Bellevue University

Pet Adoption Prediction

Ashley Mayo

Milestone Three, Week Two

March 27th, 2025

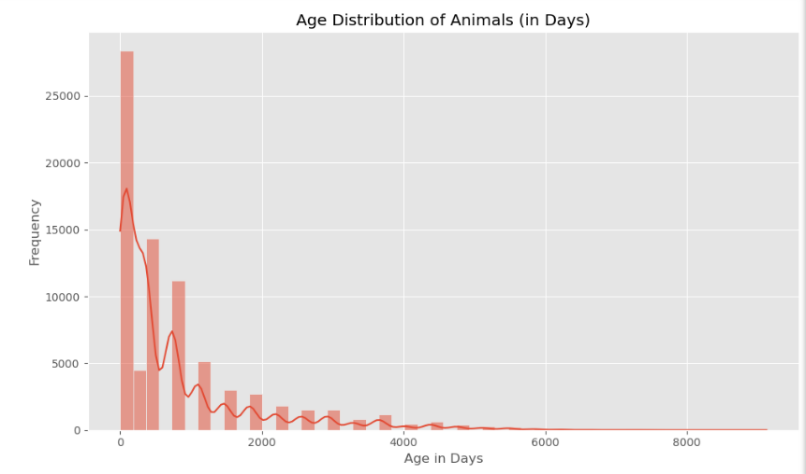
Professor Iranitalab

DSC680: Applied Data Science

**Introduction**

There is a moment in every animal shelter, the quiet pause just before a new animal is brought in. A dog abandoned on the side of the road, a cat surrendered by a family in crisis, a litter of newborn kittens left in a box by the door. These animals arrive with no voice. Some will find loving homes, embraced by families who will make them part of their world. Others, however, will not be so lucky. Some will be transferred to other facilities, others will remain in kennels for months, and for a heartbreaking few, the journey will end too soon.

The story behind this project is one that begins with a simple but powerful question: what makes the difference between an adopted animal and one that isn’t?

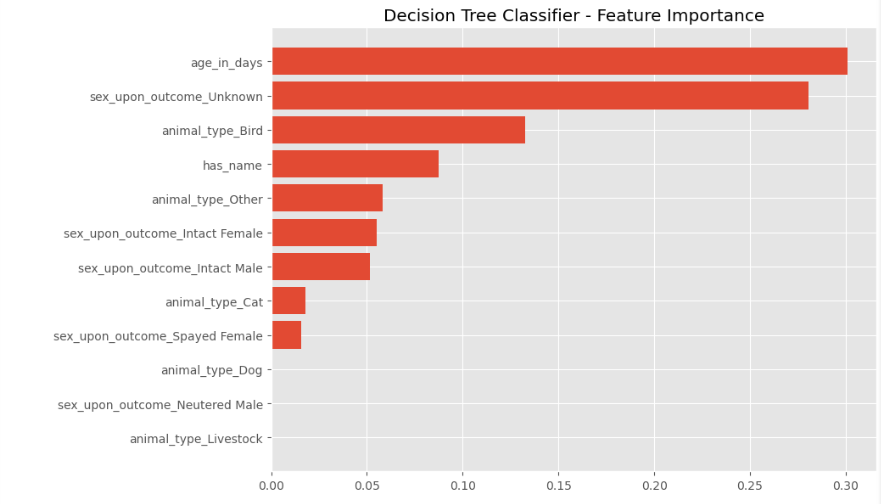
The Pet Adoption Prediction project was created in response to this reality, to use data science not only to tell these stories, but to change them. The question at the heart of this work is both simple and profound: what determines whether an animal is adopted or not? Through the lens of structured shelter data and analytical tools, this project aims to help animal shelters uncover the hidden patterns that lead to different outcomes. By identifying those patterns early, shelters can act with foresight instead of hindsight, and help more animals find a second chance at life.

A graph showing a number of different types of outcomes

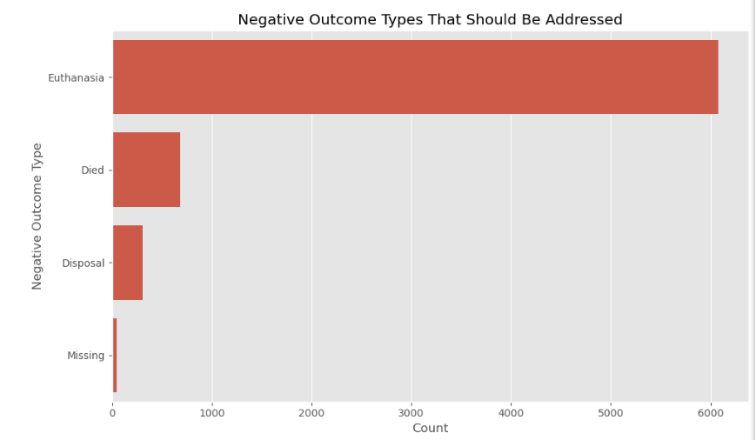
AI-generated content may be incorrect.Using data from the Austin Animal Center, the largest no-kill shelter in the United States, the project explored the dynamics of shelter outcomes (Austin Animal Center, n.d.). The dataset includes thousands of records, each containing detailed information about intake type, animal characteristics, health condition, age, sex, and outcome. Every entry represents a living being whose future depends on a series of decisions, by people, by processes, and sometimes by circumstance. The richness of this dataset offered a powerful foundation for discovering what factors truly shape an animal’s journey through the shelter system.

**Data Exploration**

Exploratory analysis revealed that age plays a central role in adoption outcomes. Most shelter animals are relatively young, yet younger age does not guarantee adoption. Notably, animals with names were adopted at significantly higher rates than those without. This subtle, humanizing detail, a name, transforms an animal from an anonymous figure into a personality. As noted by the ASPCA, naming shelter animals helps potential adopters form a bond; “It’s a lot easier to fall in love with ‘Snowflake’ than ‘Cat Number 3,298’” (PetMD, 2017).

To dive deeper into these patterns, I used decision tree and logistic regression models, selected for their balance of performance and interpretability. While I initially considered adding XGBoost, I encountered persistent technical issues that limited its implementation. Nevertheless, the decision tree model provided strong insights into the hierarchy of predictive features, while logistic regression offered clarity in understanding the relationships between variables. Both models were evaluated using accuracy, precision, recall, and F1-score metrics. The confusion matrix for the decision tree classifier showed a solid ability to distinguish adoption outcomes, with some understandable overlap in categories like transfers and returns. The feature importance rankings revealed that age in days was the most influential factor, followed by sex status, animal type, and whether the animal had a name.

Beyond the data and models, the project highlighted a deeper, more sobering narrative—the reality of unfavorable outcomes in animal shelters. Not every animal leaves through the front door with a new family. Many leave through the back, transferred indefinitely, returned repeatedly, or tragically euthanized. The data made these outcomes impossible to ignore. Euthanasia remains a significant concern; according to the ASPCA, approximately 920,000 shelter animals are euthanized each year in the United States (ASPCA, n.d.). Behind each count in the dataset is a life that deserved better. These are the stories that demand our attention just as much as successful adoptions. Every number in a bar chart could have been a wagging tail or a gentle purr, had things gone differently.

The purpose of this model is not to label or judge but to assist. It is a tool for early intervention, helping shelters identify the animals who are most vulnerable before it’s too late. When used with compassion and care, these insights can help rewrite endings before they’re written.

No algorithm can determine the worth of a life. These models are designed not to replace human judgment, but to enhance it. We must recognize and mitigate risks of algorithmic bias, especially when certain traits or breeds are overrepresented in negative outcomes. Transparency, fairness, and empathy must guide how these tools are implemented. Shelter staff must be empowered to understand not just what the model predicts, but why it predicts it.

Class imbalance was a significant hurdle, with adoption cases vastly outnumbering categories like “missing” or “disposal.” Data inconsistencies required careful cleaning, and balancing model complexity with interpretability was essential. But these challenges also reflect the reality of working with real-world data, it is messy, imperfect, and meaningful.

Looking ahead, this work has the potential to be expanded across shelters of all sizes. A simple dashboard based on this model could provide real-time recommendations, helping staff prioritize outreach for older animals, those without names, or those with specific health conditions. It could inform intake assessments, reduce unnecessary transfers, and highlight opportunities for behavioral support or foster placement. Even in its current form, the model provides enough insight to inform meaningful operational changes.

The recommendation going forward is to integrate this predictive model gradually and intentionally into shelter operations, ensuring that staff have both the training and context to use the tool effectively. A successful implementation begins with building awareness among shelter teams, helping them understand how data-driven insights can enhance, rather than replace, their intuition and experience. A pilot program can be introduced using a small set of shelter records, where predictions are reviewed alongside staff decisions to compare alignment and identify areas of strength or concern.

Shelters should begin by using the model to flag animals with high-risk profiles early in the intake process. This doesn’t mean making decisions solely based on a prediction score—but rather using that score as a starting point for discussion and planning. For example, an older animal without a name and with uncertain medical status may be less likely to be adopted. Being ahead of time allows staff to act; assigning a name, creating a social media story, prioritizing foster placement, or engaging in behavior enrichment that increases adoptability (Weiss et al., 2012).

The model’s output can be integrated into a user-friendly dashboard that presents predictions in simple terms, color-coded risk indicators, adoption probability scores, and suggested action prompts. Tools like this have already shown promise in clinical settings and can be adapted for animal shelters to drive faster and more humane outcomes (Drazenovich et al., 2020). Even a simple alert system that draws attention to animals trending toward negative outcomes can lead to faster intervention, more targeted attention, and potentially life-saving changes in approach.

Another important recommendation is to leverage the model to inform broader shelter policies and programming decisions. Over time, aggregate model predictions can reveal systemic patterns, such as the kinds of animals most likely to be transferred, the conditions that often precede euthanasia, or the intake sources that correlate with longer stays (Mohan-Gibbons et al., 2014). This information can guide resource allocation, staff training programs, partnership development, and even community outreach efforts aimed at preventing surrenders in the first place.

The overarching message is this: technology should amplify empathy, not replace it. When used thoughtfully, predictive analytics can support decision-makers in identifying vulnerable animals and shaping interventions that maximize the chance of positive outcomes (Bollen, 2019). These tools do not replace human compassion; they enhance it when guided by ethical awareness and transparent practices.

**Conclusion**

This paper set out to answer a critical question: what determines whether an animal in a shelter is adopted or not? Through data-driven analysis of thousands of records from the Austin Animal Center, we identified key predictors of adoption, such as age, sex status, and whether the animal had a name. Both the decision tree and logistic regression models provided actionable insights, showing that even small details, like giving an animal a name, can significantly impact its outcome.

Beyond the technical results, this project highlights the power of data science to inform compassionate decision-making. By identifying patterns early, shelters can intervene more effectively and increase the likelihood of adoption for vulnerable animals. While no model can replace human empathy, it can support the work of those who dedicate their lives to saving others.

**Questions an audience would ask:**

1. What were the most informative data fields used in your model?
2. Did you find any patterns in the data that were surprising?
3. Why did you choose the Austin Animal Center dataset?
4. Why was 'name' such an impactful variable?
5. What insights from the data could immediately help shelters?
6. How did data quality issues affect your modeling process?
7. Why is it important to analyze shelter data at this level?
8. How did you ensure the dataset represented different types of animals fairly?
9. How did you determine which features to include in your model?
10. Did you analyze trends in adoption by sex or neuter status?

**References**

ASPCA. (n.d.). *Pet statistics*. American Society for the Prevention of Cruelty to Animals. https://www.aspca.org/animal-homelessness/shelter-intake-and-surrender/pet-statistics

Austin Animal Center. (n.d.). *Animal Center Open Data Portal*. City of Austin. https://data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Outcomes/9t4d-g238

Bollen, J. (2019). *How predictive analytics can support decision-making in social services*. RAND Corporation. https://www.rand.org/pubs/research\_reports/RR2843.html

Drazenovich, N., Gunter, L. M., & Barber, R. T. (2020). Predicting canine adoption using animal shelter behavioral and demographic data. *Frontiers in Veterinary Science*, *7*, 355. https://doi.org/10.3389/fvets.2020.00355

Mohan-Gibbons, H., Weiss, E., & Slater, M. (2014). Preliminary investigation of food guarding behavior in shelter dogs in the United States. *Animals*, *4*(3), 563–588. https://doi.org/10.3390/ani4030563

PetMD. (2017). *Why giving shelter pets names can help them find homes*. https://www.petmd.com/news/view/why-giving-shelter-pets-names-can-help-them-find-homes-36090

Weiss, E., Gramann, S., Spain, C. V., & Slater, M. (2012). Goodbye to a good friend: An exploration of the rehoming of cats and dogs in the U.S. *Open Journal of Animal Sciences*, *2*(2), 111–126. https://doi.org/10.4236/ojas.2012.22017