Predict Post-shelter Outcomes for Dogs

Analysis of Long Beach Shelter Data from 2017-2024

C.Reilly
University of Colorado Boulder

ABSTRACT

This project uses Long Beach Animal Care Services data to predict outcomes for shelter dogs. The dataset represents over 9000 dogs that entered and left the shelter between 2017 and 2024. Shelters routinely struggle with capacity, so the goal is to provide a means for predicting outcomes for dogs upon their intake to the shelter and to identify features associated with live outcomes.

This project compared the performance of gradient boost, AdaBoost, and random forest models on predicting the shelter outcomes of dogs. The gradient boost model had the best performance with F1 scores of 0.73, 0.64, and 0.71 and AUC of 0.87, 0.85, and 0.80 for adoption, euthanasia, and rescue outcomes respectively. Sex, intake condition, and age on intake were the most influential features in the model.

CCS CONCEPTS

• Data mining • Classification • Machine learning

KEYWORDS

Gradient boost, classification, data mining, feature importance, adoption, shelter

ACM Reference format:

C Reilly. 2024. Predict Post-shelter Outcomes for Dogs: Analysis of Long Beach Shelter Data from 2017-2024. In *Proceedings of Boilerplate conference (Boilerplate '24). Boilerplate, New York, NY, USA, 2 pages.* https://doi.org/10.1145/1234567890

1 Introduction

Animals are spending longer in shelters, which reduces the number of new intakes shelters can care for. In the first half of 2024, over 1 million dogs have entered shelters in the US [7]. In order to continue providing care to new intakes, shelters need to find outcomes for dogs already in care. Of the over 1.5 million dogs that have left shelters in the first half of 2024, 12% had non-live outcomes, which includes euthanasia [7].

If shelters are better able to identify the factors that predict shorter shelter stays and positive outcomes, they can better serve the dogs in their care. For example, social media promotion can focus on dogs identified as less likely to be adopted. If medical conditions or injuries are associated with lower adoption rates, shelters can prioritize veterinary care in order to increase the adoptability of those dogs.

Previous research has used subsets of data available. Some studies have focused on specific age groups, features, or outcome types. Overall, age, behavior, and appearance are common factors associated with post-shelter outcomes in existing research. This project aims to use multiple features to predict a range of post-shelter outcomes. In addition, it focuses on information available on intake. This will result in a more widely applicable result, which shelters can use to inform programming.

2 Related Work

Previous research has explored features related to post-shelter outcomes. Behavior, age, and appearance are common factors among prior research. One subset of work focuses on specific features of dogs. McGuire et al. [2021] examined the impact of behavioral evaluation performance on dogs' adoption rate. They found that dogs exhibiting "dangerous" behavior had longer shelter stays than dogs with "concerning" or "non-concerning" behaviors.

The latter two groups did not differ in shelter stay. The applicability of these results is limited due to only one predictive feature being investigated. Cain et al. [2020] also focused on a narrow subset of features to predict adoptability. Their study analyzed the phenotype of dogs to determine whether a dog's appearance is predictive of its likelihood of adoption. Their results indicate that puppies, small dogs, and non-brachycephalic dogs are more likely to be adopted.

Another subset of research focuses on specific types of dogs. Hawes et al. [2018] chose to focus on outcomes for senior dogs. The most significant feature for dogs in their study is intake condition, leading to their conclusion that specialized vet care may increase positive outcomes for senior dogs.

The final subset focused on specific outcomes or outcome measures. Powell et al. [2021] analyzed a wider range of characteristics, including age, intake type, sex, and return frequency, but focused on predicting whether a dog would be returned to the shelter after adoption. Behavioral issues were also a predictive factor in this study and were the most common reason noted for returning a dog. Age and breed were also predictive of the likelihood of return. Both Sazara and Gao [2022] and Bradley et al. [2021] used multiple features to predict length of shelter stay. Age was a common feature. Bradley et al. [2021] also found that size and color impacted the length of shelter stay.

Logistic regression was a common method used [1, 3, 4, 5]. Gradient boosting was also used in multiple studies [1, 6, 8].

3 **Proposed Work**

This project uses data available from Long Beach Animal Care Services, located in Long Beach, CA. The dataset includes records of all dogs that have entered and left the shelter between January 2017 and July 2024. There are 23 features and 9679 observations.

The project aimed to explore which features are associated with specific outcome types and to build a model to predict outcomes based on information available upon intake to the shelter.

Preprocessing

The preprocessing stage had two main tasks: cleaning and feature engineering. Three new features were added to the dataset. 'Age on Intake' was calculated based on 'DOB' and 'Intake Date.' 'Time in Shelter' was calculated based on 'Intake Date' and 'Outcome Date.' A categorical feature ('Age Category') was also added based on this feature. Dogs were categorized as 'Puppy' (<= 6 months), 'Young' (6 months - 3 years), 'Adult' (3 - 8 years) and 'Senior' (> 8 years). Finally, 'Time in Shelter' was calculated from 'Intake Date' and 'Outcome Date.'

12 features were removed from the dataset due to a lack of predictive value. 'Reason for Intake' was also removed due to 88% of observations missing a value for that feature. 310 observations were removed because they were missing a value for the target variable or had a rare outcome type (< 100 observations).

Missing values were imputed with the mode for 'Primary Color,' 'Secondary Color,' 'Intake Subtype,' and 'Sex.' Missing values for 'Age on Intake' were imputed with the median.

The final preprocessing step grouped similar values within categorical features. There were 38 and 22 unique values for 'Primary Color' and 'Secondary Color' respectively. These were reduced to 8 and 7. 'Outcome Type' was grouped into adoption, rescue, euthanasia, and return to owner. Intake conditions of mild, moderate, or severe illness were combined with mild, moderate, and severe injury respectively. This reduced the number of subcategories with few observations in order to improve analysis. The categorical features used in the models and their values are displayed in Table I. 'Age on Intake' was the only continuous feature used for the models.

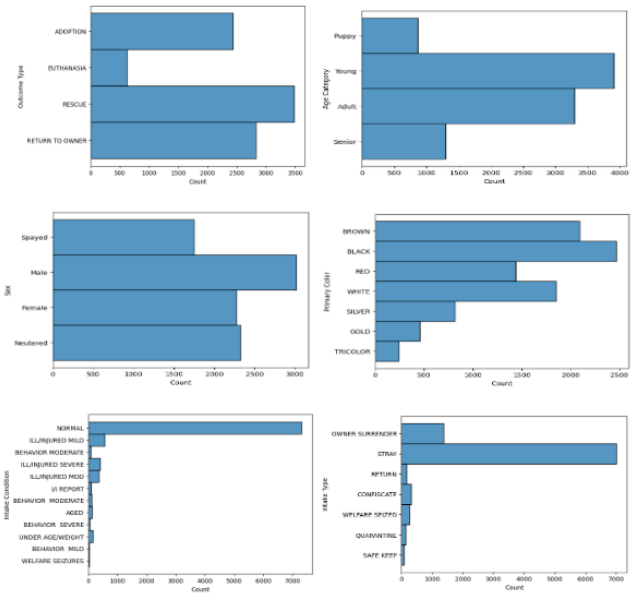
TABLE I: CATEGORICAL FEATURE VALUES

Feature	Values
Primary Color	[BROWN, BLACK, RED, WHITE, SILVER, GOLD, TRICOLOR]
Secondary Color	[WHITE, TRICOLOR, BROWN, BLACK, SILVER, GOLD, RED, None]
Sex	[Spayed, Male, Female, Neutered]
Intake Subtype	[OTC, FIELD, EVICTION, PUB SAFETY, POLICE, CRUELTY, BITE, BORN@SHELT, ABANDON, HOSPITAL, OWNER DIED, RESCUE, EMERGENCY]
Intake Condition	[NORMAL, ILL/INJURED MILD, BEHAVIOR MODERATE, ILL/INJURED SEVERE, ILL/INJURED MOD, I/I REPORT, BEHAVIOR MODERATE, AGED, BEHAVIOR SEVERE, UNDER AGE/WEIGHT, BEHAVIOR MILD, WELFARE SEIZURES]
Intake Type	[OWNER SURRENDER, STRAY, RETURN, CONFISCATE, WELFARE SEIZED, QUARANTINE, SAFE KEEP]
Outcome Type	[ADOPTION, EUTHANASIA, RESCUE, RETURN TO OWNER]

EDA

The first step in this stage was to execute descriptive analysis of the dataset. The histogram of outcome types revealed that euthanasia is the least frequent outcome type (7%). Figure 1 displays the histogram of outcome types.

FIGURE I: OUTCOME AND FEATURE HISTOGRAMS



Young (42%) and adult (35%) dogs were more frequent than puppies (9%) or seniors (14%). Intact male dogs were the most frequent sex (32%). Within color, brown (22%) and black (26%) were the most common values, followed by white (20%). Both intake condition and type have uneven distributions with normal intake condition (78%) and stray intake type (75%) occurring significantly more often than other values.

Due to the categorical nature of the predictor and outcome variables, relative frequency tables were used to assess relationships between the predictors and outcomes. There were several interesting findings. Spayed and neutered dogs were more likely to be adopted, whereas intact dogs were more likely to be placed with a rescue. 7% of dogs overall were euthanized, but approximately half of dogs with severe behavior classifications were euthanized. Severely ill or injured dogs were also more likely to be euthanized than have other outcomes. 81% of dogs classified as underage or underweight were placed with rescues. The majority (91%) of dogs in that category were puppies. 65% of puppies were placed with rescues, which is nearly twice the overall rate of 37%. Senior dogs were less likely than other age groups to be adopted (11%) and more likely to be returned to their owner (49%).

FIGURE II: DENSITY PLOTS

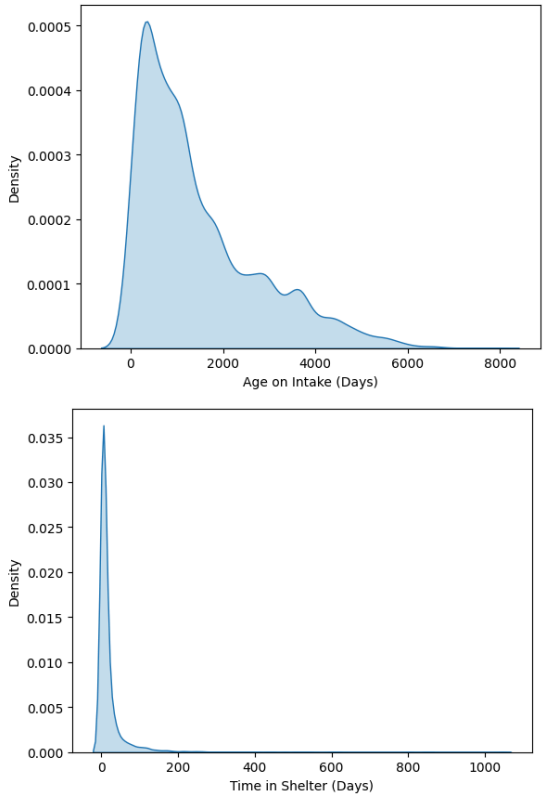
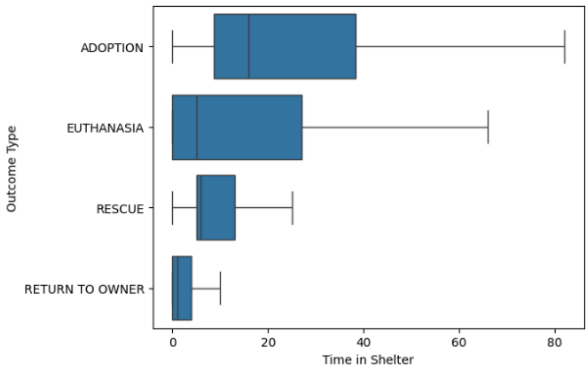


FIGURE III: TIME IN SHELTER BOXPLOT BY OUTCOME TYPE



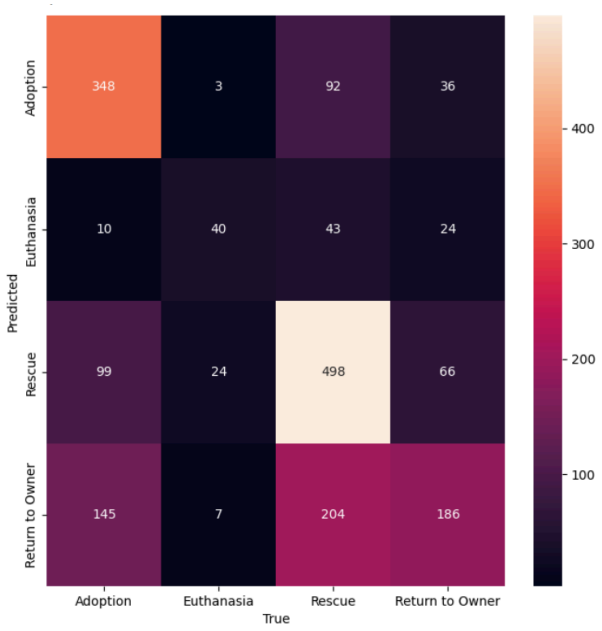
The majority of dogs across outcomes stayed in the shelter fewer than 100 days. There are differences in the time spent in shelter between outcome types, with dogs that were returned to their owners having shorter stays on average (median = 1 day). 79% of dogs that were returned to their owners were strays, which in combination with their shorter shelter stays on average, suggests that many of these dogs were within their stray hold period, during which other outcome types are not permitted. Adoption and euthanasia were associated with wider ranges of time spent in the shelter.

Modeling

Due to the imbalance between outcome types, observations with an outcome type of euthanasia were upsampled in order to provide more meaningful analysis. Upsampling was conducted within the train and test sets to avoid overfitting from including the same observation in both train and test sets. An 80-20 split was used to divide the data into train and test sets. The split was stratified by outcome type to ensure similar distributions of outcomes between the sets. Categorical features were also converted to dummy variables for analysis.

Gradient boost, AdaBoost, and random forest models were trained and initially evaluated based on confusion matrices and F1 scores. All models struggled to classify 'Return to Owner.' Figure 3 shows the confusion matrix from the gradient boosting model in which 349 of the observations that were predicted to be 'Return to Owner' were classified as either 'Adoption' or 'Rescue.'

FIGURE IV: GRADIENT BOOST MODEL CONFUSION MATRIX



This pattern repeated across multiple models, which suggests that the available features are not sufficient to differentiate 'Return to Owner' from other live outcomes. Given the short shelter stays for dogs in this class, there is limited utility in predicting this outcome. For these reasons, this outcome class was dropped.

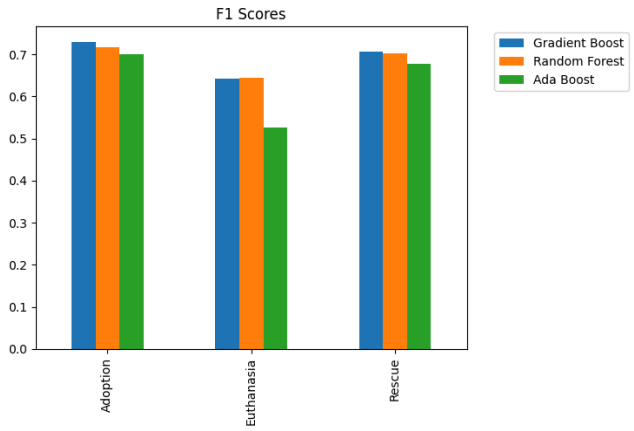
Gradient boosting, AdaBoost, and random forest models were run with the three remaining classes ('Adoption', 'Euthanasia,' and 'Rescue'). Each model was optimized with grid search cross validation.

4 Evaluation

F1 score was the primary metric because it provides a more nuanced means of evaluating model quality overall and within classes. Each model performed

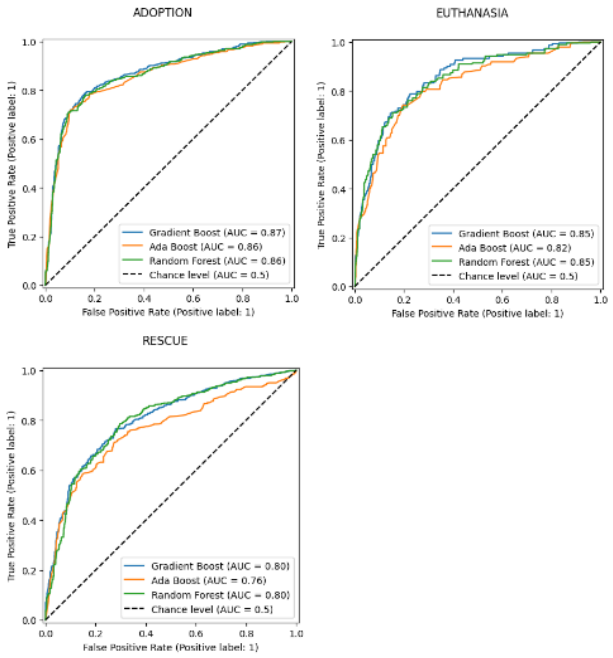
worst for 'Euthanasia,' but the disparity was greatest for the AdaBoost model (0.53). The gradient boost model performed slightly better than the random first model with scores of 0.73, 0.64, and 0.71 for 'Adoption,' 'Euthanasia,' and 'Rescue' respectively.

FIGURE V: F1 SCORE BY MODEL



AUC was also compared in order to select the optimal model. AUC was calculated for each class vs the rest. All models had similar AUC for adoption. The AdaBoost model performed worse for rescue. The gradient boost and random forest models performed similarly well. The gradient boost model was selected as the optimal model with AUC of 0.87, 0.85, and 0.80 for 'Adoption,' 'Euthanasia,' and 'Rescue' respectively.

FIGURE VI: AUC PLOT BY OUTCOME CLASS



Analysis of the permutation importance of the features indicated that sex, intake condition, age on intake, and intake type are the most significant features.

TABLE II: FEATURE PERMUTATION IMPORTANCE

Feature	Mean Permutation Importance	STD
Sex_Neutered	0.13	0.003724
Sex_Spayed	0.11	0.005842
Intake Condition_NORMAL	0.07	0.007705
Age on Intake	0.04	0.007729
Intake Type_STRAY	0.03	0.005278

5 Discussion

Proposed Timeline:

- Week 1: Project Proposal and Data Preprocessing
- Week 2: Model Building and Evaluation
- Week 3: Interpretation and Final Report

Prior to beginning the project, two potential challenges were identified. The first was the imbalance between outcome classes. EDA confirmed that euthanasia occurred least frequently. This challenge was addressed by upsampling the minority class (euthanasia). The second challenge was features with many subcategories with few observations. This was resolved by grouping similar subcategories together during data preprocessing.

During the modeling stage, 'Return to Owner' was dropped from analysis based on model performance. It was frequently misclassified as 'Adoption' and 'Rescue' and significantly reduced the efficacy of the models. EDA revealed this class of dogs had the shortest median shelter stay, which means they used relatively fewer shelter resources than dogs in the other classes.

The success of the model indicates that it is possible to predict outcomes when a dog enters the shelter. Whether a dog is spayed or neutered was the most influential feature for prediction. Shelters may be able to improve outcomes by prioritizing spaying or neutering intact dogs that enter the shelter.

If a dog's predicted class is euthanasia or rescue, the shelter can act on that information rather than waiting to see if the dog is adopted. That action may be promoting the dog's chances of adoption (features on social media, community events, adoption fee discounts), finding a rescue placement earlier, or finding alternate placements like foster. By predicting the outcome for dogs upon entering the shelter, shelters can make data-driven decisions about handling each dog's case to find its optimal outcome.

6 Conclusion

In summary, the proposed project aimed to accomplish two goals: accurate predictions of post-shelter outcomes and identification of the most influential features for predicting outcomes. Multiple models were evaluated to determine optimal outcome predictions. Gradient boosting and random forest models were the most successful models. The gradient boost model was selected for the final analysis. The model effectively predicts outcomes for dogs using features available on intake to the shelter. It achieved AUC >= .8 for each outcome class. Age, sex intake type are the most influential predictors for outcomes.

Further work can be done on the features included in this analysis. The permutation importance results suggest that the subcategories within the features should be studied further. The subcategories of spayed and neutered were important, but not male and female intact dogs. This suggests that the significance may be in spay/neuter status rather than in the dog's sex. Similarly, whether a dog had a normal intake condition was important, but each of the other subcategories was not. Further feature engineering may reveal more information about which characteristics are predictive of shelter outcomes.

This analysis was performed on data from a single shelter, so another area of further study would be to test the performance of the model on data from other shelters.

ACKNOWLEDGMENTS

I thank the Long Beach Animal Care Services for providing access to the data used in this project.

REFERENCES

[1] Janae Bradley and Suchithra Rajendran. 2021. Increasing adoption rates at animal shelters: a two-phase approach to predict length of stay and optimal shelter allocation. In *BMC Vet Res.* 17, 1 (Feb. 2021), 70. doi:10.1186/s12917-020-02728-2

[2] CJ Cain, KA Woodruff, DR Smith. 2020. Phenotypic characteristics associated with shelter dog adoption in the United States. In *Animals (Basel)* 10, 11 (Oct. 2020), 1959. doi: 10.3390/ani10111959

[3] S Hawes, J Kerrigan, and K Morris. 2018. Factors information outcomes for older cats and dogs in animal shelters. In *Animals (Basel)* 8, 3 (Mar. 2018), 36. doi: 10.3390/ani8030036

[4] Betty McGuire, Jordan Chan, Kennedy Jean-Baptiste, Philippa Kok, and Emma Rosenbaum. 2021. Results of behavioral evaluations predict length of stay for shelter dogs. In *Animals* 11, 11 (Nov. 2021), 3272. doi: 10.3390/ani11113272

[5] Lauren Powell, Chelsea Reinhard, Donya Satriale, Margaret Morris, James Serpell, and Brittany Watson. 2021. Characterizing unsuccessful animal adoptions: age and breed predict the likelihood of return, reasons for return, and post-return outcomes. In *Scientific Reports* 11, 1 (April 2021), 1. doi: 10.1038/s41598-021-87649-2

[6] Cem Sazara and Xin Gao. 2022. Predicting animal shelter pet adoption times and feature importance analysis using CatBoost. In *IEEE 11th International Conference on Intelligent Systems (IS),* October 2022, Warsaw, Poland. doi: 10.1109/IS57118.2022.10019608

[7] Shelter Animals Count. 2024 Mid-year analysis: Comparing January through June to prior years. Retrieved from https://www.shelteranimalscount.org/2024-mid-year-report.

[8] Amir Zadeh, Kara Combs, Brandon Burkey, Jordan Dop, Kevin Duffy, and Nasim Nosoudi. 2022. Pet analytics: Predicting adoption speed of pets from their online profiles. In *Expert Systems with Applications* 204, 1 (May 2022), 117596. doi: 10.1016/j.eswa.2022.117596