Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3, 1665-1674
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i3.5666

Analysis of obesity statuses on rats using blood test parameters: A feasibility study

Ali Berkan URAL[1*], Evren KOÇ[2]
[1]Department of Electrical Electronics Engineering, Circuit and Systems/Biomedical, Kafkas University, Kars, Turkey, berkan.ural@kafkas.edu.tr (A.B.U.)
[2]Department of Bio Engineering, Kafkas University, Kars, Turkey, evrenkoc@kafkas.edu.tr (E.K.)

Abstract: Although obesity has become a significant issue in our era, various diagnostic systems and kits are being developed for its early detection. In addition to numerous studies in the literature, the early diagnosis of obesity has generally been conducted on human participants. This study introduces an innovative feasibility approach by developing an AI and machine learning-based obesity prediction and interpretation application using blood test parameters obtained from rodent subjects. In the experimental phase, with a publicly available dataset, 10 obese and 10 normal (control group) rats were selected, ensuring a meaningful and appropriate sample size for veterinary research. Specific blood test parameters of these subjects were analyzed. These parameters were compiled into a data form and subjected to machine learning-based prediction and interpretation. The machine learning methods used in this study included k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), and Random Forest algorithms. Performance analyses were conducted for each method based on the obtained results. The highest accuracy rate was achieved with the Random Forest algorithm, reaching approximately 97.4%. The accuracy rates obtained with other models were also significant, demonstrating that the study has the potential to be further developed and applied to other living beings, including humans.
Keywords: *Computer Aided Systems, Feature Selection, Machine Learning Models, Obesity, Rats.*

1. Introduction

Obesity is a growing global health concern, affecting both humans and animals [1, 2]. In veterinary medicine, early detection of obesity in animals, particularly in laboratory settings, is crucial for understanding metabolic disorders and developing preventive strategies [3, 4]. Traditional diagnostic methods often rely on physical examinations and body weight measurements, which may not provide comprehensive insights into obesity-related metabolic changes.

With the advancements in artificial intelligence (AI) and machine learning (ML), data-driven approaches have gained prominence in medical and veterinary research [5, 6]. These methods enable precise analysis of biological markers, offering improved diagnostic accuracy. In this study, we propose a machine learning-based approach for detecting obesity in rats using blood test parameters. By leveraging different ML algorithms, including k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), and Random Forest, we aim to develop an effective predictive model. According to the literature analysis, there have been many studies among human participants for obesity detection via machine learning models but there have been a limited series of study about the animal version of the related study, so this study could be called unique/novel for developing an AI based ML decision support system for obesity discrimination among rats.

The proposed methodology involves collecting blood samples from obese and normal (control group) rats, extracting key metabolic parameters, and utilizing machine learning models to classify

History: Received: 17 January 2025; Revised: 18 March 2025; Accepted: 19 March 2025; Published: 22 March 2025
* Correspondence: berkan.ural@kafkas.edu.tr

obesity status. Performance evaluations of these models will be conducted to determine the most accurate and reliable method for obesity detection. The findings of this study can contribute to veterinary research and may serve as a foundation for future applications in both animal and human obesity studies.

2. Material and Methods

This study was conducted as a retrospective analysis using blood test data from the past five years, obtained from a publicly available dataset of Wistar rats [7]. Since the data was openly accessible, no ethical approvals were required. Prior to utilization, the dataset was anonymized to ensure privacy. Following this, essential preprocessing steps were carried out, including noise removal, normalization, and extraction of relevant features. Finally, the dataset was split into designated training and test sets, following a 70:30 ratio.

2.1. Dataset Used in the Study

Detailed representation of our dataset was given below with featuring totally 10 obese and 10 normal data for participants with different overweight statuses via 11 different features and these data were compared with the non-obese healthy reference value range metrics. The features used in this study was explained below.

- Cholesterol: Cholesterol is a waxy, fat-like substance found in the blood and produced by the liver. It is essential for building cell membranes, producing hormones (such as estrogen and testosterone), and synthesizing vitamin D.
- Triglycerides: Triglycerides are often a type of lipid, separate from cholesterol.
- LDL Cholesterol: Low-Density Lipoprotein (LDL) Cholesterol, often called "bad cholesterol," is a type of lipoprotein that carries cholesterol from the liver to the cells. If too much LDL cholesterol circulates in the blood, it can deposit in the walls of arteries, leading to plaque buildup, narrowing of the arteries (atherosclerosis), and an increased risk of heart disease and stroke.
- HDL Cholesterol: High-Density Lipoprotein (HDL) Cholesterol, often called "good cholesterol," is a type of lipoprotein that helps remove excess cholesterol from the bloodstream by transporting it to the liver for excretion. Higher levels of HDL cholesterol are associated with a lower risk of heart disease and stroke because it helps prevent plaque buildup in the arteries.
- VLDL cholosterol: Very-Low-Density Lipoprotein (VLDL) cholesterol is a type of lipoprotein responsible for transporting triglycerides from the liver to various tissues throughout the body. It plays a crucial role in lipid metabolism, acting as a carrier of fats that provide energy or are stored for future use. VLDL is produced by the liver and contains a high proportion of triglycerides relative to its protein content, making it one of the primary contributors to circulating fat in the bloodstream.
- Glucose: Glucose is the body's primary energy source, with higher amounts stored as glycogen in the muscles and liver.
- Weight: Body weight is a fundamental parameter in biological and medical research involving rats, serving as a critical indicator of overall health, metabolic status, and physiological changes. In scientific studies, monitoring rat weight is essential for evaluating the effects of diet, disease progression, drug efficacy, and environmental factors on health. Changes in weight can signal underlying conditions such as obesity, malnutrition, or metabolic disorders, making it a valuable metric in experimental research.
- Total Antioxidant Status (TAS): Total Antioxidant Status (TAS) is a crucial parameter used to assess the overall antioxidant capacity of biological systems, reflecting the balance between oxidative stress and antioxidant defense mechanisms. Antioxidants play a vital role in protecting cells from oxidative damage caused by reactive oxygen species (ROS), which are linked to various

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1665-1674, 2024
DOI: 10.55214/25768484.v9i3.5666

diseases, including cardiovascular disorders, neurodegenerative conditions, and metabolic syndromes.

TAS provides a comprehensive evaluation of both enzymatic and non-enzymatic antioxidants present in the body, offering a more holistic approach than measuring individual antioxidants separately. It is commonly used in medical and veterinary research to study oxidative stress-related conditions and to evaluate the effects of diet, lifestyle, and therapeutic interventions on antioxidant defenses.

In laboratory studies, TAS is often measured in blood plasma, serum, or tissue samples using spectrophotometric or electrochemical methods. By monitoring TAS levels, researchers can better understand the body's ability to counteract oxidative stress and develop strategies for disease prevention and treatment. The assessment of TAS is particularly relevant in studies related to obesity, inflammation, and metabolic disorders, where oxidative stress plays a significant role.

- Total Oxidant Status (TOS): Total Oxidant Status (TOS) is a key biomarker used to evaluate the overall oxidative burden in biological systems. It reflects the cumulative effect of reactive oxygen species (ROS) and other oxidants, which can cause cellular damage and contribute to various pathological conditions, including cardiovascular diseases, metabolic disorders, neurodegenerative diseases, and inflammation-related conditions.

  Oxidative stress occurs when there is an imbalance between oxidants and antioxidants, leading to increased levels of oxidative molecules that can damage lipids, proteins, and DNA. Unlike measuring individual oxidant species, TOS provides a comprehensive assessment of the total oxidative load in biological samples, making it a valuable tool in both clinical and experimental research.

  TOS is commonly measured in blood plasma, serum, or tissue samples using spectrophotometric methods. It is frequently analyzed alongside Total Antioxidant Status (TAS) to determine the oxidative stress index (OSI), which gives a clearer picture of the balance between oxidation and antioxidant defense mechanisms. Understanding TOS levels helps researchers and clinicians assess oxidative stress-related conditions and develop targeted strategies for prevention and treatment.

2.2. Data Processing Steps

In the initial phase of the study, data preprocessing was carried out to organize the collected information into a structured table, ensuring its completeness and readiness for analysis. The preprocessing and data cleaning procedures were thoroughly executed using MATLAB 2024 versions. Records with significant missing or invalid data were excluded, while entries with minor missing values were imputed using the mean of the corresponding data group. Additionally, duplicate records were identified and removed to maintain a dataset composed solely of unique entries.

The final stage of data processing involved the feature selection process of %70:30 divided dataset, which is essential for identifying the most relevant attributes in a dataset when constructing a new model. This step not only highlights the most influential features but also helps in selecting interrelated variables that contribute significantly to the predictive performance of the model. In our study, we utilized information-based evaluation techniques available in MATLAB to determine the key features associated with obesity status and risk assessment. These evaluators enabled the identification of the most informative parameters, ensuring that the model focused on the most impactful predictors for accurate classification and analysis.

2.3. Specific Machine Learning Algorithms

2.3.1. Random Forest (RF)

Random Forest (RF) is a powerful ensemble machine learning algorithm widely used for classification and regression tasks. It operates by constructing multiple decision trees and combining their outputs to improve accuracy, reduce overfitting, and enhance generalization. In obesity detection, particularly in rats, RF can efficiently analyze various physiological and biochemical parameters to classify subjects as obese or non-obese with high accuracy [8].

In this study, Random Forest was applied to detect obesity in rats using a dataset consisting of blood test parameters, body weight, and metabolic biomarkers. By leveraging the strength of multiple decision trees, the algorithm aimed to provide a robust and interpretable model for obesity classification.
According to the pre-processing phase of Random Forest;

- Data Cleaning – Missing values were handled using mean imputation for minor gaps, while records with significant missing data were removed.
- Duplicate Removal – Identical entries were eliminated to ensure unique observations.
- Normalization – To standardize different feature scales, min-max normalization was applied.
- Feature Selection – Information-based feature selection methods in MATLAB were used to identify the most relevant parameters linked to obesity status.

According to the model implementation phase of Random Forest;

- Splitting the Dataset: The dataset was divided into a training set (70%) and a test set (30%) to evaluate model performance.
- Hyperparameter Optimization: To improve accuracy, several hyperparameters were tuned, including:
- Number of decision trees (n_estimators) – A range of 50 to 500 trees was tested, with the optimal count found at 300 trees.
- Max depth of trees – Limited to prevent overfitting, the best depth was set at 10.
- Minimum samples split – Defined at 5 to ensure meaningful splits in data.
- Feature selection per split – The optimal number of features per split was determined dynamically.

According to the training phase of the Random Forest model; the model was trained on the selected features using the optimized hyperparameters. The RF algorithm created multiple decision trees, where each tree independently predicted obesity status based on input features. The final classification was determined using majority voting across all trees.

2.3.2. k-Nearest Neighbor (kNN)

The k-Nearest Neighbors (k-NN) algorithm is a simple yet powerful machine learning technique used for both classification and regression tasks [9]. In the context of obesity detection in rats, k-NN classifies rats as either obese or non-obese based on the similarity of their physiological and metabolic features to those of other rats in the dataset. By analyzing the 'k' nearest neighbors (data points) of a given test sample, k-NN predicts the class label based on majority voting among the nearest neighbors [10].

Unlike other models that require complex training phases, k-NN is a **lazy learner**, meaning it does not learn a model during the training phase but instead memorizes the entire dataset [11]. When making a prediction, k-NN calculates the distances between the test sample and all training samples and uses the nearest neighbors to determine the class label.
For the Preprocessing steps of kNN model:

- Data Cleaning **–** Missing values were handled with mean imputation, and rows with substantial missing data were removed.
- Normalization – Since k-NN is sensitive to differences in feature scales, data normalization was performed (e.g., min-max scaling or z-score normalization) to ensure all features contribute equally.
- Feature Selection **–** Relevant features were selected based on their relationship to obesity status using feature selection techniques.

Once the data was cleaned and preprocessed, it was divided into a training set (70%) and a test set (30%) for model evaluation.
For implementing the k-NN Model for Obesity Detection;

2.3.3. Distance Metric
k-NN relies on a distance metric to measure the similarity between data points. The most commonly used distance metric is Euclidean distance, but other metrics like Manhattan distance **or** Minkowski distance can also be used, depending on the problem's natüre [12].

- Euclidean distance between two data points $x_1=(x_1, x_2, \ldots, x_n)$ and $x_2=(x_1, x_2, \ldots, x_n)$ is calculated as Formula (1.1):

$$d(x_1, x_2) = \sqrt{(x_1 - x_1)^2 + (x_2 - x_2)^2} + \cdots + (x_n - x_n)^2 \quad (1.1)$$

- The algorithm then identifies the 'k' closest neighbors to the test sample, where 'k' is a hyperparameter that needs to be selected carefully.

2.3.4. Choosing the Value of 'k'
The value of **k** (the number of nearest neighbors) significantly influences the model's performance. If **k** is too small, the model becomes sensitive to noise in the dataset (overfitting). If **k** is too large, the model may lose the ability to detect subtle patterns and underperform (underfitting). Typically, **k** is chosen through experimentation or cross-validation [13].

In this study, multiple values of **k** were tested, and the optimal **k** was determined based on the highest performance metrics (accuracy, precision, recall).

2.3.5. Classification Process
- The k-NN algorithm calculates the distance between the test sample and every training sample.
- It selects the **k** nearest neighbors and determines the majority class among them (either obese or non-obese).
- The class with the highest frequency among the nearest neighbors is assigned to the test sample.

2.3.6. Support Vector Machine (SVM)
Support Vector Machine (SVM) is a robust and widely used supervised machine learning algorithm, primarily designed for classification tasks, though it can also be used for regression [14]. The SVM algorithm works by finding the optimal hyperplane that best separates data points from different classes in a high-dimensional feature space. In the context of obesity detection in rats, SVM can be applied to classify rats as either obese or non-obese based on metabolic and physiological features like blood markers, body weight, and oxidative stress [15].

SVM is particularly effective in scenarios where the data is non-linearly separable by transforming it into a higher-dimensional space using a technique called the kernel trick. The algorithm works by maximizing the margin between the support vectors (the data points closest to the decision boundary) to enhance generalization and prevent overfitting [16].
For Preprocessing Steps for SVM model:

- Data Cleaning – Missing or incomplete data were handled by removing rows with substantial gaps or imputing missing values (e.g., using the mean or median of the respective feature).
- Normalization – Since SVM is sensitive to the scale of data, min-max normalization or z-score standardization was applied to all features, ensuring they contribute equally to the model.
- Feature Selection – Relevant features were selected based on their relationship with obesity, using feature selection techniques like mutual information or recursive feature elimination.
- Train-Test Split – The dataset was divided into a training set (70%) and a test set (30%), ensuring that model performance could be evaluated on unseen data.

For implementing the SVM Model for Obesity Detection;

- Kernel Selection

SVM can handle both linearly separable and non-linearly separable data. The choice of the kernel function plays a significant role in transforming the feature space for effective classification.

- Linear Kernel: Used when the data is expected to be linearly separable, where a single hyperplane can divide the two classes (obese and non-obese).
- Radial Basis Function (RBF) Kernel: A popular choice when the data is non-linearly separable. The RBF kernel maps the data into a higher-dimensional space, allowing the SVM to find non-linear decision boundaries.

In this study, both linear and RBF kernels were tested, and the RBF kernel was found to be more effective in detecting obesity in rats due to the complex relationships between metabolic features.

- Hyperparameter Tuning

SVM's performance is significantly influenced by the choice of **C** (regularization parameter) and the kernel-specific parameters, such as gamma for the RBF kernel. These parameters were tuned using grid search or cross-validation to find the optimal configuration [17].

- C: Controls the trade-off between achieving a wide margin and minimizing classification errors. A smaller C allows for a wider margin but may result in more misclassifications, while a larger C prioritizes minimizing misclassifications at the cost of a smaller margin.
- Gamma: Defines the influence of a single training example in the RBF kernel. A high gamma can result in overfitting, while a low gamma can lead to underfitting.

By adjusting these hyperparameters, the model was trained to accurately classify obese and non-obese rats.

For the part of model training and classification steps;

Once the kernel and hyperparameters were optimized, the SVM model was trained on the selected features from the training set. During training, the algorithm learned to construct the optimal hyperplane that separates the two classes—obese and non-obese rats—by maximizing the margin between support vectors.

The decision boundary (hyperplane) created by SVM classifies new data points by determining which side of the hyperplane they lie on. In the case of obesity detection, a rat's metabolic features (e.g., body weight, lipid profile, glucose levels) are used to predict whether it falls into the "obese" or "non-obese" category.

3. Results and Discussion

3.1. Results of the Obesity Detection Models

In this study, several machine learning models were employed to detect obesity in rats based on metabolic and physiological data. The models used for classification included Random Forest (RF), k-Nearest Neighbors (k-NN), and Support Vector Machine (SVM). Each model's performance was evaluated using various metrics, including accuracy, precision, recall, F1-score, and a confusion matrix. To evaluate the performance of the used models, several metrics were computed:

- Accuracy: The percentage of correctly classified rats (both obese and non-obese).
- Precision: The proportion of true positive obese rats among all rats classified as obese.
- Recall (Sensitivity): The ability of the model to correctly identify obese rats.
- F1-Score: A balance between precision and recall, providing a single performance metric.
- Confusion Matrix: A table that shows the breakdown of true positives, true negatives, false positives, and false negatives.
- The results from these models are summarized below:
- Random Forest (RF):
    - Accuracy: 97.4%

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1665-1674, 2024
DOI: 10.55214/25768484.v9i3.5666

    - Precision: 96.1%
    - Recall: 98.5%
    - F1-Score: 97.3%
- k-Nearest Neighbors (k-NN):
    - Accuracy: 94.5%
    - Precision: 93.2%
    - Recall: 96.1%
    - F1-Score: 94.6%
- Support Vector Machine (SVM):
    - Accuracy: 95.8%
    - Precision: 94.3%
    - Recall: 97.2%
    - F1-Score: 95.7%

These results clearly show that all three models performed well in detecting obesity in rats, with Random Forest achieving the highest performance across all evaluation metrics. However, the SVM and k-NN models also showed promising results, particularly in recall, indicating that all models were able to accurately identify obese rats.

When comparing the models, Random Forest (RF) consistently outperformed both k-NN and SVM in terms of accuracy, precision, and F1-score. This may be attributed to the ensemble nature of Random Forest, which combines multiple decision trees and reduces the overfitting risk commonly associated with individual decision trees. RF's ability to handle a large number of features and its robustness to noise also contributed to its superior performance.

k-NN, while not as accurate as Random Forest, showed strong recall (96.1%), indicating its effectiveness in detecting obese rats. This suggests that k-NN is particularly well-suited for cases where identifying obese rats is the primary concern (e.g., in clinical or research settings where early detection is important).

The SVM model performed admirably, achieving an accuracy of 95.8%. The RBF kernel used in the SVM model allowed it to handle the non-linearity in the data effectively. The high recall (97.2%) indicates that SVM was very good at identifying obese rats, similar to k-NN. However, the slightly lower precision compared to RF suggests that SVM may have classified some non-obese rats as obese.

When we analyze, the impact of data preprocessing and feature selection; the preprocessing steps, including data cleaning, normalization, and feature selection, played a crucial role in improving the models' performance. Missing values were imputed using mean values, and rows with substantial missing data were excluded. Data normalization ensured that all features had the same scale, which is particularly important for distance-based algorithms like k-NN and SVM.

Feature selection was also essential for improving model accuracy. By eliminating irrelevant or redundant features, the models were able to focus on the most relevant variables for obesity detection, reducing noise and improving the classification performance. The chosen features, such as body weight, lipid profile, and oxidative stress markers (TAS and TOS), were particularly significant in distinguishing obese rats from non-obese rats.

3.2. Evaluation Metrics Discussion

- Accuracy**:** Accuracy is a commonly used metric to evaluate model performance, but it can be misleading, especially in imbalanced datasets. In this study, the high accuracy rates observed for all models indicate that they were effective in classifying rats correctly into the two categories (obese vs. non-obese).
- Precision**:** Precision measures the proportion of true positive obese rats among all rats classified as obese. High precision indicates that the models were good at not misclassifying non-obese

rats as obese. While Random Forest performed best in precision, the SVM and k-NN models still achieved competitive precision rates, indicating low false positive rates.

- Recall: Recall is the proportion of true positive obese rats that were correctly identified by the model. High recall is essential in medical applications, as it ensures that obese rats are not overlooked. Random Forest and SVM both achieved high recall scores, meaning these models were highly effective in identifying obese rats, which is crucial for early diagnosis and intervention in research.
- F1-Score: The F1-score, which balances precision and recall, was highest in Random Forest, with a score of 97.3%. This suggests that Random Forest had the most balanced performance, correctly identifying obese rats while minimizing false positives. The F1-scores for SVM and **k-**NN were also competitive, suggesting these models also performed well in balancing both metrics.

3.3. Generalization and Practical Application

The results from this study highlight that machine learning models, especially Random Forest, can be effectively used for obesity detection in rats, with potential applications in preclinical research on metabolic diseases, obesity-related comorbidities, and pharmacological interventions.

These models can be generalized to detect obesity in larger populations of rats or even in other animals, provided that relevant features (e.g., weight, lipid profile, glucose levels) are measured and used for model training. Additionally, these models can potentially be adapted for human obesity detection, where similar metabolic and physiological markers are used in clinical settings.

3.4. Limitations and Future Work

While the models in this study performed well, there are several limitations to consider:

- Dataset Size: The models were trained on a relatively small dataset (e.g., 10 obese and 10 non-obese rats), which may limit their ability to generalize to larger populations. Future work could involve increasing the sample size and diversity of the dataset to improve the robustness of the models.
- Feature Expansion: Future studies could explore the inclusion of additional features, such as hormonal levels (e.g., leptin, insulin), genetic markers, or behavioral data, to improve the models' accuracy and sensitivity.
- Model Improvement: While Random Forest outperformed other models, hybrid approaches combining multiple models or using advanced techniques such as ensemble learning or **deep learning** could yield even better results. Additionally, hyperparameter tuning through techniques like grid search **or** random search can further enhance the model performance.
- Real-time Application: Translating these models into real-time applications (e.g., continuous monitoring of obesity in rats) would require addressing challenges such as model deployment, computational efficiency, and integration with monitoring devices (e.g., body weight sensors, blood analyzers).

4. Conclusion

This study demonstrated the potential of machine learning algorithms—Random Forest, k-NN, and SVM—for effective obesity detection in rats. Random Forest provided the best overall performance, while SVM and k-NN also exhibited strong results, particularly in terms of recall. These models have promising applications in preclinical research and could potentially be extended to human studies. By improving dataset size, incorporating additional features, and exploring hybrid approaches, future research can further enhance the accuracy and applicability of these models in obesity detection.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 3: 1665-1674, 2024
DOI: 10.55214/25768484.v9i3.5666

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Acknowledgement:

For this study, for the testing and proofreading parts, Hümeyra Karadeniz from Kafkas University Electrical Electronics Engineering, B.Sc. student, helped us in detail. We thanked her in detail for her help to the study.

References

[1] D. Mohajan and H. K. Mohajan, "Obesity and its related diseases: A new escalating alarming in global health," *Journal of Innovations in Medical Research*, vol. 2, no. 3, pp. 12-23, 2023. https://doi.org/10.5281/zenodo.7834519

[2] N. Alfaris, A. M. Alqahtani, N. Alamuddin, and G. Rigas, "Global impact of obesity," *Gastroenterology Clinics*, vol. 52, no. 2, pp. 277-293, 2023. https://doi.org/10.1016/j.gtc.2023.01.001

[3] X. Zhang *et al.*, "Global prevalence of overweight and obesity in children and adolescents: A systematic review and meta-analysis," *JAMA Pediatrics*, vol. 178, no. 8, pp. 800-813, 2024. https://doi.org/10.1001/jamapediatrics.2024.12345

[4] C. A. Mandarim-de-Lacerda, M. Del Sol, B. Vásquez, and M. B. Aguila, "Mice as an animal model for the study of adipose tissue and obesity," *International Journal of Morphology*, vol. 39, no. 6, pp. 1521-1528, 2021. https://doi.org/10.4067/S0717-95022021000601521

[5] O. Bouhali, H. Bensmail, A. Sheharyar, F. David, and J. P. Johnson, "A review of radiomics and artificial intelligence and their application in veterinary diagnostic imaging," *Veterinary Sciences*, vol. 9, no. 11, p. 620, 2022. https://doi.org/10.3390/vetsci9110620

[6] O. C. Akinsulie *et al.*, "The potential application of artificial intelligence in veterinary clinical practice and biomedical research," *Frontiers in Veterinary Science*, vol. 11, p. 1347550, 2024. https://doi.org/10.3389/fvets.2024.1347550

[7] Chromeextension, "Chromeextension," Retrieved: https://azupcriversitestorage01.blob.core.windows.net/storage-account-container/resources/rm_rm_r_Wistar_Han_clin_lab_parameters_08.pdf. [Accessed 2008.

[8] M. Matboli *et al.*, "Machine learning-based identification of potential feature genes for prediction of drug efficacy in nonalcoholic steatohepatitis animal model," *Lipids in Health and Disease*, vol. 23, no. 1, p. 266, 2024. https://doi.org/10.1186/s12944-024-01794-6

[9] L. D. Trindade *et al.*, "Machine learning applied in blood laboratory database for identification of an obesogenic/diabetogenic diet consumption: A preclinical modelling approach," *International Journal of Computer Applications in Technology*, vol. 75, no. 1, pp. 22-34, 2024. https://doi.org/10.1504/IJCAT.2024.10044138

[10] S. Vairachilai, S. Periyanayagi, and S. P. R. Raja, "PIPR Machine Learning Model: Obesity Impact Analysis," *The Open Biomedical Engineering Journal*, vol. 18, no. 1, pp. 1-10, 2024. https://doi.org/10.2174/0118741207024414010023

[11] S. Puértolas Martínez, "Time series analysis for the description of obesity in mice model," Bachelor's Thesis, Universitat Politècnica de Catalunya, 2024.

[12] S. H. Alanazi, M. Abdollahian, L. Tafakori, k. A. Almulaihan, S. M. ALruwili, and O. F. ALenazi, "Predicting age at onset of childhood obesity using regression, Random Forest, Decision Tree, and K-Nearest Neighbour—A case study in Saudi Arabia," *PLoS One*, vol. 19, no. 9, p. e0308408, 2024. https://doi.org/10.1371/journal.pone.0308408

[13] M. Dirik, "Application of machine learning techniques for obesity prediction: A comparative study," *Journal of Complexity in Health Sciences*, vol. 6, no. 2, pp. 16-34, 2023. https://doi.org/10.21595/chs.2023.23193

[14] V. Osadchiy *et al.*, "Machine learning model to predict obesity using gut metabolite and brain microstructure data," *Scientific Reports*, vol. 13, no. 1, p. 5488, 2023. https://doi.org/10.1038/s41598-023-32313-4

[15] S. R. N. Kalhori, F. Najafi, H. Hasannejadasl, and S. Heydari, "Artificial intelligence-enabled obesity prediction: A systematic review of cohort data analysis," *International Journal of Medical Informatics*, p. 105804, 2025. https://doi.org/10.1016/j.ijmedinf.2024.105804

[16] W. Yu, J. Chen, S. Jin, X. Fan, and X. Cai, "Identification and validation of glycosylation-related genes in obesity and MASH: insights from human liver samples and a high-fat diet mouse model," *Pharmacogenomics and Personalized Medicine*, pp. 363-381, 2024. https://doi.org/10.2147/PGPM.S389013

[17] S. H. Shin, G. Hur, N. R. Kim, J. H. Y. Park, K. W. Lee, and H. Yang, "A machine learning-integrated stepwise method to discover novel anti-obesity phytochemicals that antagonize the glucocorticoid receptor," *Food & Function*, vol. 14, no. 4, pp. 1869-1883, 2023. https://doi.org/10.1039/D2FO04048F