and machine failure prediction [3].

In the context of an ordinal dependent variable—i.e., a categorical variable with an inherent order, such as satisfaction levels or academic performance—conventional logistic regression is often insufficient. One of the primary issues is the proportional odds assumption, which requires that the effect of a predictor is the same across all category thresholds—an assumption rarely met in real-world scenarios. Therefore, ordinal logistic regression (proportional odds model) was developed to better handle the structure of ordinal data [4].

Ordinal logistic regression has been widely used in various studies, including the analysis of student exam results [5], customer satisfaction [6], and disease severity levels [7]. Nonetheless, the model's performance can deteriorate when confronted with high-dimensional data, outliers, or complex interactions among variables. This is where machine learning approaches like XGBoost offer advantages. XGBoost has been adapted to handle ordinal data, either by modifying the loss function or through data transformation approaches.

On the other hand, challenges in ordinal classification lie not only in the selection of algorithms but also in the evaluation of model performance. Evaluation metrics such as accuracy are insufficient, as they do not account for the inherent order among classes. Therefore, metrics such as F1-score, Area Under the Curve (AUC), and Mean Absolute Error (MAE) are more appropriate for assessing the performance of ordinal classification models. Additionally, class imbalance is a critical issue, as models tend to be biased toward the majority class. To address this, various approaches such as oversampling, probability calibration, and threshold adjustment have been developed to improve the predictive quality for minority classes.

A previous study by Chu and Keerthi [8] developed Support Vector Machines for ordinal classification, an idea that inspired similar adaptations in boosting algorithms. Additionally, Fu et al. [9] demonstrated that applying XGBoost to ordinal data could significantly improve classification accuracy, particularly in medical and educational data processing.

In recent years, ordinal boosting approaches have gained increasing attention, involving modifications to the structure of boosting algorithms such as XGBoost to align with the ordinal nature of the data. For example, Zhu et al. [10] developed an ordinal classification method based on XGBoost that incorporates class order and demonstrated superior performance in predicting the severity of traffic accident injuries. They also emphasized the importance of understanding data characteristics, including class imbalance and feature complexity, in selecting an appropriate classification approach.

This study aims to explore and compare the effectiveness of ordinal logistic regression and XGBoost in classifying data with ordinal response variables. While this study does not introduce a novel methodological approach, its contribution lies in broadening the understanding of how these two popular methods perform across different contexts. It also serves as a reference for practitioners and researchers in selecting appropriate approaches for ordinal classification problems, while emphasizing the importance of considering data characteristics in analytical model selection.

2. RESEARCH METHODS

This study adopts a quantitative approach with a model performance comparison method. The quantitative approach emphasizes the analysis of numerical data and the application of statistical techniques to objectively measure and compare model performance [11]. The study is comparative in nature, aiming to assess and compare the prediction accuracy of two modeling methods across various data characteristics.

The methods applied in this research are Ordinal Logistic Regression and Ordinal XGBoost, used to compare model performance in predicting ordinal response variables. The analysis was conducted on ten datasets obtained from the UCI Machine Learning Repository and Kaggle, comprising predictor variables with both numerical and categorical data types. Each dataset contains an ordinal response variable with a clearly ordered structure and exhibits variation in both the number of observations and features, thus allowing for a comprehensive and in-depth model performance evaluation. The general characteristics of the datasets used in this study are summarized in Table 1.

Table 1. Summary of the Datasets Used

Dataset No.	Observations	Features	Feature Type(s)	Target Description
1.	395	33	Numeric: 9 features, Categorical: 17 features	*Student Final Grade* 0 = Low 1 = Medium 2 = High
2.	2126	22	Numeric: 21 features, Categorical: 1 feature	*Uterine Health* 0 = Normal 1 = Suspect 2 = Pathological

Dataset No.	Observations	Features	Feature Type(s)	Target Description
3.	300	17	Numeric: 13 features, Categorical: 4 features	*University Ranking* 0 = Top 1 = Middle 2 = Low
4.	400	4	Numeric: 1 feature, Categorical: 3 features	*Recruitment Criteria* 0 = Unlikely 1 = Somewhat Likely 2 = Very Likely
5.	12330	18	Numeric: 10 features, Categorical: 5 features, Boolean: 2 features	*Average Page Economic Value* 0 = Low 1 = Medium 2 = High
6.	320	9	Numeric: 8 features, Categorical: 1 feature	*Soybean Yield* 0 = Low 1 = Medium 2 = High
7.	299	10	Numeric: 6 features, Categorical: 4 features	*Age of Heart Failure Patients* 0 = Young 1 = Elderly 2 = Senior
8.	276	11	Numeric: 10 features, Categorical: 1 feature	*Cirrhosis Patient Stages* 0 = Stage 1 1 = Stage 2 2 = Stage 3 3 = Stage 4
9.	702	11	Categorical: 10 features	*Mental Health Condition* 0 = Poor 1 = Fair 2 = Good 3 = Excellent
10.	4424	12	Numeric: 6 features, Categorical: 6 features	*Student Academic Success* 0 = Dropout 1 = Enrolled 2 = Graduate

The initial stage in this study involved data preprocessing, including handling missing values, transforming variables into more appropriate data types, and selecting predictors with low correlation for model building. For example, in datasets 4, 5, and 6, some variables were eliminated from the model due to having correlation coefficients above 0.7. The datasets were then split into training and testing sets with a ratio of 80:20, maintaining stratification to preserve class proportions. Due to class imbalance, class balancing was performed on the training data using the Synthetic Minority Over-Sampling Technique (SMOTE).

2.1. Synthetic Minority Oversampling (SMOTE)

In this study, the Synthetic Minority Over-Sampling Technique (SMOTE) was used to address class imbalance by generating synthetic data points for the minority class. Class imbalance occurs when one class significantly outnumbers others, which can bias the model training and reduce classification performance.

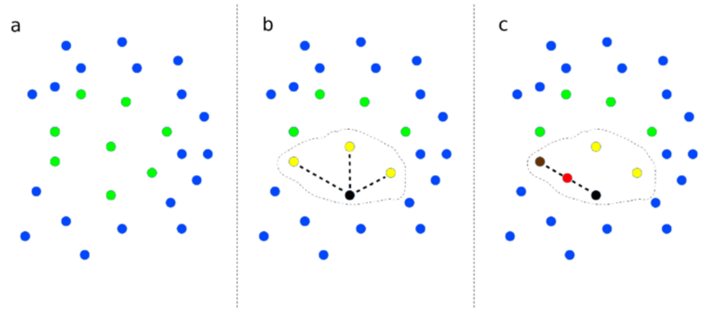


Figure 1. **How SMOTE Works [12]**

Figure 1 illustrates how SMOTE works to address data imbalance. In panel (a), minority class data points (green dots) are fewer than the majority class (blue dots). A minority point (black) is selected as shown in panel (b), and its nearest neighbors (yellow dots) are identified. Panel (c) depicts the interpolation process between the black point and one of its neighbors (brown dot), producing a new synthetic point (red) located between them. This process is repeated to enrich the minority class distribution.

2.2. Ordinal Logistic Regression

Logistic regression is a statistical method used to model the relationship between one or more predictor variables and a categorical response variable. In its binary form, logistic regression estimates the probability of an event occurring using a logit function of a linear combination of predictors. When the target variable has more than two unordered categories, multinomial logistic regression is used. However, if the categories are ordered, such as "low", "medium", and "high", ordinal logistic regression is more appropriate.

Ordinal logistic regression models the cumulative probability up to a certain category, under the proportional odds assumption, meaning the effect of predictors on the cumulative logit is constant across category thresholds. This allows ordinal information to be used more efficiently than in classical classification methods.

The cumulative probability model is as follows:

$$P(Y \leq j|x) = \pi_1(x) + \cdots + \pi_j(x) \tag{1}$$

For $j = 1, \ldots, J$, the cumulative logits are defined (Agresti, 2013) as:

$$logit[P(Y \leq j\,|x)] \quad = \quad log\,\frac{P(Y \leq j\,|x)}{1 - P(Y \leq j\,|x)} \tag{2}$$

$$logit[P(Y \leq j\,|x)] \quad = \quad log\,\frac{\pi_1(x) + \cdots + \pi_j(x)}{\pi_{j+1}(x) + \cdots + \pi_J(x)} \tag{3}$$

Which simplifies ordinal classification into several cumulative binary logistic models. Thus, the ordinal logistic regression model can be defined as:

$$logit[P(Y \leq j|x)] \,=\, \alpha_j + \beta^T x \tag{4}$$

For $j = 1, \ldots, J - 1$.

In this study, the ordinal logistic regression model was built using the stepwise selection algorithm (forward-backward) based on the Akaike Information Criterion (AIC) while addressing potential multicollinearity. This procedure aims to select the most relevant predictor variables that are not highly correlated, resulting in a simpler model without sacrificing predictive accuracy. The selected variables were then modeled using ordinal logistic regression through several steps, including parameter estimation using the maximum likelihood method **[13]**. The best model for each dataset was selected based on the lowest AIC value.

Simultaneous significance testing of parameters was conducted to evaluate whether all predictors jointly contributed significantly to the model **[14]**. Partial significance tests were used to identify which individual variables had significant effects on the response variable **[15]**. Model fit evaluation (Goodness of Fit) was then conducted to ensure that the built model adequately captured the relationship between predictors and the response variable **[16]**. The final performance evaluation included accuracy, sensitivity, specificity, AUC, and F1-score.

2.3. XGBoost Ordinal

XGBoost is a supervised learning algorithm that combines multiple weak base classifiers—each with relatively low individual accuracy—into a single, more accurate and stable predictive model. The main objective of this approach is to reduce the overall prediction error of the model **[2]**. XGBoost is constructed based on the boosting principle, where a sequence of decision trees of the CART (Classification and Regression Tree) type are trained iteratively, with each successive tree aiming to correct the errors made by its predecessors.

One of the key strengths of XGBoost lies in its ability to handle complex datasets without resulting in an overly complicated model (*overfitting*). This is achieved by expanding the loss function to its second-order derivative using Taylor series approximation, and by incorporating a regularization term to control model complexity. The objective function in XGBoost consists of two main components and can be formulated as:

$$P(x) = L(x) + E(x) \tag{5}$$

Where $L(x)$ is the loss function, which measures how well the model fits the training data, and $E(x)$ is the regularization term that quantifies the complexity of the model, such as the number of leaf nodes and the magnitude of the weights within the trees.

If the model consists of j decision trees, then the prediction for the iii-th observation is computed as the sum of the outputs from all trees:

$$y_i = \sum_{j=1}^{J} f_j(x_i) \tag{6}$$

where f_j is the j-th tree as a predictive function, $f_j \in F$ where F is the set of all possible trees that can be used as base learners, and J is the total number of trees in the model.

Thus, the final objective function of XGBoost can be expressed as:

$$P(x) = \sum_{i=1}^{n} L(y_i\,, \hat{y}_i) + \sum_{j=1}^{J} E(f_j) \tag{7}$$

This indicates that the overall objective function optimized by the model is the summation of the loss function evaluated across the entire training dataset and a complexity penalty term applied to each tree within the model. The inclusion of the regularization term serves to mitigate overfitting by controlling model complexity, thereby enhancing the generalization performance of XGBoost on previously unseen test data.

The XGBoost model for ordinal classification was developed as a comparison to ordinal logistic regression. It is an enhancement of the standard XGBoost algorithm with modifications to handle ordinal response variables. XGBoost for ordinal classification was implemented using an ensemble learning approach by constructing a series of decision trees sequentially, where each new tree corrects the errors of the previous one. In the ordinal context, XGBoost was modified by using a custom loss function that considers category ordering and penalizes misclassification based on ordinal distances [17].

To ensure a fair comparison, the evaluation of the ordinal XGBoost model used the same performance metrics: accuracy, sensitivity, specificity, AUC, and F1-score.

Each dataset was used to evaluate which model consistently provided the best performance for ordinal response variables. Stability assessment was also conducted to test the generalization capability of each model across varying data characteristics. All analytical procedures were repeated across each dataset to ensure consistency and reliability in performance evaluation.

3. RESULTS AND DISCUSSION

3.1. Data Exploration

The datasets used in this study originate from various domains such as education, health, and social sciences. Ten datasets were selected based on the presence of ordinal target variables. The number of observations in each dataset varies from hundreds to thousands, as does the number of features.

Preprocessing was carried out on each dataset, including handling missing values, transforming categorical variables into numeric format via encoding, and balancing class distribution—particularly for the target variable. Figure 2 illustrates the class proportion of a target variable reflecting stages of liver cirrhosis in Dataset 8, before and after balancing using the SMOTE algorithm.

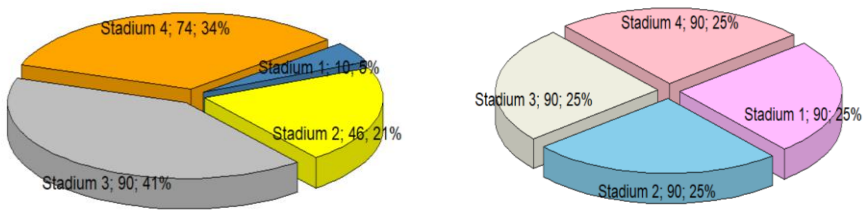


Figure 2. Class Distribution Before and After SMOTE in Dataset 8

An important assumption that must be met is the absence of multicollinearity, or high correlation among predictor variables. Multicollinearity checks were included as part of the preprocessing phase to ensure that the variables used in the analysis are mutually independent. All datasets underwent identical preprocessing steps to prepare them for analysis using two approaches: ordinal logistic regression and ordinal XGBoost, with the aim of comparing their performance.

3.2. Comparison of Ordinal Logistic Regression and XGBoost Ordinal

Implemented on each dataset. Ordinal logistic regression was applied to identify predictor variables that significantly influence the target variable in each dataset. Variable selection was performed using a stepwise algorithm to obtain the optimal set of predictors, with the selection process based on the Akaike Information Criterion (AIC).

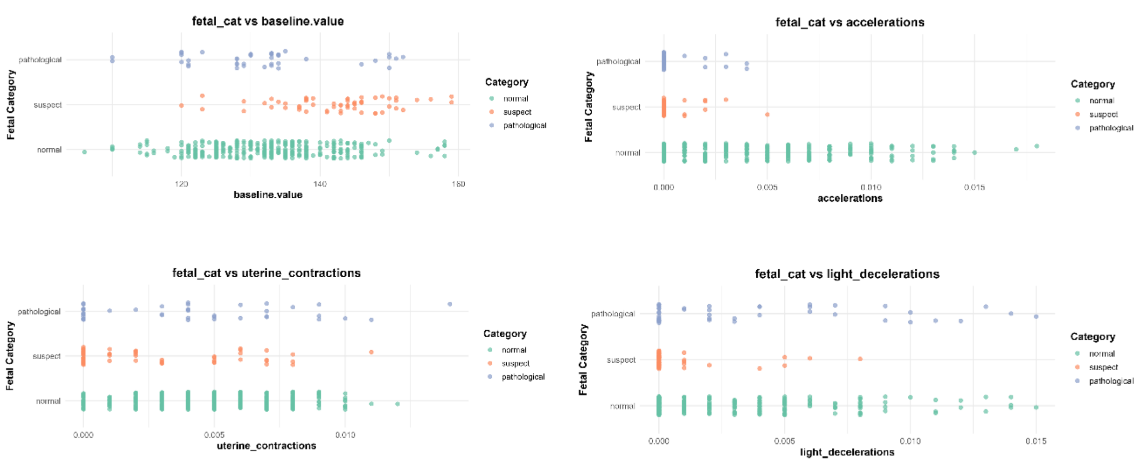
Model evaluation was conducted on the test data using accuracy and confusion matrix metrics. For some datasets, OLR was able to interpretatively identify important predictors, although model accuracy varied depending on the dataset's complexity. Ordinal classification using a modified version of standard XGBoost incorporated the order of the target categories. The XGBoost model in this study was trained with default parameters that were subsequently fine-tuned. All features were used without manual selection, as the algorithm inherently handles feature weighting and importance. The ordinal XGBoost model applied label transformation and multiclass evaluation.

Each dataset analyzed using both OLR and XO models was evaluated using accuracy, F1-score, AUC, MSE, precision, and recall. A summary of performance metrics is presented in Table 2.

Table 2. Model Performance Comparison

Dataset No.	Accuracy		F1 Score		Recall		AUC		MSE		Precision	
	OLR	XO	OLR	XO	OLR	XO	OLR	XO	OLR	XO	OLR	XO
1	0.595	0.518	0.555	0.488	0.548	0.488	0.861	0.780	0.518	0.671	0.575	0.515
2	0.852	0.943	0.705	0.383	0.571	0.889	0.951	0.990	0.204	1.29	0.729	0.904
3	0.750	0.683	0.731	0.677	0.722	0.668	0.868	0.852	0.300	1.443	0.741	0.687
4	0.713	0.487	0.618	0.387	0.430	0.406	0.743	0.479	0.475	1.163	0.642	0.394
5	0.775	0.706	0.454	0.455	0.340	0.454	0.672	0.651	0.505	1.203	0.447	0.459
6	0.547	0.75	0.592	0.700	0.383	0.699	0.644	0.863	0.594	0.375	0.422	0.700
7	0.517	0.6	0.532	0.567	0.353	0.562	0.649	0.686	0.583	0.483	0.393	0.592
8	0.536	0.482	0.525	0.366	0.552	0.355	0.652	0.659	0.625	0.946	0.394	0.336
9	0.475	0.397	0.483	0.357	0.479	0.355	0.738	0.666	0.652	1.035	0.369	0.354
10	0.566	0.583	0.565	0.489	0.569	0.501	0.622	0.695	0.43	1.054	0.410	0.490

Based on the table, it is evident that the performance of both models heavily depends on the characteristics of each dataset. In general, Ordinal XGBoost outperformed in terms of accuracy and AUC on most datasets. The highest accuracy was obtained on Dataset 2 (uterine health levels), where XO achieved 0.943 accuracy, significantly surpassing OLR at 0.852. This dataset contains 2,126 observations and 22 features. Thus, XO correctly predicted around 2,026 instances, while OLR correctly predicted approximately 1,811. Interestingly, despite the high accuracy, the model exhibited a relatively high MSE (OLR = 0.204, XO = 1.29).



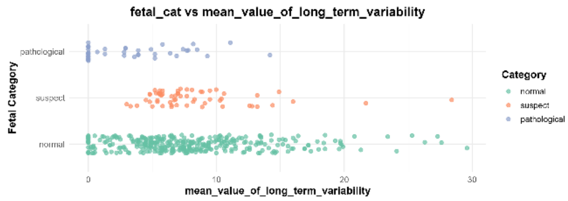


Figure 3. **Scatter Plot of Test Data Profile in Dataset 2**

Figure 3 illustrates that Dataset 2 shows clear class separation between "normal," "suspect," and "pathological," facilitating more accurate model predictions. Although the MSE was high, the clear interclass separation allowed the model to distinguish between classes effectively.

Ordinal XGBoost also demonstrated superior performance on Dataset 6, with an accuracy of 0.75 compared to OLR's 0.55. This reflects a substantial difference, with the dataset having relatively balanced target distribution and one of the highest standard deviations (1.4152) among all datasets. The model's performance was supported by a low MSE of 0.375 and a high precision of 0.7, indicating good predictive consistency—especially in data with limited complexity or class bias. Generally, there is a positive relationship among the metrics: higher accuracy tends to correlate with lower MSE, as predictions are closer to the actual labels. High precision further implies fewer false positives. The consistency of high accuracy, low MSE, and high precision in Dataset 6 reinforces that XO fits the data characteristics better than OLR. Conversely, low accuracy was found in Datasets 4, 8, and 9, particularly for XO with respective accuracies of 0.487, 0.482, and 0.397. These results are consistent with high MSE values. Moreover, precision was low, especially for the majority classes, suggesting the model struggled with distinguishing between categories. This may be due to the high variability in these datasets, making it more difficult for the model to capture patterns—especially when class distributions are imbalanced.

Besides accuracy, the F1-score metric was used to provide a more balanced evaluation between precision and recall, especially in imbalanced-class scenarios even after SMOTE balancing. In ordinal classification, F1-score reflects the model's ability to classify data correctly across all classes without bias toward the majority. From **Table 2**, Ordinal XGBoost generally yielded higher F1-scores compared to OLR in most datasets. The highest F1-score was achieved by XO on Dataset 2 at 0.9062, reflecting high balance between precision and recall for uterine health classification. However, OLR yielded better F1-scores on Dataset 3 and Dataset 4 (0.7307 and 0.6180, respectively), suggesting that in certain datasets, OLR still provides consistent and balanced classification. On the other hand, XO significantly outperformed OLR on Datasets 5, 6, and 7, indicating its superiority in handling complex feature interactions. To evaluate the model's capability to differentiate between classes, the Area Under the Curve (AUC) metric was used in a multi-class setting. A higher AUC value indicates better probabilistic predictions in identifying the correct target class.

The highest AUC was recorded by XO on Dataset 2 (0.9920), reflecting its excellent classification ability for uterine health levels. Below is the ROC/AUC visualization for each class of the target variable in Dataset 2.

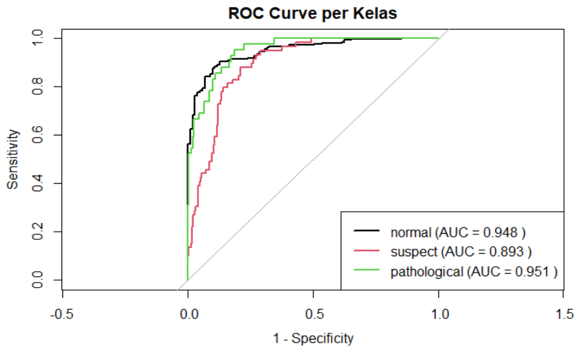


Figure 4. ROC/AUC Curves for Ordinal Logistic Regression on Dataset 2

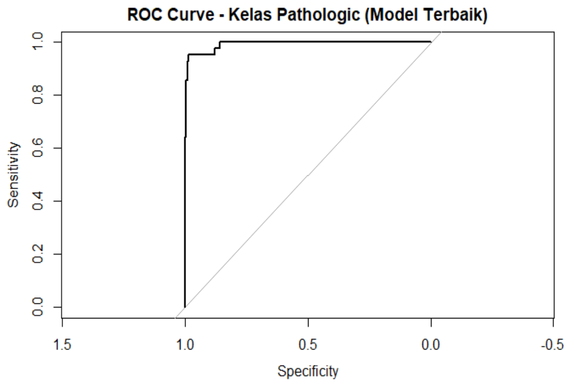


Figure 5. ROC/AUC Curves for XGBoost on Dataset 2

Additionally, XO achieved an AUC of 0.8629 on Dataset 6, significantly outperforming OLR's 0.6444, indicating a notable difference in probabilistic prediction quality. However, OLR also showed competitive performance in some datasets. For instance, in Dataset 1, OLR achieved an AUC of 0.8615, higher than XO's 0.7645, as well as in Datasets 3, 4, and 9, with AUCs exceeding 0.7. This indicates that OLR remains effective in producing good probabilistic predictions, especially when the ordinal structure is strong and the number of features is manageable.

Overall, based on evaluation results across ten datasets, both OLR and XO exhibit unique strengths depending on data characteristics. OLR tends to provide better accuracy in datasets with well-defined ordinal structures and fewer features. For example, although OLR performed poorly on Dataset 4, it outperformed XO on Dataset 2. On the other hand, XO demonstrated more consistent performance in F1-score and AUC, showing its advantage in handling complex data and non-linear relationships. These findings suggest that selecting an ordinal classification method should consider the analysis objective, data structure, and whether interpretability or predictive accuracy is prioritized.

The findings of this study are further supported by the work of Zhu et al. [10], who emphasized the importance of considering the characteristics of ordinal data and class imbalance in selecting appropriate classification methods. They demonstrated that misclassifying ordinal data as nominal can lead to decreased prediction accuracy, and that a clearly defined ordinal structure within the data significantly influences model performance—consistent with the observations regarding ordinal logistic regression in this study. Moreover, Zhu et al. [10] developed an ordinal classification approach based on XGBoost and showed that the modified XGBoost model, tailored to handle class order, outperformed traditional methods, particularly in datasets with imbalanced class distributions and nonlinear feature relationships. These findings reinforce the results of this study, which indicate that XGBoost is superior in handling complex data and exhibits more consistent performance across evaluation metrics such as F1-score and AUC, making it a more flexible choice for probabilistic prediction in ordinal classification tasks.

4. CONCLUSION

Based on the evaluation of ten ordinal datasets using two modeling approaches—ordinal logistic regression and ordinal XGBoost—several key findings were obtained. Ordinal XGBoost exhibited superior performance, particularly in datasets with complex patterns and a large number of features. This was reflected in the highest accuracy and AUC scores achieved by XO in Datasets 2 and 6. In Dataset 2, the model also recorded the highest F1-score of 0.9062, indicating its capability in distinguishing between multiple classes. On the other hand, ordinal logistic regression remained competitive and even outperformed XO in specific datasets. This model recorded higher accuracy in six out of ten datasets, namely Datasets 1, 3, 4, 5, 8, and 9, with relatively low MSE values. OLR performed better in datasets with more linear structures or fewer features, such as Datasets 1, 3, and 4. Overall, the performance of each model is highly dependent on the dataset's characteristics. For example, in datasets with high feature variability like Dataset 2, XO was able to capture the data patterns more effectively. Conversely, in datasets with fewer features or more balanced class distributions, OLR provided more stable performance. Therefore, in practical applications of ordinal classification modeling, it is advisable to evaluate more than one approach, taking into account data structure and selecting the model that offers the most consistent and accurate performance.

REFERENCES

[1] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed., Hoboken: Wiley, 2013.

[2] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.

[3] D. Nielsen, *Tree Boosting with XGBoost: Why Does XGBoost Win "Every" Machine Learning Competition?* [Master's thesis, NTNU], 2016.

[4] A. Agresti, *Analysis of Ordinal Categorical Data*, 2nd ed., Hoboken: Wiley, 2010.

[5] Q. Cutts, E. Cutts, S. Draper, P. O'Donnell, and P. Saffrey, "Manipulating mindset to positively influence introductory programming performance," *ACM SIGCSE Bulletin*, vol. 41, no. 3, pp. 431–435, 2010.

[6] Y. Liu, J. Wu, Y. Chen, and H. Li, "Ordinal logistic regression model for customer satisfaction analysis," *Journal of Applied Statistics*, vol. 42, no. 5, pp. 1014–1023, 2015.

[7] D. G. Kleinbaum and M. Klein, *Logistic Regression: A Self-Learning Text*, 3rd ed., New York: Springer, 2010.

[8] W. Chu and S. S. Keerthi, "Support vector ordinal regression," *Neural Computation*, vol. 19, no. 3, pp. 792–815, 2007.

[9] Y. Fu, T. M. Hospedales, T. Xiang, and S. Gong, "Learning multimodal latent attributes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 10, pp. 2532–2545, 2018.

[10] S. Zhu, K. Wang, and C. Li, "Crash injury severity prediction using an ordinal classification machine learning approach," *International Journal of Environmental Research and Public Health*, vol. 18, pp. 1-21, 2021

[11] J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed. Thousand Oaks, CA: SAGE Publications, 2014.

[12] H.-S. Ryu and J. Lee, "Detection of political hate speech in Korean language," *Language Resources and Evaluation*, pp. 1–32, 2024.

[13] M. N. Shofiyah and M. S. Chamid, "Faktor-Faktor yang Mempengaruhi Tingkat Stres Siswa Saat Pembelajaran Daring Menggunakan Metode Regresi Logistik Ordinal," *Jurnal Sains dan Seni ITS*, vol. 11, no. 2, pp. 23-28, 2022.

[14] D. U. Setyawati, B. D. Korida, and B. R. A. Febrilia, "Analisis Regresi Logistik Ordinal Faktor-Faktor yang Mempengaruhi IPK Mahasiswa," *Jurnal Varian*, vol. 3, no. 2, pp. 65-72, 2020.

[15] T. Pentury, S. N. Aulele, and R. Wattimena, "Analisis Regresi Logistik Ordinal (Studi kasus: Akreditasi SMA di Kota Ambon)," *Barekeng : Jurnal Ilmu Matematika dan Terapan*, vol. 10, no. 1, pp. 55–60, Mar. 2016.

[16] D. Holmes and L. M. Haines, *Applied Multivariate Data Analysis*. Berlin: Springer, 2013.

[17] D. A. Al-Qudah, A. M. Al-Zoubi, A. I. Cristea, J. J. Merelo-Guervós, P. A. Castillo, and H. Faris, "Prediction of sentiment polarity in restaurant reviews using an ordinal regression approach based on evolutionary XGBoost," *PeerJ Computer Science*, vol. 11, p. e2370, 2025.