УДК 004.8, 368.8

Kuan Madiyarov
(Russian Federation)
Scientific supervisor – A. I. Pestunov,
Ph.D. physics and mathematics Sciences, Associate Professor
Novosibirsk State University of Economics and Management

COMPARISON OF ACCURACY IN MACHINE LEARNING MODELS FOR INSURANCE RISK PREDICTION IN POULTRY FARMING

Machine learning models for predicting insurance risks in poultry farming are analysed in this study. Methods include Random Forest for structured data and MobileNetV2 for image diagnostics, as well as CatBoost, XGBoost, and LightGBM for boosting-based predictions. The results indicate that CatBoost achieves the highest accuracy in structured data analysis, while MobileNetV2 effectively detects poultry diseases. The developed model enhances insurance risk assessment, improving policy optimization.

Several studies have demonstrated the effectiveness of machine learning models in predicting insurance risks. Rocheteau and Kim highlighted the use of deep learning for medical image analysis, showing that transfer learning approaches significantly improve diagnostic accuracy. The use of machine learning in insurance claim prediction was analysed by Abdulkadir and Fernando, with an emphasis on the role of deep learning and gradient boosting in risk assessment.

Predictive models are used by insurers to assess and manage risks in poultry farming, including extreme weather, infectious diseases, and poor farm oversight. Traditional actuarial methods are now enhanced with machine learning, enabling more precise risk estimation.

Financial losses stem from high mortality due to weather, bacterial infections, and poor environmental conditions like ammonia and CO_2 levels. Predictive models help insurers quantify risks and tailor policies effectively. This study compares multiple machine learning models for forecasting poultry insurance claims, integrating structured data and image-based diagnostics. The author developed and trained the predictive model to address poultry insurance challenges.

This study employs machine learning for structured insurance data and poultry disease classification. The models include **Random Forest**

(RF) for structured data, **CatBoost** for categorical data, **XGBoost** for efficient parallel boosting, and **LightGBM** for large-scale datasets. For disease detection, **MobileNetV2** applies transfer learning with pre-trained TensorFlow Hub weights. Model performance was assessed using **MAPE** and **R^2 Score** for structured data, measuring prediction accuracy and variance explanation. For poultry disease classification, **Accuracy** and **F1-Score** evaluated model effectiveness.

The performance metrics displayed in Table 1 were derived from an evaluation of different machine learning models for predicting insurance claims.

Table 1. Performance metrics of machine learning models for structured data

№	Model	MAPE	R^2 Score
1.	Random Forest	7.12	0.9913
2.	CatBoost	2.98	0.9970
3.	XGBoost	5.65	0.9806
4.	LightGBM	4.89	0.9841

As seen in Table 1, CatBoost outperforms other models, achieving the lowest MAPE and highest R^2 score. This suggests that gradient boosting models are particularly effective for structured insurance data prediction. Random Forest, while accurate, does not match the precision of boosting-based techniques.

For poultry disease classification, MobileNetV2 achieved an accuracy of 89.2% and an F1-score of 0.87. This demonstrates that deep learning techniques can be successfully applied to automate disease detection in poultry farming. However, further improvements, such as dataset expansion and hyperparameter tuning, could enhance the classification performance.

Conclusion

A comparative study of machine learning models for poultry insurance risk prediction is presented, with an emphasis on structured data and image-based disease classification. The results confirm that CatBoost outperforms other models in structured data prediction, achieving the lowest MAPE and highest R^2 score, making it the most effective model for structured insurance data. Traditional ensemble methods like Random Forest remain useful but demonstrate lower predictive power compared to boosting-based models.

For image-based disease classification, MobileNetV2 achieved high accuracy, demonstrating its potential for automated poultry disease detection.