

Austin AniML Rescue



Adopt



Foster



Volunteer



Donate

Purpose

Help the Austin Animal Center better understand information on their animals



Data Source Description

Sources:

- City of Austin public data portal
- Two csv files, updated daily

CSV files:

Intakes = Animals coming to the shelter

Outcomes = Animals leaving the shelter

Outcomes



Intakes

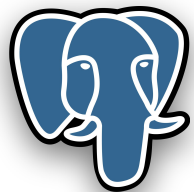


Research Question

Can we predict the outcome for an animal based on other characteristics?



Technologies and tools



SQL



dmlc
XGBoost

github.com/ilaha/Austin_AnimL_Rescue

- Data Files
- Exploratory Data Analysis
- SQL Database Code
- Machine Learning Model
- Tableau Dashboard
- Readme.md



*GitHub
Public Repo*

Austin AniML Rescue



Database

Original csv files

- *Intakes*
- *Outcomes*

Each \approx 130,000 rows



Drop duplicate
animal_id



Each \approx 105,000 rows

PostgreSQL Database

Intakes	
<i>animal_id</i>	varchar
animal_name	varchar
intake_date	date
intake_date_2	date
found_location	varchar
intake_type	varchar
intake_condition	varchar
animal_type	varchar
sex_upon_intake	varchar
age_upon_intake	varchar
breed	varchar
color	varchar

Outcomes	
<i>animal_id</i>	varchar
animal_name	varchar
outcome_date	date
outcome_date_2	date
date_of_birth	date
outcome_type	varchar
outcome_subtype	varchar
animal_type	varchar
sex_upon_outcome	varchar
age_upon_outcome	varchar
breed	varchar
color	varchar

Queries

SELECT
columns

INNER JOIN
animal_id



SQLAlchemy



Python

Austin AniML Rescue



**Data Preparation &
Exploratory Analysis**

Data exploration: Original data

Data Types:

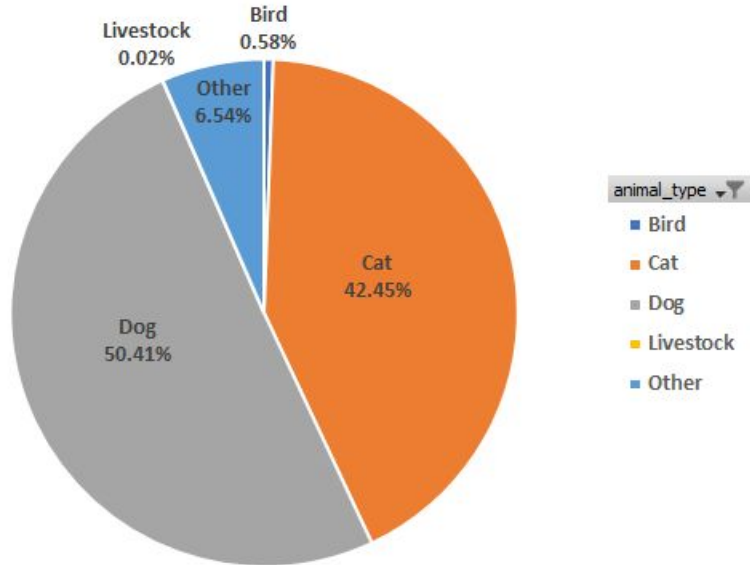
Dates

Categorical Text Strings

	animal_id	animal_name	animal_type	breed	color	intake_type	date_of_birth	intake_date	found_location	intake_condition	sex_upon_intake
0	A786884	*Brock	Dog	Beagle Mix	Tricolor	Stray	2017-01-03	2019-01-03	2501 Magin Meadow Dr in Austin (TX)	Normal	Neutered Male
1	A682524	Rio	Dog	Doberman Pinsch/Australian Cattle Dog	Tan/Gray	Stray	2010-06-29	2014-06-29	800 Grove Blvd in Austin (TX)	Normal	Neutered Male
2	A696408	*Pearl	Dog	Chihuahua Shorthair	Tricolor	Stray	2013-02-04	2015-02-04	9705 Thaxton in Austin (TX)	Normal	Intact Female
3	A736287	*Twilight	Cat	Domestic Shorthair Mix	Torbie	Stray	2016-08-08	2016-10-08	South First And Stassney in Austin (TX)	Normal	Intact Female
4	A810994	NaN	Other	Bat	Brown	Wildlife	2017-12-24	2019-12-25	7900 Rm 1826 Rd in Travis (TX)	Normal	Unknown

Data exploration: Animal Types

Intake: Animal Types (N = 105,367)



Animal Type	Number of Unique Breeds
Bird	79
Cat	98
Dog	2180
Livestock	7
Other	121

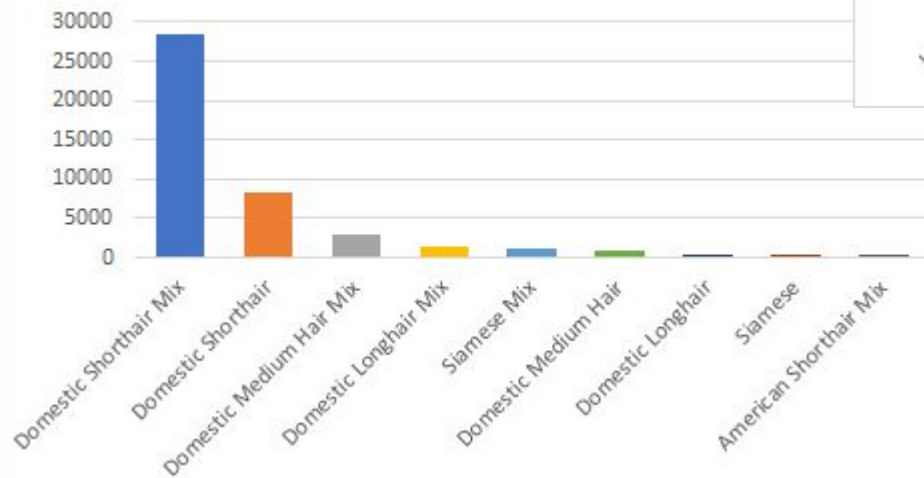
Data Preparation

- Converted dates from string to datetime
- Dropped unnecessary columns and fields
- Used OneHotEncoder to encode categorical features

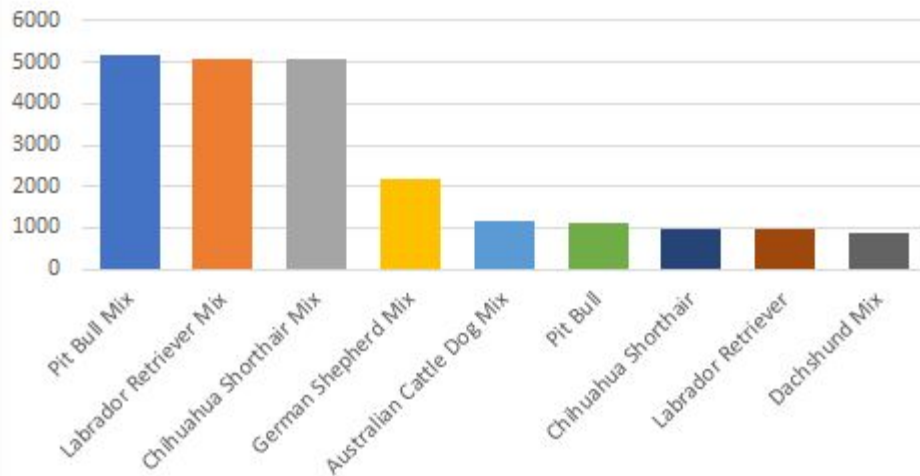


Data Exploration: Animal breed

Cat Breeds



Dog Breeds



Labrador
Retriever/Pit Bull

Pit Bull/Labrador
Retriever

Labrador
Retriever/Pit Bull

Data Exploration: Bin Outcome Types

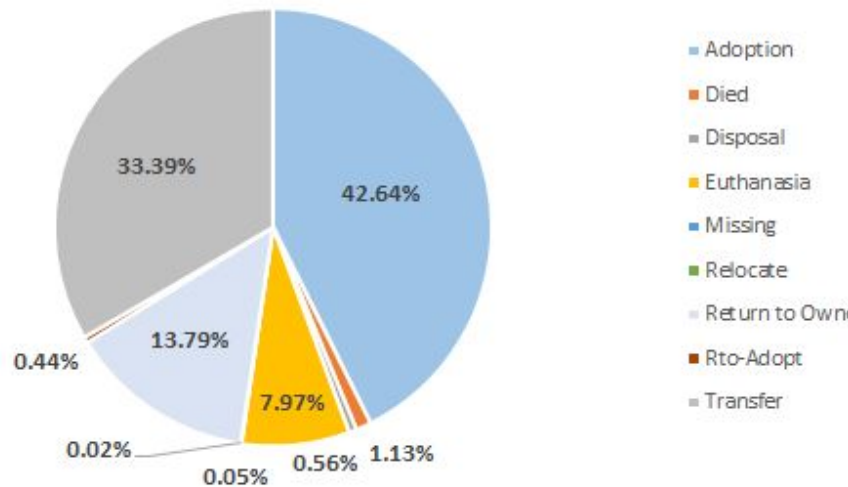
Initial Outcomes

- Adoption
- Transfer
- Return to Owner
- Euthanasia
- Died
- Disposal
- Rto-Adopt
- Missing
- Relocate

Binned Outcomes

- Adoption
- Transfer
- Euthanasia/Died
- Other

Outcome Types (N = 105,349)



Austin AniML Rescue



Machine Learning Model & Data Analysis

Machine Learning Model

Classification Models Tested for All Outcomes

- Random Forest
- AdaBoost
- Neural Nets
- CatBoost
- XGBoost

Classification Models Tested for Binary Outcomes

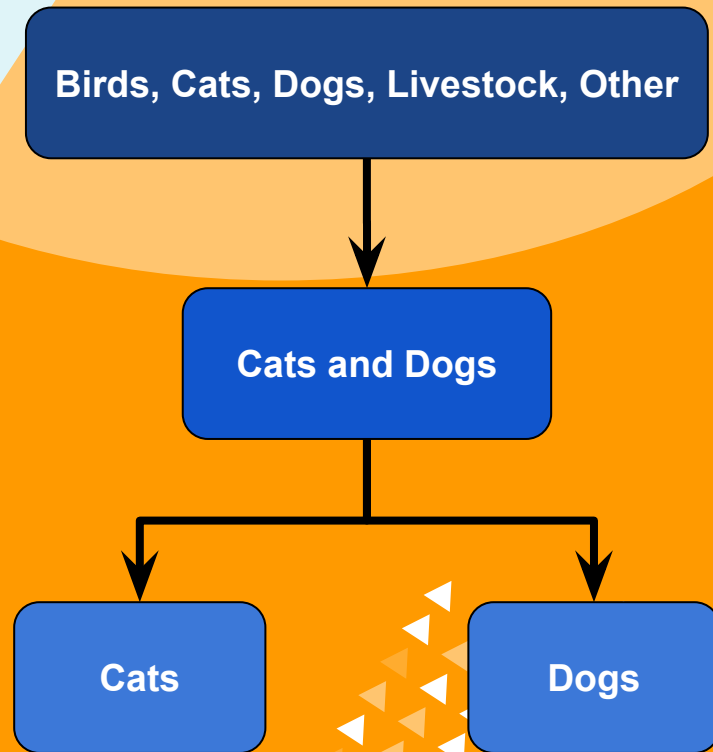
- Logistic Regression
- Support Vector Machine (SVM)
- XGBoost

Other Techniques Used

- Principal Component Analysis (PCA)
- Cross-Validation (5 cycles)
- Resampling
 - Undersampling
 - Oversampling
 - SMOTE-ENN

Feature Reduction: Reduce the Test Sample

- Tested Cats and Dogs only
- Tested Cats and Dogs separately
- Cut out features based on initial model results
 - Top 50 and Top 20 features
 - Retrain XGBoost with just top features



Feature Reduction: Principal Component Analysis

Principal Component Analysis

- reduce dimensionality
- tested reducing features into different numbers of components (range: 2 - 500)

```
# Get weights of each feature
pca.explained_variance_ratio_
```

```
array([0.00691358, 0.00573547, 0.00517823, 0.00498553, 0.00476146,
       0.00432708, 0.00429765, 0.00404978, 0.0039602 , 0.0037923 ,
       0.0037695 , 0.00375394, 0.00366746, 0.00361406, 0.00358121,
       0.00339461, 0.00337121, 0.00333684, 0.00326007, 0.00325312,
       0.003233 , 0.00322205, 0.00318446, 0.00315718, 0.0031251 ,
       0.00309489, 0.00307662, 0.00305787, 0.00302912, 0.00297431,
       0.00297108, 0.00294539, 0.00293849, 0.00292763, 0.00292547,
       0.00292012, 0.0029075 , 0.00289275, 0.00288294, 0.00288045,
       0.0028758 , 0.0028648 , 0.00284679, 0.002839 , 0.0028374 ,
       0.00281911, 0.00281174, 0.00280317, 0.00277887, 0.00277451,
       0.00276804, 0.00276527, 0.00275451, 0.0027435 , 0.00273593])
```

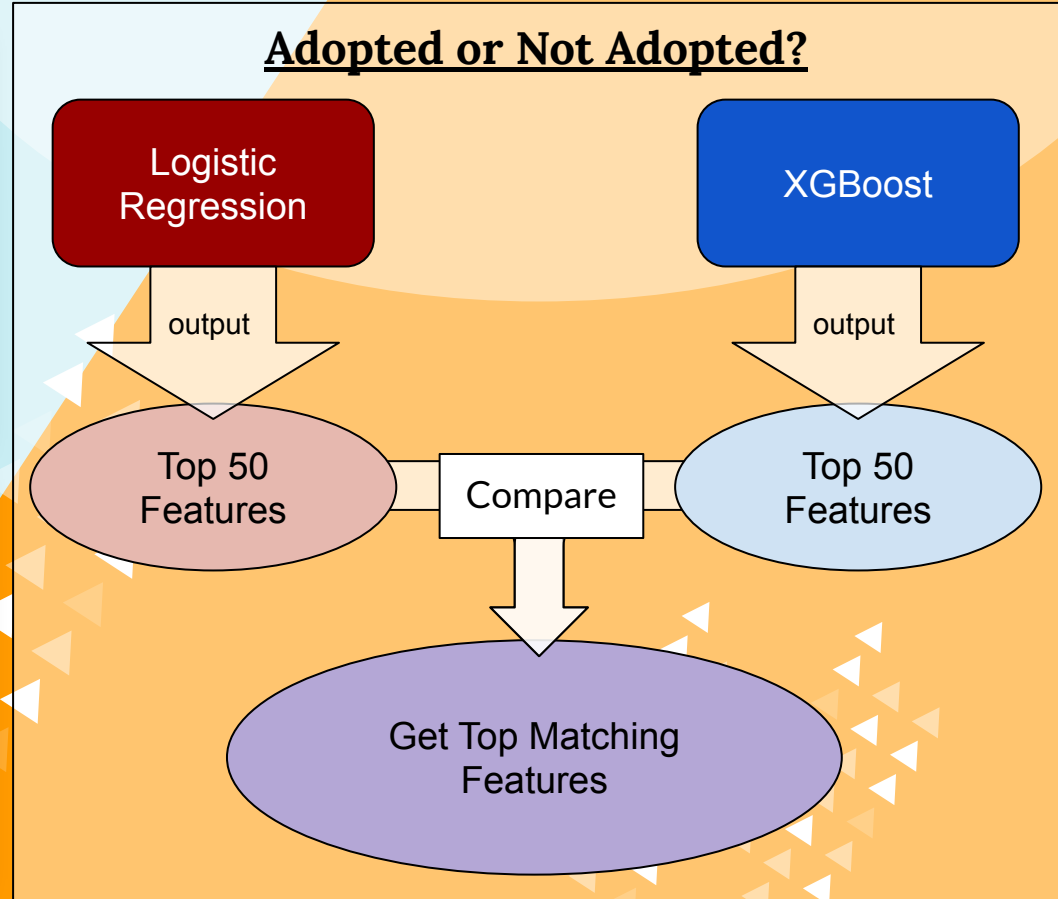

Finding Most Predictive Features

Classification Models Tested for Binary Outcomes

- Logistic Regression
- Support Vector Machine (SVM)
- XGBoost

Find Top Features that Influence These Outcomes

- Adopted? (Yes/No)
- Euthanized/Died? (Yes/No)
- Transferred? (Yes/No)



Finding Most Predictive Features

Resampling Euthanasia/Died

- Euthanasia/Died category had less than half the amount that Adoption & Transfer categories had
- Methods Tested
 - Undersampling
 - Oversampling
 - SMOTEENN
- Cats - resampling didn't help much
- Dogs - undersampling helped!

Dogs Euthanasia/Died Classification Report 4

	precision	recall	f1-score	support
0	0.98	0.99	0.99	7900
1	0.24	0.09	0.13	199
accuracy			0.97	8099
macro avg	0.61	0.54	0.56	8099
weighted avg	0.96	0.97	0.96	8099



Dogs Undersampling Classification Report 4

	precision	recall	f1-score	support
0	0.46	0.59	0.52	2906
1	0.73	0.61	0.67	5193
accuracy			0.60	8099
macro avg	0.59	0.60	0.59	8099
weighted avg	0.63	0.60	0.61	8099

Austin AniML Rescue



Results and Recommendations

Result of analysis

Result of Model for Cats

Cats XGBoost Classification Report 4				
	precision	recall	f1-score	support
Adoption	0.71	0.70	0.70	3225
Euthanasia/Died	0.35	0.29	0.31	431
Other	0.00	0.00	0.00	60
Transfer	0.67	0.71	0.69	3214
accuracy			0.67	6930
macro avg	0.43	0.42	0.43	6930
weighted avg	0.66	0.67	0.67	6930

Result of Model for Dogs

Dogs XGBoost Classification Report 4				
	precision	recall	f1-score	support
Adoption	0.68	0.91	0.78	3894
Euthanasia/Died	0.17	0.12	0.14	117
Other	0.00	0.00	0.00	17
Transfer	0.55	0.22	0.31	2076
accuracy			0.66	6104
macro avg	0.35	0.31	0.31	6104
weighted avg	0.62	0.66	0.61	6104

Analysis

Top Predictive Features for Cats by Outcome

Top Features for Cats: **Adoption**

sex_upon_intake_Unknown
intake_condition_Normal
age_upon_intake_in_days
sex_upon_intake_Spayed Female
color_Gray Tabby
intake_type_Owner Surrender
breed_American Shorthair Mix
sex_upon_intake_Neutered Male
color_Gray/White
color_Orange
color_Gray
color_Calico/White

Top Features for Cats: **Euthanasia/Died**

intake_condition_Normal
intake_condition_Injured
intake_condition_Nursing
intake_condition_Sick
sex_upon_intake_Unknown
color_Tortie/Black
age_upon_intake_in_days
color_Cream Tabby
breed_American Shorthair Mix
breed_Domestic Medium Hair Mix
breed_Domestic Shorthair Mix
intake_type_Euthanasia Request
breed_Domestic Longhair Mix
color_Torbie/White
intake_condition_Other
color_Seal Point
color_Cream

Top Features for Cats: **Transfer**

sex_upon_intake_Unknown
intake_condition_Injured
sex_upon_intake_Spayed Female
age_upon_intake_in_days
color_Gray/White
color_Brown Tabby/Black
breed_American Shorthair Mix

Analysis

Top Predictive Features for Dogs by Outcome

Top Features for Dogs: Adoption

sex_upon_intake_Unknown
breed_Shih Tzu Mix
breed_Shih Tzu
breed_Lhasa Apso Mix
breed_Dachshund
breed_Pug
breed_Pug Mix
breed_Pekingese
intake_type_Euthanasia Request
breed_Black Mouth Cur/Doberman Pinsch
breed_Golden Retriever
breed_Cocker Spaniel
breed_Maltese

Top Features for Dogs: Euthanasia/Died

intake_condition_Normal
intake_type_Euthanasia Request
breed_Labrador Retriever/Pointer
sex_upon_intake_Unknown
breed_Standard Schnauzer
intake_type_Stray
breed_Pit Bull Mix
age_upon_intake_in_days
color_White/Red
color_Brown Brindle/Black
color_Black Brindle/White
breed_Rottweiler Mix
breed_Shetland Sheepdog Mix
intake_condition_Injured
breed_Welsh Terrier Mix
intake_condition_Pregnant

Top Features for Dogs: Transfer

sex_upon_intake_Unknown
breed_Shih Tzu Mix
breed_Shih Tzu
breed_Lhasa Apso Mix
breed_Pug
breed_Dachshund
breed_Pomeranian Mix
breed_Pug Mix
breed_Pekingese
breed_Cocker Spaniel
breed_Golden Retriever
breed_Maltese
breed_Pekingese Mix
breed_Black Mouth Cur/Doberman Pinsch
breed_Beagle/Jack Russell Terrier

Recommendation 01

- Automatically pull new files from City of Austin public portal

Recommendation 02

- Capture data for animals currently in the shelter and show predicted outcomes for those animals on dashboard

Recommendation 03

- Continue refining model to better determine characteristics of animals that are adopted

*Recommendations
for future
analysis*



*Anything the team
would have done
differently?*

- Establish communication with the animal center
- Revised dataset to include animal's personality

- Define well a single research question rather than attempt to pursue many questions at the start





Data Dashboard: *Tableau Public*

Austin AniML Rescue



**Thank you for your
attention.**