

Animal Health Prediction Using Hybrid KNN Based Vector Neighbor Classification Model: A Machine Learning Approach

¹J. Rathi and ²A. Sumathi

¹Department of Computer Science, KPR College of Arts Science and Research, Arasur, Coimbatore, Tamil Nadu, India

²Department of Information Technology, KPR College of Arts Science and Research, Arasur, Coimbatore, Tamil Nadu, India

Article history

Received: 22-04-2025

Revised: 14-07-2025

Accepted: 24-07-2025

Corresponding Author:

J. Rathi

Department of Computer Science,
KPR College of Arts Science and
Research, Arasur, Coimbatore,
Tamil Nadu, India

Email:

sharveshrathivadivel@gmail.com

Abstract: Animal condition disease prediction is a crucial task in veterinary medicine, and accurate prediction can significantly improve animal health and reduce economic losses. This paper proposed a novel hybrid classification model that combines an Improved Fuzzy-based feature selection process with a Hybrid KNN-based Vector Neighbor Classification model to enhance animal condition disease prediction. The proposed model aims to improve the analysis of both normal and diseased categories by leveraging the strengths of fuzzy logic and machine learning. The Improved Fuzzy feature selection process utilizes a fuzzy-based algorithm to select relevant features from a high-dimensional dataset, reducing dimensionality and improving model performance. The Hybrid KNN-based Vector Neighbor Classification model integrates the benefits of K-Nearest Neighbors (KNN) and Vector Neighbor Classification (VNC) to classify animal conditions into normal and diseased categories. The proposed hybrid model is evaluated on a dataset of animal health records, and the results demonstrate improved accuracy and robustness compared to existing classification models. The experimental results of the proposed HKNN-VNC method are evaluated on a real-world Animal condition classification dataset, achieving impressive performance metrics: Accuracy: 98.62%, Precision: 98.62%, Recall: 100%, and F1-score: 99.30%. These results demonstrate the effectiveness of the HKNN-VNC approach in accurately classifying animal conditions.

Keywords: Data Classification, Animal Disease, KNN, Random Forest, Hybrid Model

Introduction

The well-being of animals is closely tied to sustainable agriculture, livestock, and industry, which has far-reaching implications for global health and security. Accurate and timely diagnosis of animal diseases is essential for maintaining farm profitability and animal welfare. While traditional diagnostic methods have been the cornerstone of veterinary medicine, they have limitations that are becoming increasingly apparent (Shailaja *et al.*, 2018). One of the major challenges in animal health is the lack of access to prompt diagnostic approaches, particularly in remote or underserved areas where experienced veterinarians and well-equipped laboratories may not be available. Additionally, animals often only exhibit clinical symptoms when diseases have progressed to an advanced stage, making early intervention more difficult and emphasizing the need for innovative diagnostic solutions (Ezanno *et al.*, 2021).

Data mining for animal health is a crucial area of research that harnesses the power of data analytics and machine learning to improve the health and wellbeing of animals, which in turn has a direct impact on human health and wellbeing. By analyzing large datasets related to animal health, including medical records, genomic information, and environmental factors, researchers and practitioners can identify early warning signs of disease outbreaks, detect patterns and trends in animal health, inform evidence-based decision making, improve animal welfare and conservation, and ultimately enhance human health and wellbeing. This exciting field has the potential to drive positive change for both animals and humans alike, and its applications are vast, ranging from detecting zoonotic diseases and antimicrobial resistance to supporting conservation efforts and improving food safety, making it an essential area of study in today's interconnected world (Zhang *et al.*, 2021; Morota *et al.*, 2018; Chen *et al.*, 2017; Gholami & Javidan, 2017).

Machine learning disease classification is revolutionizing pet healthcare for dogs, cats, and other animals. For dogs, machine learning can detect diseases such as arthritis, diabetes, and cancer. For cats, it can identify conditions like feline leukemia, kidney disease, and asthma. Additionally, machine learning can also classify diseases in other pets like birds (e.g., avian influenza), horses (e.g., colic, laminitis), and small mammals (e.g., respiratory issues in rabbits). By analyzing medical data, images, and genomic information, machine learning algorithms can accurately diagnose and predict diseases, enabling veterinarians to provide personalized treatment and improve animal health outcomes (Ortiz-Pelaez and Pfeiffer, 2008; Li *et al.*, 2021; Buza *et al.*, 2014).

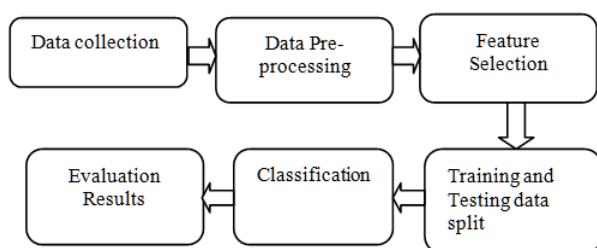


Fig. 1: Workflow of machine learning classification process

Figure (1) describes a machine learning workflow for animal health classification starts with Data Collection, gathering datasets related to animal health, followed by Data Pre-processing, which involves cleaning, transforming, and preparing the data for analysis. Next is Feature Selection, where relevant features are chosen to represent the data, and then the data is split into Training and Testing sets. The machine learning model is trained on the Training set, which enables it to learn patterns and relationships within the data. The Testing set is then used to assess the model's performance, evaluating its accuracy and effectiveness in making predictions or classifications (De La Rocque *et al.*, 2008; Bök and Micucci, 2024; Niu *et al.*, 2021).

Machine learning algorithms have revolutionized the field of zoonotic disease research, enabling researchers to uncover hidden patterns and relationships in vast amounts of data from diverse sources, including animal and human health records, environmental conditions, and genetic information. By harnessing the power of machine learning, researchers can now better understand and forecast zoonotic diseases, which are illnesses that can transmit from animals to humans. This paper aims to utilize machine learning techniques to predict outbreaks of zoonotic diseases and classify animal diseases based on their symptoms using traditional ML algorithms. While significant progress has been made in predicting chronic human diseases like heart disease and diabetes using machine learning, relatively little development has occurred in the area of animal disease prediction. However, in the wake of COVID-19, it is crucial that to prioritize monitoring and predicting animal diseases that

have the potential to become zoonotic, highlighting the urgent need for advancements in this field (Rahman *et al.*, 2020).

Existing animal disease prediction models often struggle with accurately classifying complex health conditions due to limitations in feature selection and classification algorithms. Specifically, many models rely on simplistic feature selection methods and lack robustness in handling uncertain or noisy data, leading to reduced accuracy and reliability in animal health predictions. The HKNN-VNC model addresses this research gap by integrating Improved Fuzzy-based feature selection with Hybrid KNN-based classification, providing a more comprehensive and robust approach to animal disease prediction.

This paper proposes a novel hybrid classification model that enhances animal condition disease prediction by integrating an Improved Fuzzy-based feature selection process with a Hybrid KNN-based Vector Neighbor Classification model. This combination enables a more comprehensive analysis of both normal and diseased categories, leading to improved accuracy and robustness in animal health predictions.

Related Works

Yadav & Jadhav (2019), highlighted the significance of medical image classification in clinical treatment and teaching tasks. However, traditional methods have reached their performance limit, requiring extensive time and effort for feature extraction and selection. Deep neural networks, particularly Convolutional Neural Networks (CNNs), have shown promise in various classification tasks. Nevertheless, medical image datasets are challenging to obtain due to the need for expert labeling. To address this, the authors investigated the application of CNN-based algorithms on a chest X-ray dataset for pneumonia classification. Through experimental evaluation, they compared three techniques to improve performance.

Mhlanga (2020), conducted a dynamic analysis of a model to inform the design of effective prevention and intervention strategies for zoonotic diseases. The study's objective was to minimize the number of infected individuals through health promotional programs in affected communities. To achieve this, the researcher performed a sensitivity analysis of the model parameters and employed the Pontryagin's Maximum Principle to determine the optimal level of control measures. The resulting optimality system was solved numerically, revealing that high-efficacy health promotional programs are crucial for successfully preventing and controlling zoonotic diseases among villagers.

Liu & Sun (2021), introduced a novel model to tackle the diagnosis of small round blue cell tumors (SRBCTs), a challenging task due to high noise and diversity in small samples. To address this, their model utilizes a

masking mechanism to focus on relevant areas in pathological images and incorporates a length-aware loss function to enhance robustness to feature diversity. The framework was evaluated on a demanding SRBCTs dataset, which is difficult to classify even for experienced pathologists, demonstrating the model's potential to improve diagnostic accuracy.

Tamburis *et al.* (2021), highlighted the growing significance of Veterinary Informatics, which is driving the development of integrated veterinary information management systems (VIMS) for animal data capture, storage, analysis, and retrieval. The authors developed a decision tree algorithm using the University Veterinary Hospital's database at Federico II University of Naples, with the goal of creating a predictive model for accurately identifying neoplastic diseases and zoonoses in cats and dogs in the Campania Region.

Afshari Safavi (2022) investigated the potential of machine learning algorithms to predict the occurrence of LSDV infection based on meteorological and geological factors. The study employed the Extra Trees Classifier algorithm to identify the most relevant predictive features, selecting key attributes from meteorological, animal population density, dominant land cover, and elevation data. The results showed that certain machine learning techniques achieved high accuracy (up to 97%) in forecasting LSDV occurrence in unseen test data.

Rehman *et al.* (2023), addressed the growing concern of rising animal diseases, many of which have the potential to become zoonotic and impact both animals and humans. While machine learning has been widely applied to human disease detection, its application to animal diseases has been limited. This study bridges this gap by introducing a novel approach using machine learning techniques to classify and predict the spread of animal diseases, which is crucial for mitigating the impact of zoonotic diseases on both human and animal populations. The authors' contribution is significant, as zoonotic diseases can have far-reaching consequences for public health and animal welfare.

Nadar *et al.* (2023), highlighted the crucial role of veterinary doctors in ensuring the health and well-being of animals throughout their lifespan. However, in rural India, accessing veterinary services is challenging due to remote locations, forcing farmers to travel long distances for animal treatment. This hinders animal husbandry in these areas, where farm animals are a vital source of nutrition and income. A website connecting veterinarians with livestock owners could be a valuable solution. Timely access to veterinary services is essential to address issues like pest infestation, disease, and toxicity, which can lead to poor productivity, animal death, and significant economic losses for owners and the nation. Making veterinary services available nationwide, either online or by deploying practitioners to rural areas, is critical to addressing these challenges.

Methods

The proposed research methodology performs the Animal Condition Classification using data cleaning, feature selection and Hybrid KNN based Vector Neighbor Classification for Animal Disease Prediction algorithm process is derived in this section. The overall process flow is depicted in Figure (2), providing a visual representation of the methodology.

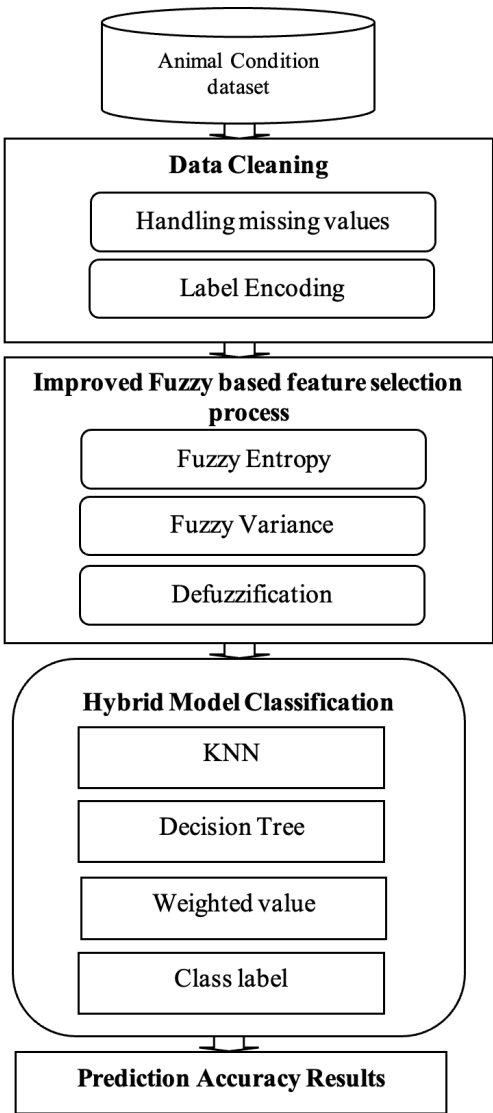


Fig. 2: HKNN-VNC process flow

Data Cleaning

Data cleaning is a crucial step in the classification process, as it directly impacts the accuracy and reliability of the predictive models. The “Animal Condition Classification Dataset” from Kaggle offers a unique data challenge in animal health assessment. In the dataset, data cleaning is necessary to remove errors, inconsistencies, and noise that can lead to biased or incorrect classifications. By cleaning the data, the paper

ensures that the symptoms are consistently represented, missing values are handled, which enables the classification algorithms to learn patterns and relationships in the data more effectively. If the data is not cleaned, the models may learn from errors and produce suboptimal results, leading to incorrect classifications of animal conditions, which can have serious consequences for animal health and welfare. Therefore, data cleaning is an essential step in preparing the dataset for classification and ensuring the development of robust and reliable predictive models.

The dataset undergoes thorough cleaning, starting with handling missing values by replacing them with “None” or “Unknown”, followed by categorization of symptoms into groups (e.g., respiratory, gastrointestinal, neurological) to reduce dimensionality using Imputation by Constant Value method. Categorical variables (e.g., Yes/No, Dog/Cat) are encoded into numerical variables using one-hot encoding or label encoding, and numerical data (e.g., temperature, weight loss) is scaled to a standard range (e.g., 0-1) to prevent feature dominance. Relevant symptoms and features are selected and redundant or irrelevant ones are removed. The dataset is then balanced to ensure equal representation of different animal conditions resulting in a clean, organized, and balanced dataset ready for analysis and modeling.

Improved Fuzzy Based Feature Selection Process

The improved fuzzy-based feature selection process for the Animal Condition Classification Dataset begins with fuzzification, converting crisp symptom values into fuzzy sets using membership functions, followed by calculating the importance of each symptom using fuzzy measures such as fuzzy entropy and fuzzy variance, then ranking symptoms based on their importance scores, selecting the top-ranked symptoms as the most relevant features, and finally, defuzzification, converting the selected fuzzy features back into crisp values, thereby handling uncertainty and ambiguity in symptom values and ensuring a more robust and accurate feature selection for animal condition classification, capturing subtle relationships between symptoms and improving model performance.

Let dataset D_i be the i^{th} symptom value, where $i=1,2,\dots,n$. Define m fuzzy sets Fz_j , where $j=1,2,\dots,a$, with membership functions μ_{ij} , where μ_{ij} represents the degree of membership of D_i in Fz_j is shown in Eq. (1):

$$\mu_{ij} = \frac{1}{1 + \left(\frac{D_i - c_j}{w_j}\right)^2} \quad (1)$$

Where c_j is the center and w_j is the width of the j^{th} fuzzy set.

Calculate the importance score Fsc_j for each fuzzy set Fz_j using the fuzzy entropy measure is shown in Eq. (2):

$$Fsc_j = \frac{1}{n} \sum_{i=1}^n \mu_{ij}^2 \quad (2)$$

Rank the fuzzy sets based on their importance scores Fsc_j is shown in Eq. (3):

$$RF_j = \frac{Fsc_j}{\sum_{k=1}^m Fsc_k} \quad (3)$$

Feature selection is performed by selecting the top-ranked fuzzy sets as the most relevant features, leveraging the fuzzification model to prioritize features based on their fuzzy membership degrees, and retaining only the most informative and discriminative features while discarding redundant or irrelevant ones, thereby optimizing the feature subset for enhanced performance and accuracy in subsequent analysis or modeling tasks. Select the top-ranked fuzzy sets as the most relevant feature is shown in Eq. (4):

$$SF_j = \begin{cases} 1, & RF_j \geq th \\ 0, & otherwise \end{cases} \quad (4)$$

Where th is a threshold value.

Defuzzification is the process of converting the selected fuzzy features back into crisp values, involving the transformation of fuzzy membership degrees into precise numerical values, typically using techniques such as Centroid, Weighted Average, or Maximum Membership Principle, to determine the most representative crisp value for each fuzzy feature, thereby facilitating the conversion of fuzzy outcomes into tangible, actionable insights or decisions, and enabling the seamless integration of fuzzy logic with conventional crisp datasets and models is shown in Eq. (5):

$$D_i^* = \frac{\sum_{j=1}^m SF_j \mu_{ij} D_i}{\sum_{j=1}^m SF_j \mu_{ij}} \quad (5)$$

The resulting D_i^* are the selected feature values.

The proposed improved fuzzy-based feature selection approach employs fuzzy logic to select the most relevant features from the animal condition disease dataset, leveraging fuzzification to handle uncertainty and ambiguity in feature importance. The method prioritizes features based on their fuzzy membership degrees, retaining only the most informative and discriminative ones. Defuzzification then converts the selected fuzzy features into crisp values, enabling actionable insights. This robust fuzzy-based feature selection process improves disease diagnosis and prediction accuracy, facilitating effective animal condition monitoring and treatment.

Hybrid KNN Based Vector Neighbor Classification

This paper introduces a novel Hybrid KNN-based Vector Neighbour classification method, which builds upon the strengths of both KNN and Random Forest classification models. By combining the distance-based similarities of KNN with the decision tree-based ensemble learning of Random Forest, this new approach leverages the advantages of both techniques to improve classification accuracy and robustness. The proposed method extends the traditional KNN algorithm by

incorporating Random Forest's ability to handle high-dimensional data and complex interactions, while also utilizing KNN's simplicity and interpretability. The resulting Hybrid KNN-based Vector Neighbour classification model demonstrates improved performance and versatility, making it a promising approaches for various classification tasks.

The hybrid model combines the KNN and Random Forest predictions using a weighting parameter, where the KNN component calculates the distance using the Euclidean distance metric, assigns similarity weights using the inverse distance, predicts the class label using the weighted sum of the nearest neighbors' labels, and the Random Forest component trains decision trees using the CART algorithm, predicts the class label using each tree, and combines the predictions using a weighted sum, and the hybrid model combines the predictions using the weighting parameter, alpha, to produce the final class label, leveraging the strengths of both algorithms to improve the accuracy and robustness of the predictions.

Let assume x be the input feature vector, y be the target variable, K be the number of nearest neighbors, N be the number of decision trees, F be the number of features used in each decision tree and C be the number of classes.

The first process of calculate the distance between x and each training sample x_i using Euclidean distance of KNN component is defined as shown in Eq. (6):

$$d(x, x_i) = \sqrt{\sum_{j=1}^F (x_j - x_{ij})^2} \quad (6)$$

After that, select the K nearest neighbors based on the distance metric.

Calculate the similarity weights for each nearest neighbor is shown in Eq. (7):

$$we_i = \frac{1}{d(x, x_i)} \quad (7)$$

Predict the class label using the weighted sum of the nearest neighbors' labels is shown in Eq. (8):

$$Y_{knn_k} = \sum_{i=1}^k we_i \times y_i \quad (8)$$

The Random Forest Component using CART constructs multiple decision trees on random subsets of the training data, splitting nodes using the Gini impurity criterion and recursive partitioning, and prunes trees using cost-complexity pruning. Each decision tree predicts the class label for the new sample, and the predictions are combined using a weighted sum, where each tree's prediction is weighted equally. The Random Forest algorithm repeats this process for a particularamount of iterations, combining the predictions from all iterations using the weighted sum. This approach leverages the strengths of multiple decision trees to improve classification accuracy and robustness, while also reducing overfitting and variance.

Train a decision tree F_i on a random subset of the training data. Predict the class label using each decision

tree is shown in Eq. (9):

$$P_i = F_i(x) \quad (9)$$

Combine the predictions using a weighted sum is shown in Eq. (10 & 11):

$$Y_{RF} = \sum_{i=1}^N P_i \times w_i \quad (10)$$

Where,

$$w_i = \frac{1}{N} \quad (11)$$

To combine the both models of KNN and Random Forest predictions using a weighting parameter of hybrid model is shown in Eq. (12):

$$y_{hybrid} = \alpha \times Y_{knn_k} + (1 - \alpha) \times Y_{RF} \quad (12)$$

The weighting parameter α in the hybrid model is a crucial component that determines the contribution of each model's predictions to the final output. To ensure reproducibility and transparency, α should be optimized using a systematic approach, such as cross-validation on the training dataset, to find the optimal value that maximizes the hybrid model's performance. The optimal α value can be determined by iterating over a range of values (e.g., 0 to 1 with increments of 0.1) and evaluating the model's performance on a validation set. The α value yields the finest performance can then be selected and reported, ensuring that the hybrid model's results are reliable and reproducible.

Algorithm: Hybrid KNN based Vector Neighbor Classification (HKNN-VNC)

Input: Input dataset D , Disease Class C , Features F

Output: Prediction (P) result

Preparation:

1. Data collection
2. Data cleaning
3. Improved Fuzzy-based feature selection process
4. Hybrid Classification

Steps:

While (Number of features)

Dt_cleaning $\leftarrow$ data cleaning(D)

For $F = 1$ to fn **do** // where fn is number of features in dataset

$F_s \leftarrow$ Improved Fuzzy-based feature selection(D).

$y_{knn} \leftarrow$ KNN(F_s) // KNN prediction result

$y_{RF} \leftarrow$ RF(F_s) // RF prediction result

$y_{hybrid} \leftarrow y_{knn} + y_{RF}$ // Hybrid model

$P \leftarrow$ Prediction Disease label (class 0 or 1)

Predicted Disease $\leftarrow P(C)$

Show the Disease class label.

End For

End While

The proposed Hybrid Model combines the KNN and Random Forest predictions using a weighting parameter (α) to produce the final prediction using Eq. 12, where KNN calculates distances and similarity weights, and Random Forest uses decision trees to predict class labels, and the hybrid model leverages the strengths of both to improve accuracy and robustness.

Results and Discussion

The hybrid classification approach with the improved fuzzy-based feature selection technique was used for result evaluation. To compare the proposed HKNN-VNC method with existing Naive Bayes (NB), KNN and Random Forest (RF) algorithms was performed. Real-world Animal Condition Classification dataset are attainable in this experiment. The performance of the proposed HKNN-VNC technique was evaluated using precision, recall, and accuracy metrics. The results are presented in Table (1) and Figure (3), which compare the precision and recall measures. Additionally, the classification accuracy and F1-score comparisons are summarized in Table (2) and Figure (4).

Table (1) shows the evaluation metrics of precision and recall measures for various classification methods, showing that all methods achieve high precision scores above 97%, with HKNN-VNC performing the best at 98.62%. In terms of recall, KNN, RF, and HKNN-VNC attain perfect scores of 100%, while NB lags behind with a recall score of 78.17%, indicating that HKNN-VNC and RF are the top performers in balancing precision and recall.

Table 1: Evaluation metrics of precision and recall measures

Methods	Precision	Recall
NB	97.42	78.17
KNN	97	100
RF	97.97	100
HKNN-VNC	98.62	100

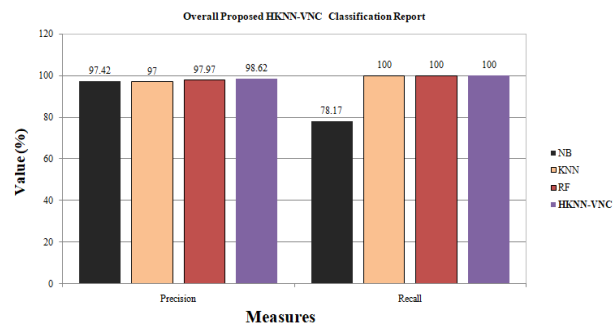


Fig. 3: Performance Analysis of Precision and Recall measures

Fig. (3), illustrates the performance analysis of various machine learning methods reveals high precision and recall rates, with Random Forest (RF) and HKNN-VNC achieving the highest scores. Specifically, RF attained a precision of 97.97% and perfect recall of 100%, while HKNN-VNC demonstrated the highest

precision at 98.62% with equally flawless recall of 100%, indicating the effectiveness of these methods in classification tasks.

Table (2) refers the evaluation metrics of accuracy and F1-score measures, revealing that HKNN-VNC achieves the highest accuracy at 98.62% and F1-score at 99.30%, closely followed by KNN and RF with accuracy scores of 97.41% and 93.0%, and F1-scores of 98.68% and 97.98%, respectively. In contrast, NB exhibits relatively lower performance with an accuracy of 88.0% and F1-score of 76.24%, indicating that HKNN-VNC outperforms the other methods in terms of overall accuracy and F1-score.

Table 2: Evaluation metrics of Accuracy and F1-score measures

Methods	Accuracy	F1-score
NB	88.0	76.24
KNN	97.41	98.68
RF	93.0	97.98
HKNN-VNC	98.62	99.30

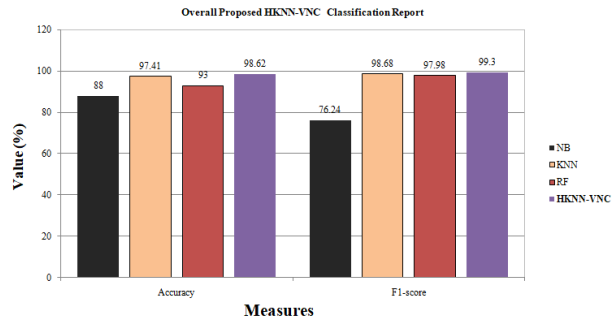


Fig. 4: Performance Analysis of Accuracy and F1-score measures

Fig. (4), illustrates the performance analysis reveals that proposed HKNN-VNC outperforms other methods with the highest accuracy of 98.62% and an F1-score of 99.30%, indicating excellent balance between precision and recall. KNN and RF also demonstrate performance, with KNN achieving an accuracy of 97.41% and F1-score of 98.68%, and RF attaining an accuracy of 93.0% and F1-score of 97.98%. In contrast, Naive Bayes (NB) shows relatively lower performance with an accuracy of 88.0% and F1-score of 76.24%, suggesting its limitations in this classification task.

The HKNN-VNC model improves over standalone KNN or RF models by leveraging a hybrid approach that combines the strengths of KNN with vector neighbor classification, enabling more accurate and robust classification of animal health conditions. Unlike standalone KNN, which may struggle with noisy or uncertain data, HKNN-VNC's hybrid approach enhances its ability to handle complex data. Additionally, compared to RF, HKNN-VNC's use of vector neighbor classification allows for more nuanced and context-specific classification, resulting in improved accuracy and F1-score performance.

Conclusion

This paper introduces a novel Hybrid KNN-Based Vector Neighbor Classification (HKNN-VNC) method that significantly enhances in classification accuracy and robustness when applied to animal condition datasets. By integrating the strengths of both KNN and Random Forest algorithms, this innovative approach leverages the benefits of distance-based classification and ensemble learning to improve the prediction of animal health outcomes. The Hybrid model exhibits improved performance in handling high-dimensional data and noise making it a promising solution for animal health monitoring and disease diagnosis. The combination of KNN's simplicity and Random Forest's robustness results in a powerful and versatile classification model, capable of handling complex animal condition datasets. This research contributes to the advancement of animal health informatics, offering a new tool for researchers and practitioners to improve animal welfare and reduce the economic impact of animal diseases. Future work will advance Deep Neural Networks for animal condition datasets using CNNs and multi-modal learning, improving animal health predictions.

Acknowledgment

The authors acknowledge the support and cooperation rendered by all the members directly and indirectly.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

J. Rathi: Responsible for content creation, design, data collection, and ensuring the novelty of the work.

A. Sumathi: Refined the content, verified the novelty, and ensured the logical flow of the work.

Ethics

This article is written adhering to all the ethical standards that are necessary,

References

Afshari Safavi, E. (2022). Assessing machine learning techniques in forecasting lumpy skin disease occurrence based on meteorological and geospatial features. *Tropical Animal Health and Production*, 54(1), 1–13.
<https://doi.org/10.1007/s11250-022-03073-2>
 Bök, P.-B., & Micucci, D. (2024). The future of human and animal digital health platforms. *Journal of Reliable Intelligent Environments*, 10(3), 245–256.
<https://doi.org/10.1007/s40860-024-00232-0>

Buza, T., Arick, M., Wang, H., & Peterson, D. G. (2014). Computational prediction of disease microRNAs in domestic animals. *BMC Research Notes*, 7(1).
<https://doi.org/10.1186/1756-0500-7-403>
 Chen, M., Hao, Y., Hwang, K., Wang, L., & Wang, L. (2017). Disease Prediction by Machine Learning Over Big Data From Healthcare Communities. *IEEE Access*, 5, 8869–8879.
<https://doi.org/10.1109/access.2017.2694446>
 de La Rocque S., S., Rioux, J. A., & Slingenbergh, J. (2008). Efectos del cambio climático en los sistemas zoonosarios y consecuencias en materia de vigilancia y control. *Revue Scientifique et Technique de l'OIE*, 27(2), 339–354.
<https://doi.org/10.20506/rst.27.2.1807>
 Ezanno, P., Picault, S., Beaunée, G., Bailly, X., Muñoz, F., Duboz, R., Monod, H., & Guégan, J.-F. (2021). Research perspectives on animal health in the era of artificial intelligence. *Veterinary Research*, 52(1). <https://doi.org/10.1186/s13567-021-00902-4>
 Gholami, M., & Javidan, R. (2017). An intelligent model for prediction of brucellosis. *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, 570–576.
<https://doi.org/10.1109/kbei.2017.8324865>
 Li, X., Zhang, Z., Liang, B., Ye, F., & Gong, W. (2021). A review: antimicrobial resistance data mining models and prediction methods study for pathogenic bacteria. *The Journal of Antibiotics*, 74(12), 838–849.
<https://doi.org/10.1038/s41429-021-00471-w>
 Liu, Y., & Sun, S. (2021). SagaNet: A small sample gated network for pediatric cancer diagnosis. *Proc. 38th Int. Conf. Mach. Learn.*, 6947–6956.
 Mhlana, A. (2020). Assessing the Impact of Optimal Health Education Programs on the Control of Zoonotic Diseases. *Computational and Mathematical Methods in Medicine*, 2020, 1–15.
<https://doi.org/10.1155/2020/6584323>
 Morota, G., Ventura, R. V., Silva, F. F., Koyama, M., & Fernando, S. C. (2018). Big Data Analytics and Precision Animal Agriculture Symposium: Machine Learning and Data Mining Advance Predictive Big Data Analysis in Precision Animal Agriculture1. *Journal of Animal Science*, 96(4), 1540–1550.
<https://doi.org/10.1093/jas/sky014>
 Nadar, A., Sane, A., Muga, G., Masih, E., & Rukhande, S. (2023). Animal Healthcare and Farm Animal Disease Prediction Using Machine Learning. *2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE)*, 1–6.
<https://doi.org/10.1109/icnte56631.2023.10146635>
 Niu, B., Liang, R., Zhou, G., Zhang, Q., Su, Q., Qu, X., & Chen, Q. (2021). Prediction for Global Peste des Petits Ruminants Outbreaks Based on a Combination of Random Forest Algorithms and Meteorological Data. *Frontiers in Veterinary Science*, 7, 1–13.
<https://doi.org/10.3389/fvets.2020.570829>

- Ortiz-Pelaez, A., & Pfeiffer, D. U. (2008). Use of data mining techniques to investigate disease risk classification as a proxy for compromised biosecurity of cattle herds in Wales. *BMC Veterinary Research*, 4(v), 1–13.
<https://doi.org/10.1186/1746-6148-4-24>
- Rahman, Md. T., Sobur, Md. A., Islam, Md. S., Ievy, S., Hossain, Md. J., El Zowalaty, M. E., Rahman, A. T., & Ashour, H. M. (2020). Zoonotic Diseases: Etiology, Impact, and Control. *Microorganisms*, 8(9), 1405.
<https://doi.org/10.3390/microorganisms8091405>
- Rehman, S., Rathore, B., & Lal, R. (2023). Animal Disease Prediction using Machine Learning Techniques. *International Journal for Research in Applied Science and Engineering Technology*, 11(6), 1441–1456.
<https://doi.org/10.22214/ijraset.2023.53544>
- Shailaja, K., Seetharamulu, B., & Jabbar, M. A. (2018). Machine Learning in Healthcare: A Review. *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 910–914.
<https://doi.org/10.1109/iceca.2018.8474918>
- Tamburisi, O., Masciari, E., Esposito, C., & Fatone, G. (2021). Data Mining for Animal Health to Improve Human Quality of Life: Insights from a University Veterinary Hospital. *Proceedings of the 10th International Conference on Data Science, Technology and Applications*, 694–701.
<https://doi.org/10.5220/0010517200002993>
- Yadav, S. S., & Jadhav, S. M. (2019). Deep convolutional neural network based medical image classification for disease diagnosis. *Journal of Big Data*, 6(1), 1–18. <https://doi.org/10.1186/s40537-019-0276-2>
- Zhang, S., Su, Q., & Chen, Q. (2021). Application of Machine Learning in Animal Disease Analysis and Prediction. *Current Bioinformatics*, 16(7), 972–982.
<https://doi.org/10.2174/1574893615999200728195613>