SMU Data Science Review

Volume 8
Number 3 *Winter 2024*

Article 2

2024

Enhancing Animal Shelter Operations with Time Series and Machine Learning

Sakava L. Kiv
Southern Methodist University, sak2000.kiv@gmail.com

Donald L. Anderson
Southern Methodist University, donanderson@mail.smu.edu

Shivam Negi
Walmart Global Tech, shivamnegi92@gmail.com

Jacquelyn Cheun
Southern Methodist University, jcheun@mail.smu.edu

Recommended Citation

Kiv, Sakava L.; Anderson, Donald L.; Negi, Shivam; and Cheun, Jacquelyn (2024) "Enhancing Animal Shelter Operations with Time Series and Machine Learning," *SMU Data Science Review*: Vol. 8: No. 3, Article 2.
Available at: https://scholar.smu.edu/datasciencereview/vol8/iss3/2

Enhancing Animal Shelter Operations with Time Series and Machine Learning

Sakava L. Kiv[1], Donald L. Anderson[2], Shivam Negi[3], Jacquelyn Cheun[4]
[1] Master of Science in Data Science, Southern Methodist University,
Dallas, TX 75275 USA
{skiv, donanderson, jcheun}mail.smu.edu
shivamnegi92@gmail.com

Abstract. Enhancing animal shelter operations through machine learning involves employing a variety of advanced techniques aimed at increasing efficiency, promoting animal welfare, and optimizing resource allocation. This paper explores predictive analytics for adoption rates using regression models to estimate the likelihood of adoption based on historical data, encompassing variables such as breed, health status, and previous adoption trends. Additionally, classification algorithms are utilized to categorize animals by adoption probability, facilitating better resources and marketing prioritization. Clustering algorithms are employed to group animals according to behavior patterns and/or physical health, enabling tailored medical care and enrichment activities that improve their mental and physical well-being. Time series (ARIMA) forecasting is applied to predict future animal intake capacity influencing inventory management and room allocation. Random forest and neural networks (LSTM) are integrated as supplemental predictive models, offering deep learning capabilities for more complex data patterns and serve as an additional tool in shelter decision-making processes. A key goal of this research is to reduce overcrowding/euthanasia rates, increase adoption rates, and increase Return to Owner (RTO) rates at the Dallas Animal Shelter by forecasting animal intake and predicting length of stay. The integration of these machine learning methodologies demonstrates enhancements in the operational effectiveness of animal shelters, ensuring that both animal care and resource utilization are maintained at high standards. The findings highlight the potential of these techniques to enhance shelter management, promote more successful adoption outcomes, and reduce overcrowding/euthanasia rates.

1 Introduction

This research focuses on enhancing animal shelter operations through the application of machine learning and times series techniques. This topic is significant because animal shelters play a vital role in the community by caring for and rehoming abandoned animals.

Enhancing the animal shelter operations not only promotes animal welfare but also helps shelter management staff to plan and identify areas of opportunities for resource allocation. For example, increasing staff for busy times, preparing room capacity for different types of animals based on breed, behavior, size, and health status. This is critical for shelters that often operate with limited resources.

In 2023, animal shelters and rescue organizations across the United States faced significant challenges due to the influx of animals. According to data from Shelter Animals Count (2023), over 6.5 million animals were received by animal shelters and rescue organizations, comprising 3.3 million cats and 3.2 million dogs. Of this total, 48% were strays, while 25% were voluntarily surrendered by their owners. The situation has become increasingly critical, as shelters have entered their fourth consecutive year of experiencing an imbalance between the number of animals received and adoption rates, particularly concerning dogs.

Since January 2021, an additional 900,000 animals have entered and remained in shelters, exacerbating the national capacity crisis (Best Friends Animal Society, 2023). Notably, for the first time since 2016, the number of dogs euthanized has surpassed the number of cats euthanized. Non-live outcomes, which include animals that died in care, were lost in care, or were euthanized, have risen significantly since the onset of the COVID-19 pandemic, accounting for 850,000 animals (Shelter Animals Count, 2023).

Despite these challenges, there is a positive trend in adoption rates. In 2023, 61% of dogs and cats taken into shelters or rescues found permanent homes (Best Friends Animal Society, 2023). The strain on animal shelters operating at or above capacity leads to overcrowding, adversely affecting the well-being of animals, their chances of adoption, and the quality of care they receive (Kilgour & Flockhart, 2022). Advocacy for pet-friendly housing policies and increased access to veterinary care is essential to keep pets in loving homes and out of shelters (Horecka, K., & Neal, S. 2022). These statistics highlight the pivotal role that animal shelters play in providing care, finding suitable homes, and addressing the challenges posed by overcrowding. There is a critical need for community support and systemic changes to improve conditions and outcomes for animals in shelters. By addressing these issues, we can enhance the welfare of animals and ensure that shelters continue to fulfill their essential role in our communities.

Investigating this topic is crucial because many animal shelters face challenges such as overcrowding, limited funding, and high euthanasia rates. By using machine learning and time series techniques, shelters can better predict adoption rates, optimize animal care, manage inventory, increase staff levels for certain periods of influx of animals, and plan financial needs more effectively. Specifically, reducing capacity and euthanasia/overcrowding rates at shelters like the Dallas Animal Shelter is a worthy goal that can be achieved through improved capacity planning, resource allocation, and best practices for the animal shelter management domain.

Experts have explored various applications of machine learning in animal shelter operations. For instance, Best Friends Animal Society has developed a sophisticated data model using artificial intelligence and machine learning to predict outcomes and prioritize resources in U.S. shelters (Best Friends Animal Society, 2024). Predictive analytics have been used to estimate adoption likelihood using features such as length of stay Bradley, J., & Rajendran, S. (2021), while classification algorithms can help determine animal outcome (Mitrović, K., Milošević, D., & Greconici, M., 2019). These classification algorithms can be used to help allocate resources to animals that need them most and drive marketing campaign strategies to increase adoption. Clustering algorithms facilitate tailored care based on behavior patterns (Menaker, T., Monteny, J., de Beeck, L. O., & Zamansky, A., 2022). Also, time series forecasting is mentioned to analyze Return to Owner (RTO) Kremer, T. (2021). Time series could be used to predict animal intake and track interventions such as managed shelter intake best practices effect on adoption outcomes (Kreisler, R. E., Pugh, A. A., Pemberton, K., & Pizano, S., 2022). These studies have demonstrated that time series and machine learning can aid the decision-making processes in shelters, leading to effective management and better outcomes for animals.

While progress has been made, some animal shelters have yet to fully embrace combining different machine learning and time series methods. The barriers include being underfunded, lacking sufficient staff to collect data for use in machine learning algorithms (Bradley, J., & Rajendran, S., 2021). This research aims to use time series analysis and machine learning techniques such as LSTM (Long-Short Term Memory) artificial neural networks and including others such as regression models, classification algorithms, clustering algorithms, and random forest to provide solutions to problems of overcrowding, low adoption rates, and euthanasia rates at the Dallas Animal Shelter.

The application of this research will explain how integrating these machine learning and time series techniques can enhance animal shelter management at Dallas Animal Shelter. By improving the accuracy of predictive models and enabling more informed decision-making, shelters can achieve higher standards of animal care and resource utilization. This can lead to more successful adoption outcomes and reduce overcrowding, specifically for Dallas Animal Shelter and can also be used for other shelters potentially.

2 Literature Review

The literature review focuses on the integration of machine learning and time series analysis on some key areas for animal shelter enhancements: Animal intake, Euthanasia, Adoption, Return to Owner (RTO), Transfers, Overcrowding, and Diversion strategies. An overview of the literature reveals a growing body of research focused on the application of machine learning in animal shelter management. The literature highlights the interconnectedness of animal intake, euthanasia, adoption rates, return to owner rates, transfers, overcrowding, and diversion strategies in shelter operations. The integration of machine learning and time series analysis can aid in enhancements for managing these factors. By developing predictive models and using data-driven insights, shelters can reduce overcrowding, reduce euthanasia rates, and enhance the overall welfare of the animals in their care. This comprehensive approach highlights the potential of time series and machine learning techniques in enhancing shelter management practices.

2.1 Animal Shelter

2.1.1 Animal Intake

Animal intake in shelters is a critical metric that affects various operational aspects. Intake rates can be highly variable due to factors such as seasonality, local socio-economic conditions, and community outreach efforts (Kilgour & Flockhart, 2022). Time series analysis can help understand these dynamics by examining monthly intake and outcome patterns. This approach can identify trends, correlations, and anomalies in intake data. Using time series analysis, specialized staff can optimize resource allocation by identifying peak intake periods, analyzing possible seasonal trends for stocking supplies, and predicting natural disasters or crises with historical data. Moreover, time series analysis can enhance welfare strategies by developing targeted outreach during high intake periods for specific animals. It also identifies trends in behavior or health for training programs and analyzes the impact of spay/neuter initiatives on intake rates. Additionally, time series analysis can inform policy decisions by guiding local legislation on animal control, evaluating Trap-Neuter-Return (TNR) program effectiveness, and identifying high abandonment areas for targeted education outreach. In retrospect, peak intake periods can be identified, and targeted outreach programs can be developed. The effectiveness of interventions can be evaluated, and data-informed adjustments can be made. Understanding animal intake patterns is critical for supporting animal welfare and shelter operations.

2.1.2 Euthanasia and Adoption Outcomes

Euthanasia rates are a sensitive yet essential aspect of shelter operations. High intake rates often correlate with higher euthanasia rates, particularly in shelters that struggle with overcrowding. A significant study titled, *Increasing Adoption Rates at Animal Shelters: A Two-Phase Approach To Predict Length of Stay and Optimal Shelter Allocatio*n, discusses the grim reality that nearly 50% of animals entering shelters are euthanized annually, with 10%-25% due to overcrowding (Bradley, J., & Rajendran, S., 2021). To reduce euthanasia rates for healthy adult cats, shelters have explored novel approaches, including waiving adoption fees. A case study of a free adoption-drive for cats aged ≥1 year in a Western Australian shelter revealed positive outcomes. The free adoption-drive successfully rehomed 137 cats, increasing weekly adoptions by 533%. First-time adopters constituted a larger portion of the free cohort due to mixed-media promotions. Both free and normal-fee adopters selected cats of similar age, sex, and pelage. Post-adoption, both groups retained over 90% of the cats, with comparable incidences of medical and behavioral issues. Legislative compliance regarding collars, registration, and roaming did not differ between adopter groups. Overall, there was no evidence of adverse outcomes associated with free adoptions. Shelters should consider occasional free adoption-drives during overflow periods (Crawford, H. M., Fontaine, J. B., & Calver, M. C., 2017). This study aims to increase adoption rates by predicting the length of stay for animals using machine learning algorithms which could help manage euthanasia rates and using time series to find optimal times for marketing campaigns as needed. Adoption rates are important for balancing intake and reducing overcrowding in shelters. Factors influencing adoption rates include the shelter's location, marketing efforts, and the animals' health and behavior. According to (Protopopova, Alexandra & Gunter, Lisa., 2017), shelters that employ targeted marketing strategies, such as highlighting the personality traits and histories of animals, see higher adoption rates. Machine learning helps shelters analyze historical adoption data to find patterns and factors that increase adoption rates. This allows shelters to adjust their strategies more effectively. Time series forecasting can also help predict adoption trends and optimize resources for adoption events.

2.1.3 Return to Owner (RTO)

The Return to Owner (RTO) rate is another critical metric that reflects the effectiveness of a shelter's efforts to reunite lost pets with their owners. Animals with microchips, older animals, healthy animals, neutered animals, and animals brought to the shelter by another public agency exhibit a higher likelihood of being reunited with their owners. Notably, neither the sex of the animal nor the season of impound significantly affect the return-to-owner (RTO) rates for dogs or cats (Hill, C. R., Weng, H.-Y., Protopopova, A., & Ly, L. H., 2023). Machine learning can play a role in enhancing RTO rates by predicting which animals are most likely to be reclaimed by their owners and identifying the most effective outreach methods. Kremer (2021) discusses the development of a web-based tool focused on RTO data analysis. Machine learning and time series methods can help predict RTO rates.

2.1.4 Transfers

Transfers between shelters and rescue organizations are a crucial component of managing shelter populations, particularly in cases of overcrowding. Transferring animals to facilities with more resources or higher adoption rates can reduce the strain on individual shelters and increase the likelihood of positive outcomes for the animals. (Kreisler, R. E., Pugh, A. A., Pemberton, K., & Pizano, S., 2022) emphasize the importance of incorporating best practices for transfers, including health assessments and transportation protocols. Machine learning and time series analysis can assist in identifying optimal times and candidates for transfers, enhancing the efficiency and effectiveness of these operations. For example, the potential of advanced data analysis techniques, such as cluster algorithms and time series analysis, to address these challenges. Cluster algorithms can be utilized to identify specific breeds, behaviors, and other characteristics that determine the suitability of animals for transfer to specialized shelters or fosters.

This targeted approach ensures that animals are placed in environments best suited to their needs, thereby improving their chances of adoption and overall well-being. Moreover, time series analysis can play a crucial role in predicting intake rates, enabling shelters to anticipate periods of high intake and prepare accordingly. By forecasting these trends, shelters can proactively manage their capacity and arrange transfers before reaching critical limits, thus maintaining a balanced and efficient operation.

2.1.5 Overcrowding

Overcrowding in animal shelters poses significant challenges to animal welfare and operational efficiency. It leads to increased stress, disease transmission, and resource strain. Preventing disease spread in shelters is a challenging task, often measured by infection rates. Shelters, housing transient mixed-species animals, face biological instability. Overcrowding, poor hygiene, and housing multiple species create an ideal environment for pathogen transmission. Disease outbreaks can lead to crises. While some diseases are more impactful, detailed studies are needed. Depopulation during outbreaks is sometimes considered, but understanding efficient containment strategies within shelter constraints is crucial (Horecka, K., & Neal, S. 2022). Time series analysis further aids in understanding and anticipating periods of high occupancy, allowing for better preparation and resource allocation strategies.

2.1.6 Diversion Strategies

Diversion strategies are essential for reducing the intake pressure on shelters and preventing overcrowding. These strategies include self-rehoming via online platforms (Ly, L. H., & Protopopova, A., 2023). Diversion programs offer help to pet owners by giving access to affordable vet care, behavior training, and short-term care for pets. These programs include efforts to teach the community about caring for pets responsibly and the importance of spaying or neutering. Advanced machine learning can help identify which people and places need the most help, making sure the right resources get to the right areas.

In summary, integration of machine learning techniques in animal shelter operations can lead to significant improvements in efficiency, resource allocation, and animal welfare. Specifically, the application of machine learning and time series will enable shelters to better anticipate adoption rates, manage inventory, and optimize capacity planning, resulting in decreased euthanasia/overcrowding rates.

2.2 Data Science Techniques used in Animal Shelter Management

2.2.1 Machine Learning (ML)

Machine learning (ML) is a critical area within artificial intelligence (AI). It involves the development of algorithms and statistical models that allow computers to perform tasks autonomously. Unlike traditional programming, ML systems are not explicitly instructed on how to complete tasks. Instead, they learn from data and adapt their performance over time (Sarker, 2021). There are several key types of machine learning. The first is supervised learning, where algorithms are trained on labeled datasets. Each piece of training data is associated with a specific output label. The goal is to create a mapping from inputs to outputs, enabling the prediction of labels for new data (Sarker, 2021).

Unsupervised learning is another important type of ML. In this approach, algorithms work with data that lacks labeled responses. The focus is on discovering hidden structures within the data, such as by clustering similar data points or reducing dimensionality (Sarker, 2021). Semi-supervised learning is a hybrid approach in ML.

It combines both labeled and unlabeled data during the training process. This method is especially useful when obtaining a fully labeled dataset is costly or time-consuming, as it leverages the benefits of both types of data (Sarker, 2021). Reinforcement learning is another significant ML technique. In reinforcement learning, an agent learns by interacting with an environment. The agent makes decisions to maximize cumulative rewards, drawing on principles from behavioral psychology (Sarker, 2021).

Machine learning has wide-ranging applications across various industries. In healthcare, ML algorithms are used for disease diagnosis, outcome prediction, and personalized treatment planning. In cybersecurity, ML helps detect and prevent threats by analyzing patterns and anomalies in network traffic (Sarker, 2021). ML is also transforming smart cities by optimizing traffic management, energy consumption, and public safety. In e-commerce, it powers recommendation systems, customer segmentation, and demand forecasting (Sarker, 2021). As ML continues to advance, it brings both new opportunities and challenges. Understanding its core principles and diverse applications is essential. Harnessing ML's potential can lead to substantial improvements in many areas of life (Sarker, 2021).

In animal shelter management, ML can be used to predict animal intake, improve adoption rates, reduce euthanasia and overcrowding rates by making informed decisions based on historical and real-time data. The ability to predict outcomes and identify trends can enhance shelter operations' efficiency and effectiveness.

2.2.2 Time Series Analysis

The autoregressive integrated moving average (ARIMA) model is a key method in time series forecasting, particularly for linear data. Introduced by Box and Jenkins in 1970, ARIMA has been widely applied in fields like finance, economics, and social sciences to predict future values based on historical data. The model uses three parameters: autoregression (p), differencing (d), and moving average (q), which together help in capturing linear patterns in the data (Lee & Tong, 2011).

However, ARIMA has limitations when applied to nonlinear time series, as it is inherently designed for linear data. To address this, hybrid models that combine ARIMA with nonlinear techniques, such as artificial neural networks (ANN) and support vector machines (SVM), have been developed. These hybrids take advantage of both linear and nonlinear approaches, providing more accurate forecasts for complex datasets (Lee & Tong, 2011). For example, ARIMA-ANN and ARIMA-SVM models have been used in various applications, including stock price prediction and energy demand forecasting. In these hybrid models, ARIMA is first used to model the linear component of the time series. The residuals, which represent the variance not explained by the linear model, are then modeled using ANN or SVM to capture any remaining nonlinear dynamics (Lee & Tong, 2011). The final forecast is produced by combining the linear predictions from ARIMA with the nonlinear predictions from ANN or SVM, resulting in a more comprehensive and accurate prediction. This approach effectively overcomes the limitations of using ARIMA alone, especially in real-world scenarios where data often exhibits both linear and nonlinear characteristics (Lee & Tong, 2011).

In conclusion, while ARIMA is a powerful tool for linear time series forecasting, its integration with nonlinear methods like ANN and SVM significantly enhances its effectiveness. This combination results in more accurate forecasts, making it suitable for a wider range of applications (Lee & Tong, 2011).

2.2.3 Time Series Analysis and Machine learning Integration Enhancements

Studies have shown that predictive analytics, particularly through regression models, can effectively predict the length of stay of each animal based on a range of factors such as breed, size, and health status (Bradley & Rajendran, 2021). This could allow shelters to prioritize resources and marketing efforts towards animals with higher adoption probabilities and/or help create campaigns to help the animals with lower probabilities to get adopted or prepare resources to enhance behavioral and physical health for animals predicted for longer length of stay rate.

Classification algorithms can be utilized to categorize animals by characteristics which can determine outcome (Mitrović, K., Milošević, D., & Greconici, M., 2019). This can facilitate better resource allocation and improve overall operational efficiency. By accurately predicting adoption outcomes, shelters can optimize their capacity planning and reduce overcrowding, leading to improved animal welfare.

Clustering algorithms have been employed to group animals based on behavior patterns (Menaker, T., Monteny, J., de Beeck, L. O., & Zamansky, A., 2022). This can enable tailored medical care and enrichment activities. This personalized approach enhances the mental and physical well-being of animals, increasing their chances of adoption.

Time series forecasting techniques for (RTO) have been mentioned as future research endeavor Kremer, T. (2021). Time series ARIMA technique, can be applied to predict future animal intake capacity. By analyzing historical data, shelters can anticipate fluctuations in demand and effectively manage their inventory and financial resources.

The integration of random forest and neural networks as supplemental predictive models aid in the shelter management decision making process, offering deep learning capabilities for more complex data patterns such as length of stay (Bradley, J., & Rajendran, S., 2021). These advanced techniques can enable shelters to make more informed decisions and achieve better outcomes for both animals and shelter operations.

While the literature provides promising evidence of the effectiveness of machine learning and times series techniques in animal shelter management, there is still a need for further research, particularly in addressing challenges such as data scarcity and model interpretability. Additionally, the impact of external factors, such as the COVID-19 pandemic (Rodriguez, Jeffrey & Davis, Jon & Hill, Samantha & Wolf, Peter & Hawes, Sloane & Morris, Kevin., 2022), on shelter operations warrants investigation to ensure the resilience and adaptability of these predictive models in dynamic environments.

3 Methods

3.1 Data Sources and Preprocessing

The data we initially downloaded spans from October 2014 to September 2024 with approximately 300,000 records. With our added columns (intake_month, intake_year, and intake_trend) we have a total of 34 features with a combination of numerical and categorical data.

1. animal_id
2. animal_type
3. animal_breed
4. kennel_number
5. kennel_status
6. activity_number
7. activity_sequence
8. source_id
9. census_tract
10. council_district
11. intake_type
12. intake_subtype
13. reason
14. staff_id
15. intake_date
16. intake_time
17. due_out
18. intake_condition
19. hold_request
20. outcome_type
21. outcome_date
22. outcome_time
23. receipt_number
24. impound_number
25. service_request_number
26. outcome_condition
27. chip_status
28. animal_origin
29. additional_information
30. month
31. year
32. intake_month
33. intake_year
34. intake_trend

This study utilizes historical data from the Dallas Animal Shelter, publicly available online as open data. The dataset encompasses comprehensive records of animal intake, adoption rates, demographic information, and health status. A primary metric, intake trend, was created to track animal intake on a monthly and yearly basis.

Initial exploratory data analysis revealed inconsistencies in column names and feature counts. To ensure data integrity, a thorough cleaning process aggregated features into uniform names aligned with the Dallas Animal Shelter Open data sets. Data was downloaded in CSV format but can be accessed through API calls as well.

The time series model utilizes data spanning from October 2014 to September 2024. However, as the dataset is updated daily, it may include more recent dates. The cleaned and formatted data provides a robust foundation for subsequent time series analysis and modeling. The Dallas Animal Shelter Open Data serves as the primary source for this research, offering a rich and dynamic data set for exploring animal intake trends and patterns.

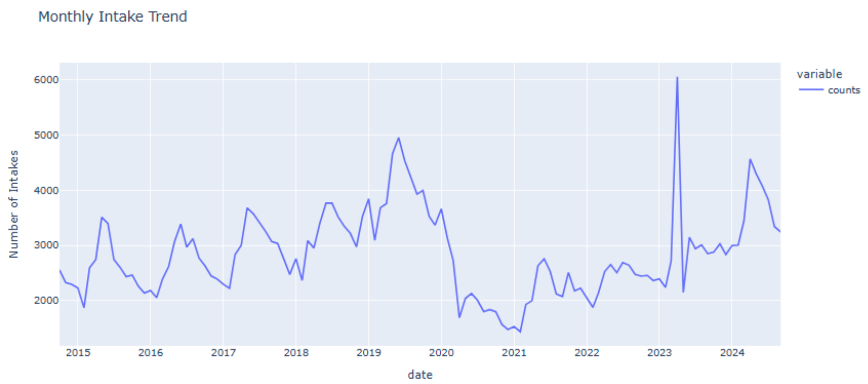


Fig.1. The time series exhibits stationary behavior with seasonal fluctuations, peaking in summer, and event-driven influences, notably the COVID-19 pandemic corresponding with historically low intake rates around February 2020 - January 2021. These insights inform data-driven strategies, requiring stationarity testing, seasonal decomposition, and nonlinear modeling to capture complex relationships and develop accurate forecasting models.

3.2 Model Training Approach

This study employs a multifaceted model training approach to optimize capacity planning and resource allocation at the Dallas Animal Shelter. Regression models can be utilized for predictive analytics, estimating the likelihood of adoption, and predicting length of stay based on historical data. Specifically, incorporating variables such as breed, health status, and previous adoption trends. Furthermore, classification algorithms employ adoption probability categorization to enhance resource allocation, promoting tailored health interventions, behavioral support, and strategic marketing. Also, clustering algorithms group animals according to characteristics and behavior patterns, enabling tailored medical care and enrichment activities.

Now, time series forecasting, specifically ARIMA, predicts future animal intake capacity, aiding room allocation strategies, transfer to partner shelters, foster care decisions, and marketing campaigns. Finally, Neural networks supplement base models like ARIMA and Random Forest, offering deep learning capabilities for complex data patterns.

By utilizing these model training strategies and data sources, this research endeavors to improve capacity planning and resource allocation at the Dallas Animal Shelter. The primary objective is to mitigate overcrowding and euthanasia rates, thereby supporting informed operational decisions and refining organizational efficiency.

4 Results

This analysis reveals that animal intake data in Texas exhibits stationarity due to longer and warmer climate patterns. A time series is considered stationary if its statistical properties, such as mean, variance, and autocorrelation, remain constant over time. Stationarity is crucial for modeling and forecasting.

Several factors contribute to stationarity in animal intake data. Longer warm periods lead to prolonged breeding cycles for dogs and cats reducing seasonality, resulting in varied intake throughout the years. Moreover, warmer climates encourage outdoor activities, increasing the likelihood of stray animals being brought into shelters. Additionally, warmer environments increase wandering behavior trends in animal intake. Therefore, animals are generally more active in warm climates, leading to wandering intake patterns. Furthermore, heat waves causing short-term spikes, also contribute to stationarity.

Statistical tests and visual inspections confirm stationarity. The Augmented Dickey-Fuller (ADF) test, Autocorrelation Function (ACF) plot, and Spectral Density plots provide evidence of stationarity, justifying the selection of the ARIMA (2,0,0) model without seasonality at least for shorter periods of time forecast. These findings have significant implications for modeling and forecasting animal intake, enabling more accurate predictions and informed decision-making.

4.1 Visual and Statistical Analysis of Stationarity

Before we start let's define an AR (1) model. The AR (1) model is defined by the equation:

$$X_t = \phi_1 X_{t-1} + \epsilon_t \tag{1}$$

In equation (1), X_t is the value of the time series at time t. ϕ_1 is the autoregressive coefficient, indicating the influence of previous (X_{t-1}) on the current value (X_t). ϵ_t denotes the white noise error term at time t, which is assumed to be independently and identically distributed with a mean of zero and constant variance. Understanding AR (1) models is important for gaining insight into more general AR (P) models, as well as more complex models such as ARMA and ARIMA.

Now, to further verify stationarity, additional checks were performed using statistical and visual tests. The evidence of stationarity in the time series was supported by the Augmented Dickey-Fuller (ADF) test (Fig.1), Autocorrelation Function (ACF) plot (Fig.2), and Spectral Density plots (Fig.3). The stationarity tests and visual test provide strong evidence for the stationarity of the animal intake time series data.

Collectively, these results help to justify the assumption of stationarity in the animal intake time series data, supporting the use of the ARIMA (2,0,0) model for forecasting. The stationarity verification is crucial for accurate forecasting, as non-stationary data can lead to biased predictions. By confirming stationarity through statistical and visual tests, this analysis ensures reliable forecasting results using the selected ARIMA model.

The ARIMA (2,0,0) model excels at capturing short-term patterns, rendering it an effective framework for animal intake forecasting over shorter periods, including daily forecasts, weekly projections, and monthly predictions.

Additionally, its robustness and accuracy make it potentially suitable for longer-term forecasting to a certain degree. However, for more complex long-term forecasts, incorporating a Seasonal ARIMA (SARIMA) model or other models such as ARIMA ANN can provide enhanced accuracy. These models can offer valuable insights for strategic planning and decision-making in animal shelter operations with a focus on predicting animal intake.

```
ADF Statistic: -3.114119992780748
p-value: 0.02552573977173465
Critical Values 1%: -3.4870216863700767
Critical Values 5%: -2.8863625166643136
Critical Values 10%: -2.580009026141913
The time series is stationary according to the ADF test.
```

Fig.1. The Augmented Dickey-Fuller (ADF) test results indicate rejection of the null hypothesis of non-stationarity, with a p-value of 0.02552, which is less than the significance level of 0.05. Furthermore, the ADF statistic of -3.114 is less than the critical value at 5% significance, confirming stationarity.

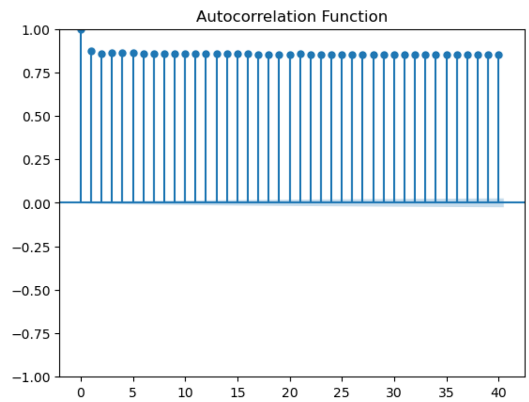


Fig.2. The Autocorrelation Function (ACF) plot exhibits slow dampening, indicating significant autocorrelation at lower lags and remaining significant even at higher lags as seen here. Also, the

ACF indicates our AR model has a $|\phi_1| < 1$ which is close to 1. Despite the slow dampening in the ACF plot, the time series remains stationary as long as the absolute values of $|\phi_1| < 1$.

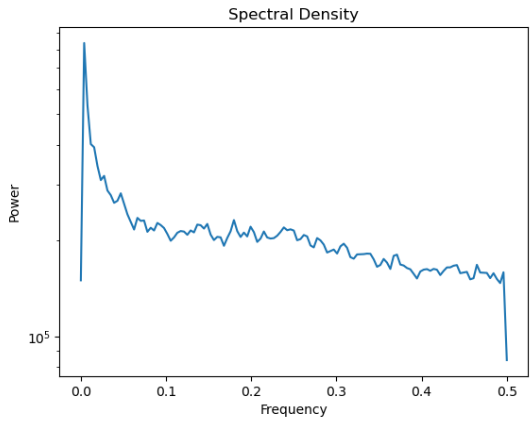


Fig.3. Observe that the Spectral Density plot shows a peak near zero indicating a stationarity process.

4.2 Model Selection and Performance Evaluation for Time Series Forecasting

This section presents the results of the time series forecasting analysis, demonstrating the effectiveness of the selected models in predicting animal intake. The ARIMA (2,0,0) time series model exhibited superior performance in forecasting animal intake data, achieving Average Square Error (ASE) accuracy scores of 217,925 and 141,964 for 5-month and 10-month forecasts, respectively. This model outperformed other tested models, including SARIMA, Facebook Prophet (open-source software for forecasting time series data), Random Forest, XGBoost (Xtreme Gradient Boosting), and LSTM (Long-Short Term Memory) a type of ANN, which yielded higher ASE and RMSE values (Table 1). Keep in mind that the lower the ASE or RMSE the better the model performs, since we are capturing reduced error bounds. Notably, the other models could be tuned for more accuracy, but were kept with default settings.

The ARIMA (2,0,0) model's forecasting capabilities are visualized in Fig. 4 (5-month prediction) and Fig. 5 (10-month prediction). To facilitate practical shelter management decisions, the ASE value of 217,925 for the ARIMA (2,0,0) 5-month prediction forecast was converted to a Root Mean Square Error (RMSE) of 467. This translates to an average error of approximately 467 animals per month. For instance, if 2,000 animals are predicted to be taken in next month, the actual number could range from 1,533 to 2,467, indicating a moderate 23% variance in intake. Although this error level is moderate, the model provides valuable insights into overall trends and seasonal patterns, serving as a foundation for refinement and improved accuracy.

The demonstrated forecasting capabilities of the ARIMA (2,0,0) model support its suitability for animal intake prediction, enabling informed decision-making in shelter operations. Future refinements can focus on integrating additional factors or exploring alternative models, such as SARIMA or machine learning algorithms, to enhance forecasting accuracy.

Table 1. A comprehensive assessment of the performance of various time series models, the open-source forecasting tool Prophet, and adapted machine learning methods tailored for time series forecasting, is presented. The primary evaluation metrics, Absolute Sums of Errors (ASE) and Root Mean Squared Errors (RMSE), are calculated for both 5-month and 10-month forecasting horizons, facilitating a systematic comparison of the models' efficacy.

Models	ASE 5-Month	RMSE 5-Month
ARIMA (2,0,0)	217,925	467
ARIMA (1,1,1), S = 12	1,094,009	1,046
Prophet	1,464,871	1,210
Random Forest	875,221	936
XGBoost	240,621	491
LSTM	816,520	904

Models	ASE 10-Month	RMSE 10-Month
ARIMA (2,0,0)	141,964	377
ARIMA (1,1,1), S = 12	1,094,009	1,046
Prophet	628,917	793
Random Forest	879,204	938
XGBoost	240,621	491
LSTM	848,476	921

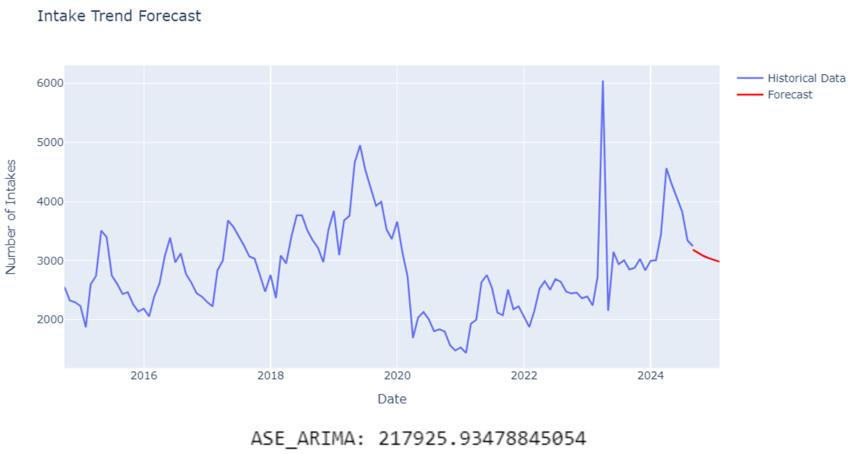
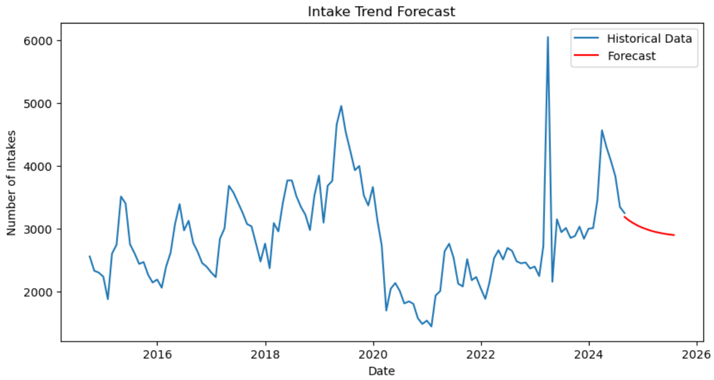


Fig.4. Observe that for this ARIMA (2,0,0) for a 5-month forecast shows a decrease in trend for animal intakes. This fits our intuition as winter months are ahead and thus lower animal intake.




```
# Calculate the Average Squared Error (ASE)
actual = series[-10:].values  # Convert to numpy array if it's a pandas Series
ASE_ARIMA = np.mean((actual - preds) ** 2)
print("ASE_ARIMA:", ASE_ARIMA)
```

```
ASE_ARIMA: 141964.85144605752
```

Fig.5. For this ARIMA (2,0,0) for a 10-month period, we see a trend towards a mean value.

4.3 Practical Use Case for Animal Shelter

The ARIMA (2,0,0) model's Average Squared Error (ASE) value of 217,925 translates to a Root Mean Square Error (RMSE) of 467. This metric indicates an average monthly deviation of approximately 467 animals from predicted intake. Considering a predicted intake of 2,000 animals for the next month, the actual figure may range from 1,533 to 2,467. This reflects a moderate 23% variance, providing valuable insights into overarching trends and seasonal patterns. The model's performance suggests moderate accuracy, suitable for strategic planning, with room for refinement to enhance predictive capabilities. Future directions include exploring additional predictors, seasonal decomposition, and fine tuning the models to improve accuracy and inform shelter management strategies. To further optimize the model, regular updates with new data are necessary, investigating sources of error and exploring ensemble methods can also enhance performance, enabling more informed decision-making and effective shelter management.

4.5 K-Modes Clustering Analysis for Animal Shelter Data

After research and analysis, K-Modes clustering was used to uncover patterns in animal intake and outcomes at the Dallas Animal Shelter. This method adapts the popular K-Means algorithm for categorical variables, overcoming limitations of distance-based clustering techniques. By clustering data, this method aimed to group animals with similar characteristics, identifying distinct categories for informed decision-making.

The dataset comprised detailed records of animals passing through the shelter, including type, breed, intake condition, and outcome. Guided by domain knowledge, features were selectively removed lacking meaningful differentiation for clustering animal profiles, primarily administrative and time-based variables.

To determine the optimal number of clusters (k) for K-Modes, the Elbow Method was implemented, plotting cost (inertia) against varying k values to pinpoint the "elbow" where cost reduction slowed significantly. Two initialization methods were employed to ensure robust results, providing a reliable foundation for subsequent analysis.

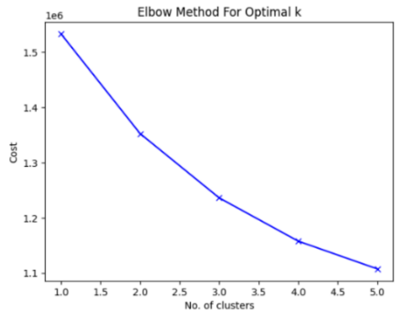


Fig.6. Cao Initialization is designed to be faster and more efficient for large datasets. It helps to reduce the chance of getting poor initial clusters by ensuring that the initial cluster centers are well spread out. This method works well with imbalanced datasets (where some categories appear more frequently than others).

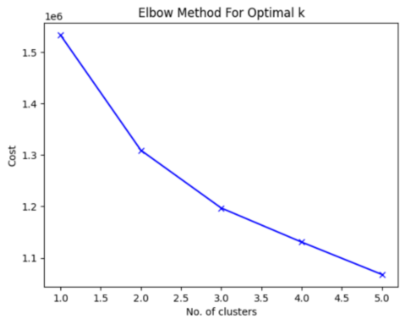


Fig.7. The Huang Initialization, a default method for K-Modes, randomly selects initial cluster centers from the dataset, suitable for smaller datasets where random selection biases are minimized. However, its effectiveness may diminish with larger datasets, potentially leading to suboptimal clustering performance. Figures 6 and 7 illustrate the impact of initialization methods on determining the optimal number of clusters (k). Both Cao and Huang initializations reveal a consistent pattern, indicating k=3 as the optimal cluster number. This is evidenced by a sharp decline in cost preceding k=3, followed by diminishing returns beyond this point, characterized by a flattening curve. The Elbow Method confirms k=3 as the optimal cluster number, marked by a pronounced decrease in cost before this point and negligible reductions thereafter. The lack of a distinct elbow after k=3 suggests that additional clusters yield marginal improvements, supporting the conclusion that three clusters optimally capture the underlying structure of the data.

Now taking these concepts of optimal number of clusters and fitting K-Modes, based on the Elbow Method, the optimal number of clusters were set to 3. Furthermore, the K-Modes algorithm was applied to the dataset, assigning each record to one of the three clusters.

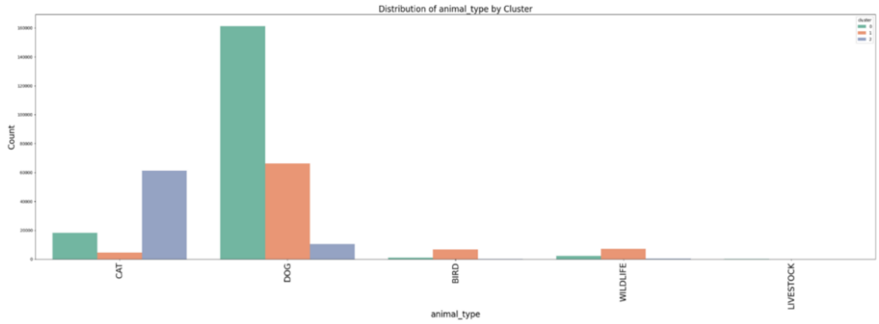


Fig.8. Visualizing the Cluster Distribution: Count plots were generated to show the distribution of each categorical variable within the clusters. This provided a clear visualization of how certain breeds, conditions, and outcomes were concentrated within specific clusters.

A hierarchical clustering approach revealed distinct patterns in animal shelter data, highlighting differences in animal characteristics and outcomes. Dogs predominantly comprised two clusters, whereas cats were overrepresented in another. Notable correlations emerged between intake conditions and outcomes, such as animals undergoing rehabilitation being more likely to be adopted, whereas certain conditions were strongly associated with euthanasia. To clearly interpret the characteristics of each cluster, we conducted a comprehensive analysis of categorical variable distributions across clusters. See Fig.9.

Cluster 0 was primarily composed of Pit Bull breeds, mostly strays taken in under the "at large" subtype, frequently categorized as treatable and rehabilitatable, with adoption being the most common outcome.

In contrast, Cluster 1 was dominated by mixed-breed dogs, mainly owner surrenders, with an elevated proportion of euthanized animals having conditions marked as cared for by an APP, which is an Advanced Practice Provider who diagnosis the outcome as "within normal limits" (APP WNL).

Cluster 2 was largely comprised of domestic shorthair cats, mostly strays, with transfer being the most frequent outcome, surpassing adoption, and euthanasia. These insights provide a deeper understanding of the complex relationships between animal characteristics, intake conditions, and outcomes in animal shelters, informing data-driven decision-making and policy development. By examining the distribution of various categorical variables across clusters, this analysis highlights the dominant characteristics of each cluster, offering valuable implications for animal welfare organizations and policymakers to act upon.

cluster	animal_type	animal_breed	intake_type	intake_subtype	intake_condition	outcome_type	outcome_condition	animal_origin
0	DOG	PIT BULL	STRAY	AT LARGE	TREATABLE REHABILITABLE NON-CONTAGIOUS	ADOPTION	TREATABLE REHABILITABLE NON-CONTAGIOUS	OVER THE COUNTER
1	DOG	MIXED BREED	OWNER SURRENDER	GENERAL	APP WNL	EUTHANIZED	APP WNL	FIELD
2	CAT	DOMESTIC SH	STRAY	AT LARGE	APP WNL	TRANSFER	APP WNL	OVER THE COUNTER

The above results show the most common category for each feature in every cluster, helping us understand what type of data points dominate each cluster.

Fig.9. The above table row results show the most common category for each feature in every cluster, helping us to identify and interpret what type of data points dominate each cluster.

The cluster-specific category frequency distribution provides valuable insights into the dominance of specific categories within each cluster. By normalizing the counts across clusters, this analysis reveals how certain attributes are distributed among the clusters, offering a deeper understanding of the underlying structure of the data. Notably, detailed results are extensive and can be reviewed directly in the code notebook for a comprehensive understanding of category distributions across clusters. Preliminary findings suggest notable variations in category frequencies across clusters.

To determine the significance of each categorical feature in defining the clusters, we employed the chi-square test, examining the association between each feature and the cluster labels generated by the K-Modes algorithm. The chi-square statistic, degrees of freedom, and p-value enabled the identification of features exerting substantial influence on the clustering results. Specifically, features with elevated chi-square values and decreasing p-values were deemed more pivotal in shaping cluster composition. Ranking features in decreasing order of importance based on chi-square values revealed discernible patterns, showcasing the differential impact of various attributes on cluster formation. Analysis of Table 2 shows the relative importance of each feature, with ‘animal_breed’ being the most important and ‘animal_origin’ being the least important in this case.

Table 2. Feature importance in decreasing order with Chi-square value metric.

Feature	Chi-square value
animal_breed	252,217.60
animal_type	202,161.97
intake_subtype	158,061.35
outcome_condition	153,029.50
intake_condition	130,057.17
intake_type	98,170.54
outcome_type	84,303.58
animal_origin	46,467.66

In conclusion, K-Modes clustering effectively grouped animals with similar intake conditions, breeds, and outcomes, providing the shelter with actionable insights. The three clusters revealed distinct profiles of animals: rehabilitated dogs likely to be adopted, surrendered dogs often euthanized, and stray cats often transferred. These insights can help the shelter make better decisions regarding resource allocation, focusing efforts on rehabilitation and adoption strategies for animals in specific clusters.

4.6 Predicting Animal Outcomes Using Multi Class Logistic Regression

This study developed a predictive model utilizing multiclass logistic regression to classify outcomes for animals entering Dallas Animal Shelter, including Adoption, Return to Owner (RTO), Euthanasia, and Other. By using historical data, the model aims to predict outcome types for new animal entries, informing resource allocation and enhancing animal welfare.

The analysis employed a curated version of the animal shelter dataset, with irrelevant or redundant features removed, specifically animal_id, census_tract, council_district, reason, hold_request, outcome_date, due_out, intake_time, and outcome_time. Time-related features were derived from intake_date, including intake_day, intake_day_of_week, intake_day_of_year, intake_week, intake_quarter, intake_semester, and intake_is_weekend. Features were separated into categorical and numerical sets for targeted transformations.

The outcome_type target variable was encoded into four distinct classes for multiclass classification: Adoption, Returned to Owner (RTO), Euthanized, and Other. The dataset was partitioned into training (64%), validation (16%), and test sets (20%) for model development and evaluation.

Categorical features underwent One-Hot Encoding (OHE), while numerical features were standardized using StandardScaler to ensure parity in scale for logistic regression. Optuna facilitated hyperparameter tuning, yielding optimal parameters for the logistic regression model after 10 trials. The refined model employed a solver of saga, penalty of l2, and regularization strength (C) of 1.0481552889251464. The optimized model was retrained on the combined training and validation sets.

The test set evaluation yielded an accuracy of 85%, with macro average precision, recall, and F1-score of approximately 0.86. Weighted average precision, recall, and F1-score were around 0.85. Class-specific performance metrics demonstrated robust results, with Adoption yielding precision of 0.77, recall of 0.87, and F1-score of 0.82; Euthanized yielding precision of 0.97, recall of 0.96, and F1-score of 0.96; Other yielding precision of 0.86, recall of 0.79, and F1-score of 0.82; and Returned to Owner (RTO) yielding precision of 0.88, recall of 0.82, and F1-score of 0.85. See Fig. 10.

Classification Report:

	precision	recall	f1-score	support
ADOPTION	0.77	0.87	0.82	20779
EUTHANIZED	0.97	0.96	0.96	13425
OTHER	0.86	0.79	0.82	24610
RETURNED TO OWNER	0.88	0.82	0.85	9252
accuracy			0.85	68066
macro avg	0.87	0.86	0.86	68066
weighted avg	0.86	0.85	0.85	68066

Fig. 10. Classification Report results

These findings indicate the model's effectiveness in predicting animal shelter outcomes with an overall accuracy of 85%, providing valuable insights for resource allocation and animal welfare improvement. The multiclass logistic regression model's performance was evaluated by examining the classification results for each outcome category. Now analyzing the confusion matrix Fig. 11, we discovered the following:

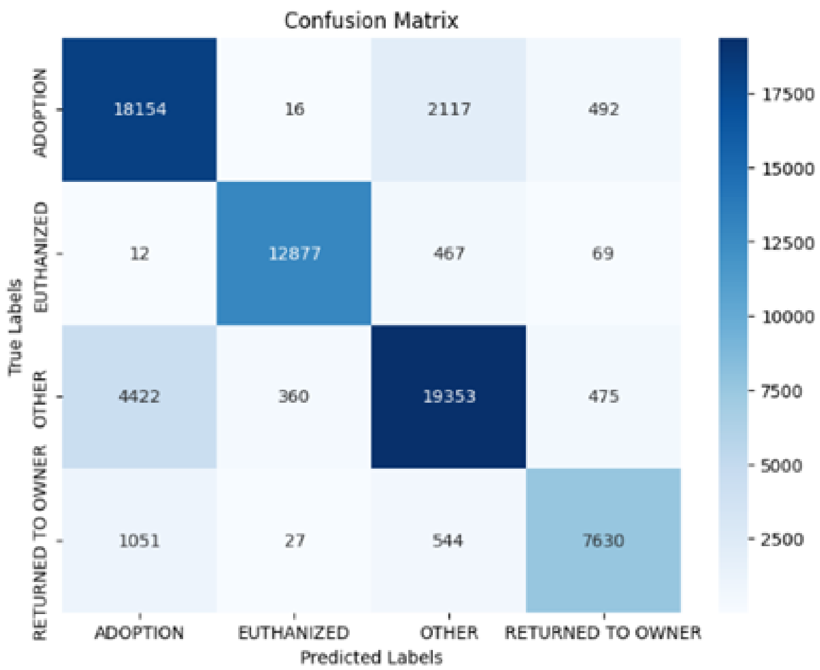


Fig. 11. Confusion Matrix plot results

Adoption outcomes were accurately predicted for 18,154 instances, representing true positives. However, 16, 2,117, and 492 instances were misclassified as Euthanized, Other, and Returned to Owner, respectively. These errors account for false positives in the respective categories.

Conversely, Euthanized outcomes were correctly predicted for 12,877 instances. Misclassification occurred in 12, 467, and 69 instances, which were incorrectly predicted as Adoption, Other, and Returned to Owner.

The model accurately predicted 19,353 instances as Other. However, 4,422, 360, and 475 instances were misclassified as Adoption, Euthanized, and Returned to Owner.

Returned to Owner outcomes were correctly predicted for 7,630 instances. Misclassification occurred in 1,051, 27, and 544 instances, which were incorrectly predicted as Adoption, Euthanized, and Other.

The classification results demonstrate varying degrees of accuracy across outcome categories. To further investigate the model's performance, predicted probabilities for each class were retained for subsequent analysis.

The predicted probabilities provide valuable insights into potential improvements in animal care and shelter operations. Specifically, identifying animals with low adoption probabilities can inform targeted adoption efforts and resource allocation. This information can be leveraged by shelters to prioritize interventions and optimize outcomes.

In conclusion, the logistic regression model achieved an overall accuracy of 85%, indicating solid performance in predicting animal outcomes. Hyperparameter tuning via Optuna enhanced the model's generalizability to unseen data. The results demonstrate the utility of multiclass logistic regression in animal shelter outcome prediction and highlight areas for future improvement.

5 Discussion

5.1 Interpretation of Results

The application of time series analysis, specifically the ARIMA (2,0,0) model, yielded a 5-month forecast with a Root Mean Square Error (RMSE) of 467. This predictive model has significant implications for the Dallas Animal Shelter, enabling them to anticipate and prepare for future capacity levels. By leveraging this forecast, the shelter can inform strategic decision-making, such as optimizing resource allocation to ensure adequate staffing, facilities, and services to meet the predicted demand. The forecast can also inform capacity planning, allowing the shelter to develop strategies to manage capacity, such as implementing adoption promotions or expanding partnerships with rescue organizations. Moreover, by anticipating potential capacity shortages or surpluses, the shelter can develop contingency plans to mitigate risks and ensure the welfare of animals in their care. The predictive model's accuracy provides a reliable basis for the shelter's strategic planning and decision-making, ultimately enhancing its ability to provide effective and efficient animal care services.

Furthermore, the K-Modes clustering algorithm effectively partitioned the dataset into three distinct clusters, each characterized by unique intake conditions, breeds, and outcomes. The clusters revealed distinct animal profiles, including rehabilitated dogs likely to be adopted, surrendered dogs often euthanized, and stray cats frequently transferred. Lastly, the logistic regression model's overall accuracy of 85% indicates robust performance in predicting animal outcomes, enabling data-driven decision-making and strategic planning.

5.2 Implications

This study's findings demonstrate the efficacy of integrating time series analysis and machine learning methodologies in optimizing animal shelter operations. The synergistic application of predictive analytics, classification algorithms, clustering techniques, and time series forecasting has been shown to support decision-making processes, with the goal of contributing to improved animal welfare outcomes. Specifically, the implementation of these methodologies has the potential to increase adoption rates while concurrently reducing overcrowding and euthanasia rates, thereby alleviating the strain on shelter resources.

5.3 Limitations

One of the challenges encountered during the analysis was the availability and quality of data. While efforts were made to utilize historical data from the Dallas Animal Shelter, limitations in data collection and completeness may have impacted the accuracy of the predictive models. Additionally, the complexity of implementing machine learning techniques in real-world shelter settings may pose practical challenges for some organizations.

5.4 Sustainability and Scale

To ensure the long-term sustainability and scalability of these solutions, future research directions may include addressing data limitations, exploring ethical considerations, and conducting evaluations of the impact of machine learning integration on overcrowding/euthanasia rates and adoption outcomes.

Moreover, to maintain the accuracy and effectiveness of these models, it is essential to establish a framework for ongoing model maintenance and refinement. This may involve hiring staff or recruiting volunteers with domain expertise in time series forecasting to update and refine models on a regular basis, thereby ensuring the continued validity and reliability of predictions. By prioritizing sustainability and scalability, animal shelters can harness the full potential of machine learning and time series techniques to drive lasting improvements in animal welfare outcomes.

6 Conclusion

In conclusion, this research has demonstrated the transformative potential of integrating machine learning methodologies into animal shelter operations. By leveraging predictive analytics, classification algorithms, clustering algorithms, and time series forecasting techniques, animal shelters can optimize efficiency, resource allocation, and animal welfare outcomes. The findings of this study feature the utility of machine learning in enhancing shelter management, promoting data-driven decision-making, and ultimately improving adoption outcomes while reducing overcrowding and euthanasia rates.

This research contributes to the advancement of animal shelter management by providing empirical evidence of the effectiveness of machine learning techniques in enhancing operational effectiveness and promoting better outcomes for animals in need. The integration of these advanced technologies presents a promising avenue for future research and practical implementation in animal shelter operations. As such, this research serves as a valuable resource for specialized staff and animal shelters seeking to optimize their operations and improve animal welfare outcomes, highlighting the potential for data science to drive meaningful change in this critical domain.

Acknowledgments. Jacquelyn Cheun, PhD. – Capstone Professor, Bivin Sadler, PhD. – Time Series Professor, Shivam Negi, MSc. – Capstone Advisor, Robert Slater, PhD. – Quantifying The World Professor, Jacob Turner, PhD. – Applied Statistics Professor

References

1. Shelter Animals Count. (2023). 2023 Annual Report. Shelter Animals Count. https://www.shelteranimalscount.org/sac-releases-2023-annual-analysis
2. Best Friends Animal Society. (2023). *Annual report 2023*. Retrieved from https://bestfriends.org/annual-report
3. Kilgour, R. J., & Flockhart, D. T. T. (2022). Direct and indirect factors influencing cat outcomes at an animal shelter. Frontiers in Veterinary Science, 9. https://doi.org/10.3389/fvets.2022.766312
4. Horecka, K., & Neal, S. (2022). Critical Problems for Research in Animal Sheltering, a Conceptual Analysis. *Frontiers in Veterinary Science,* 9. https://doi.org/10.3389/fvets.2022.804154
5. Bradley, J., & Rajendran, S. (2021). *Increasing Adoption Rates at Animal Shelters: A Two-Phase Approach to Predict Length of Stay and Optimal Shelter Allocation*. BMC Veterinary Research, 17, Article number: 70. doi:10.1186/s12917-020-02728-2
6. Mitrović, K., Milošević, D., & Greconici, M. (2019). Comparison of machine learning algorithms for shelter animal classification. In 2019 IEEE 13th International Symposium on Applied Computational Intelligence and Informatics (SACI) (pp. 211-216). IEEE. https://doi.org/10.1109/SACI46893.2019.9111575
7. Menaker, T., Monteny, J., de Beeck, L. O., & Zamansky, A. (2022). Clustering for automated exploratory pattern discovery in animal behavioral data. Frontiers in Veterinary Science, 9. https://doi.org/10.3389/fvets.2022.884437

8. Kremer, T. (2021). A new web-based tool for RTO-focused animal shelter data analysis. Frontiers in Veterinary Science, 8. https://doi.org/10.3389/fvets.2021.669428
9. Kreisler, R. E., Pugh, A. A., Pemberton, K., & Pizano, S. (2022). The impact of incorporating multiple best practices on live outcomes for a municipal animal shelter in Memphis, TN. Frontiers in Veterinary Science, 9. https://doi.org/10.3389/fvets.2022.786866
10. Crawford, H. M., Fontaine, J. B., & Calver, M. C. (2017). Using Free Adoptions to Reduce Crowding and Euthanasia at Cat Shelters: An Australian Case Study. *Animals*, 7(12), 92. DOI: 10.3390/ani7120092
11. Protopopova, Alexandra & Gunter, Lisa. (2017). Adoption and relinquishment interventions at the animal shelter: A review. Animal Welfare. 26. 35-48. 10.7120/09627286.26.1.035.
12. Hill, C. R., Weng, H.-Y., Protopopova, A., & Ly, L. H. (2023). Factors Affecting the Likelihood of Dogs and Cats Returning to Their Owners at a Municipal Animal Shelter in the United States. Journal of Shelter Medicine and Community Animal Health, 2(1). https://doi.org/10.56771/jsmcah.v2.64
13. Ly, L. H., & Protopopova, A. (2023). Predictors of successful diversion of cats and dogs away from animal shelter intake: Analysis of data from a self-rehoming website. Animal Welfare, 32, e13. doi:10.1017/awf.2023.8
14. Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions1. SN Computer Science, 2(3), 160. https://doi.org/10.1007/s42979-021-00592-x
15. Lee, Y.-S., & Tong, L.-I. (2011). Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. *Knowledge-Based Systems, 24*(1), 66-72. https://doi.org/10.1016/j.knosys.2010.07.006
16. Rodriguez, Jeffrey & Davis, Jon & Hill, Samantha & Wolf, Peter & Hawes, Sloane & Morris, Kevin. (2022). Trends in Intake and Outcome Data From U.S. Animal Shelters From 2016 to 2020. Frontiers in Veterinary Science. 9. 863990. 10.3389/fvets.2022.863990.