Will you take me home?

Metis Bootcamp - Classification Module

Andy Wang

Introduction

❖Problem:

- Animal shelters have limited funding and volunteers

❖Goal:

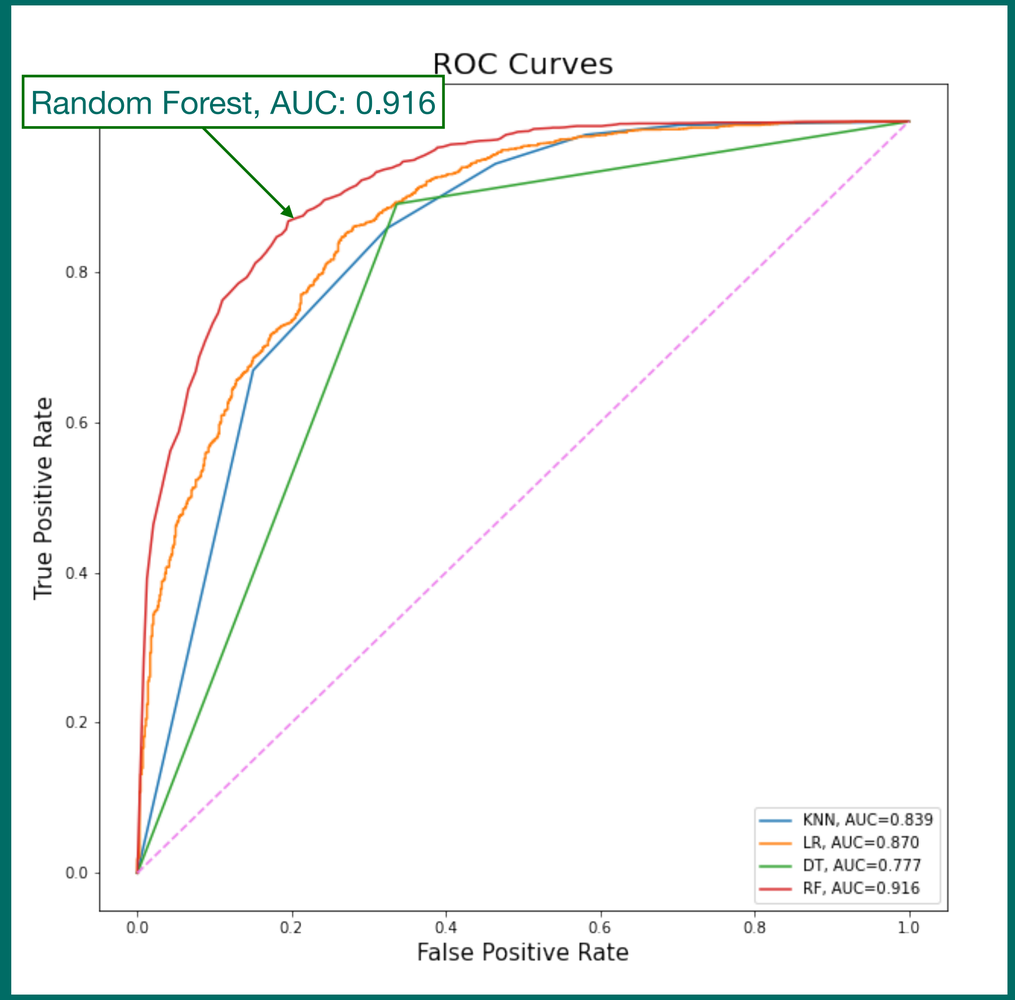
- Build a classification model to identify which features led to positive outcome (adoption or return to owner) that would assist shelter to utilize resources properly

Data

- Data from the Sonoma County Animal Services
- 15,281 observations

Methodology

- ROC AUC for model selection
  - Better distinguishes positive and negative classes
- Accuracy as performance metric
  - The features can then be utilized when the model is able to accurately predict the outcome

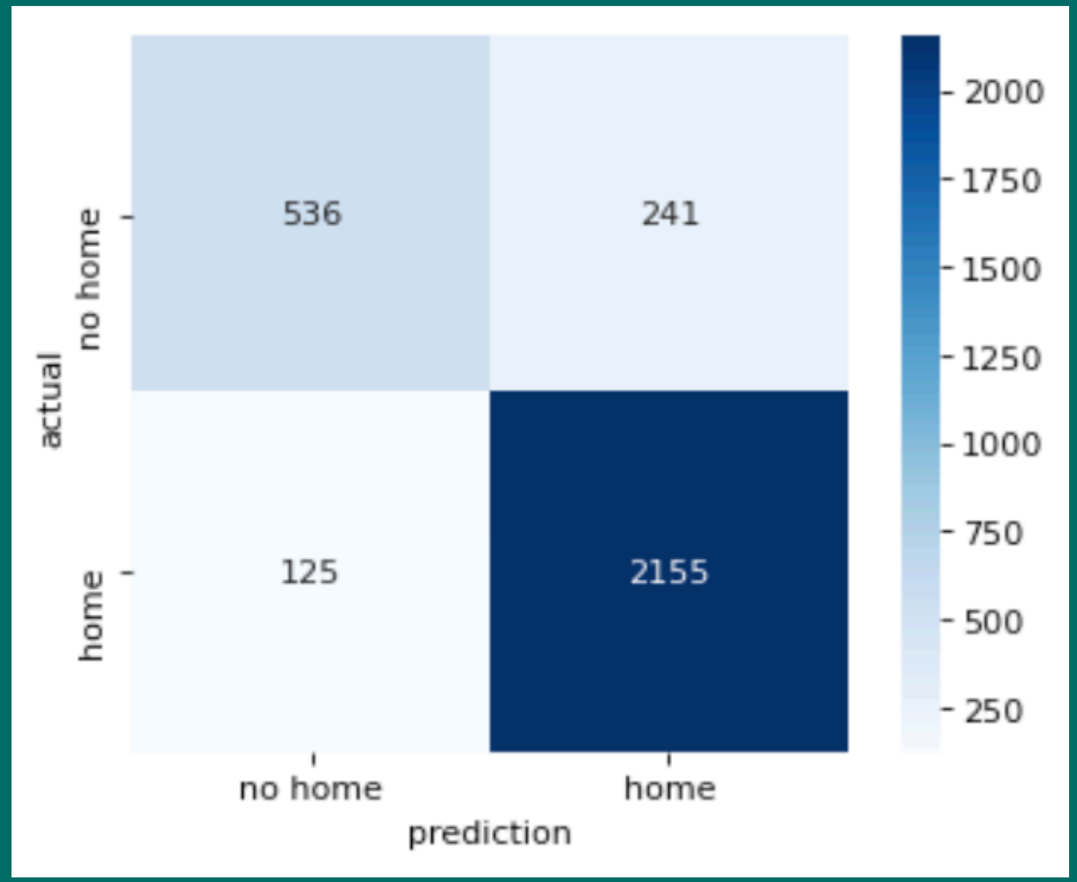


Model

❖Hyperparameter tuned

Random Forest

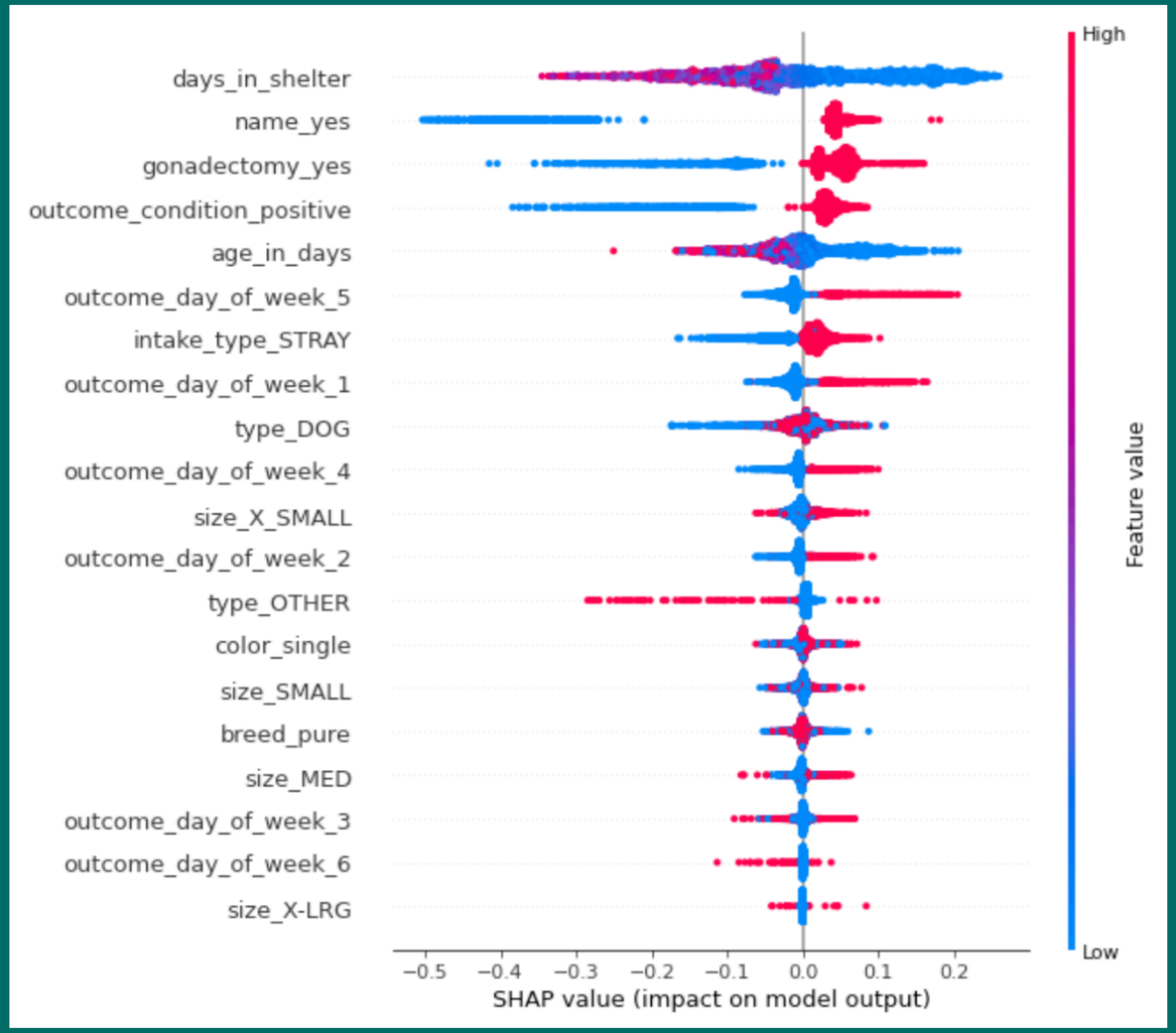
	Test
Accuracy	0.88
AUC	0.93



Feature Importance

Features:

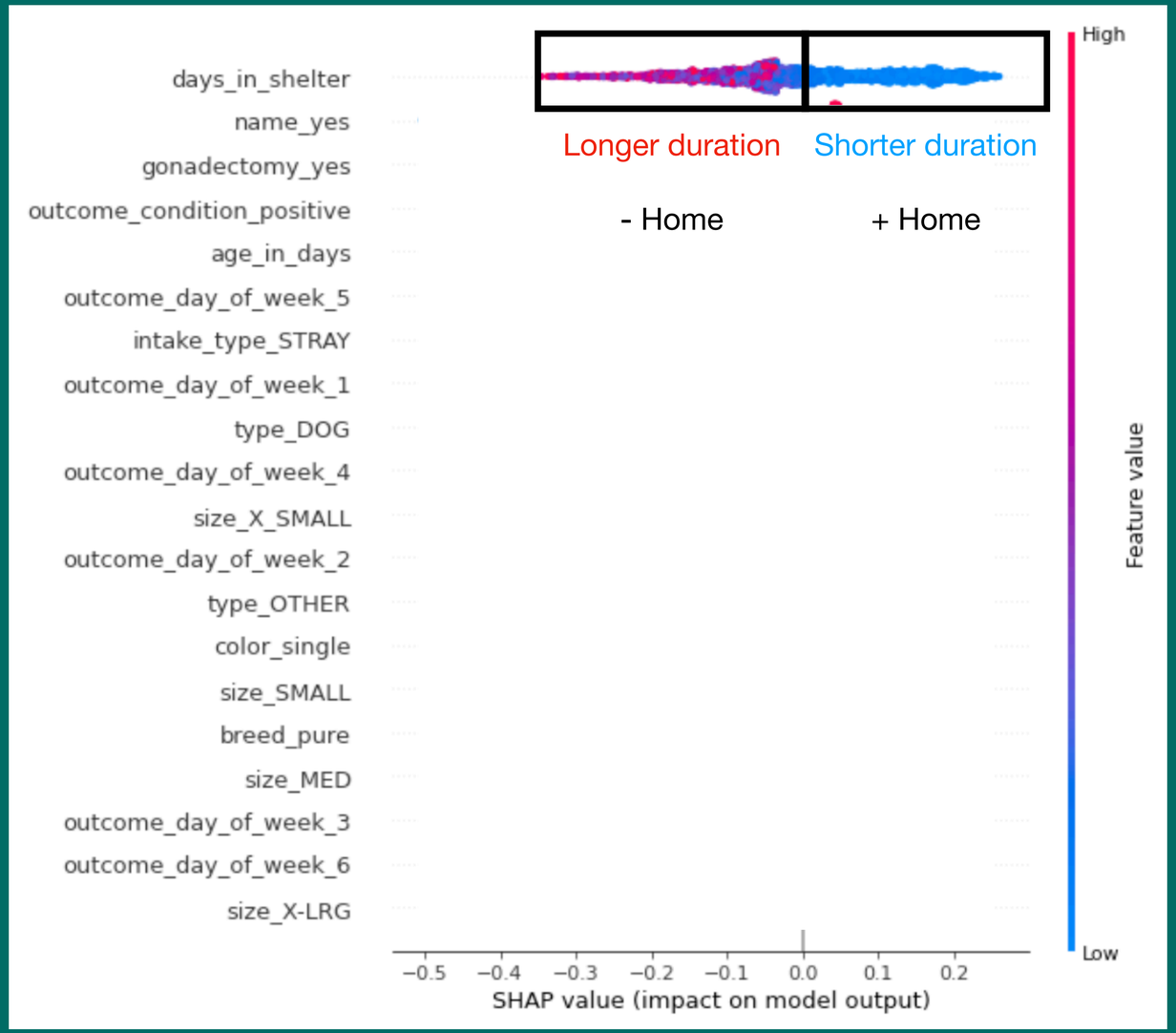
- ❖Type of animal
- ❖Intake type
- ❖Days in shelter
- ❖Outcome condition
- ❖Outcome day of week
- ❖Characteristics of animals (Name, breed, age, color, sex, size)



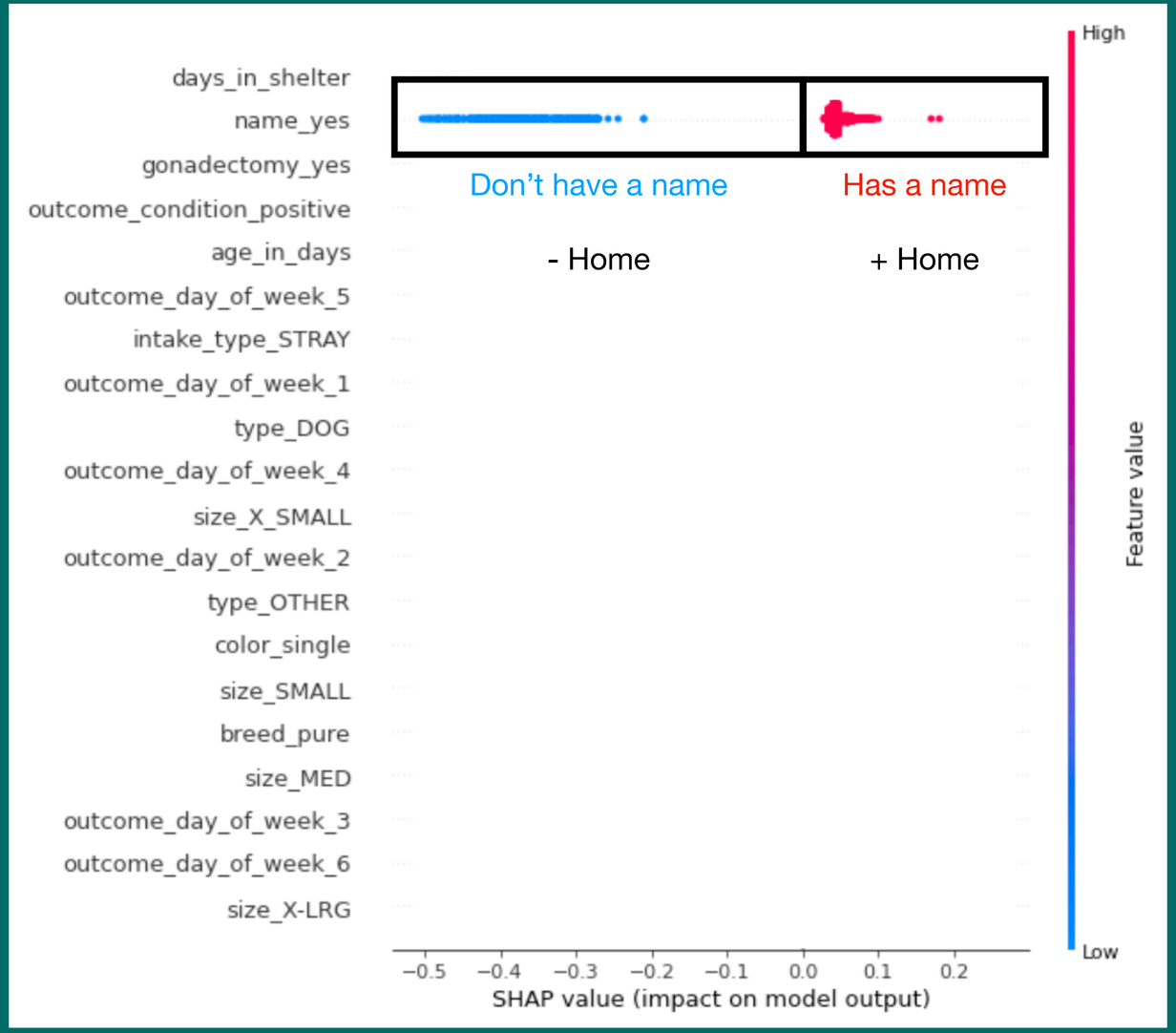
Feature Importance

Positive Outcome (Animals find homes)	Negative Outcome (Animals not find homes)
Adoption	Transfer to another shelter
Return to owner	Euthanized
	Appt made to interact with the animals
	Died
	Disposal
	Escaped/stolen

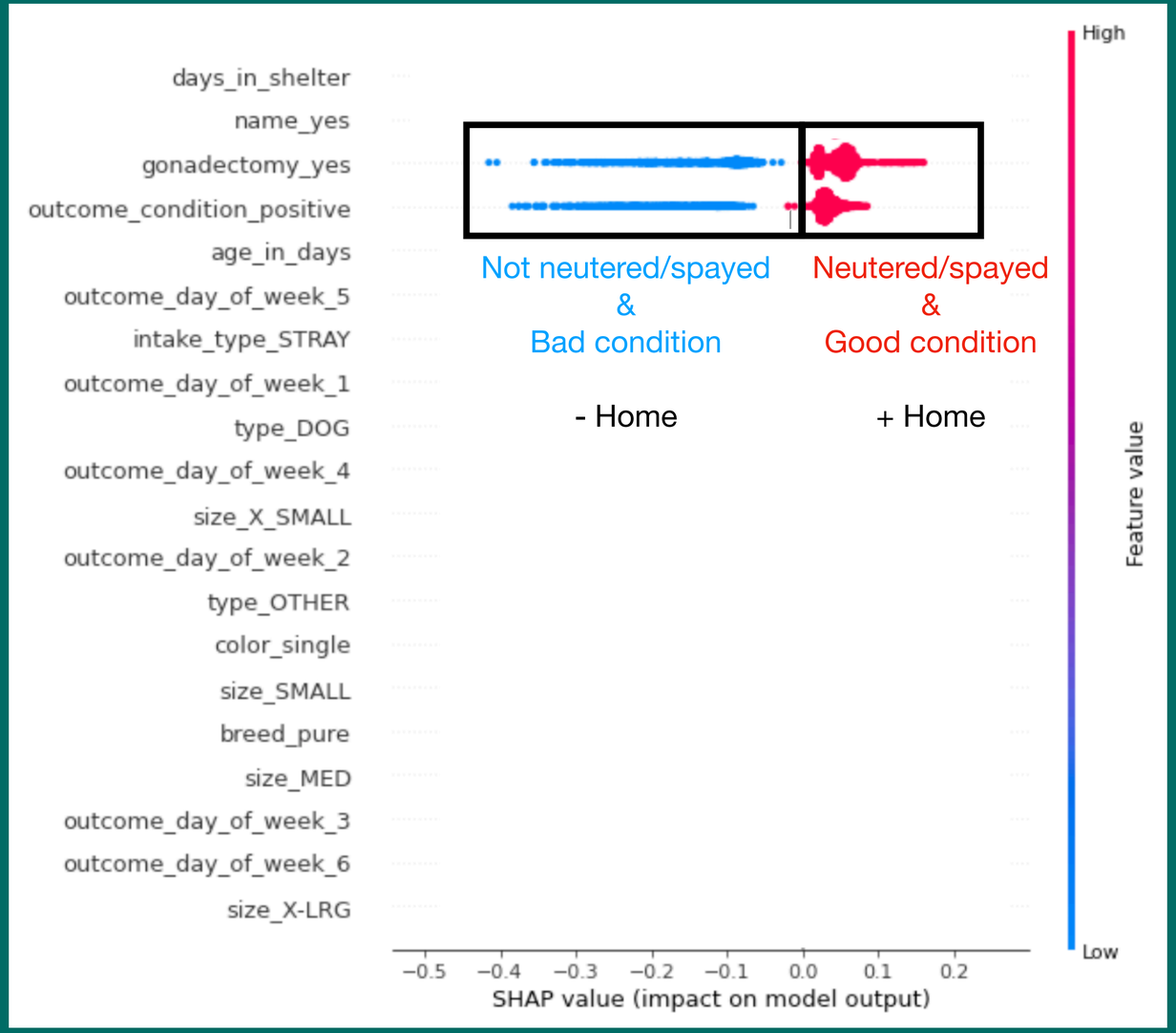
Feature Importance



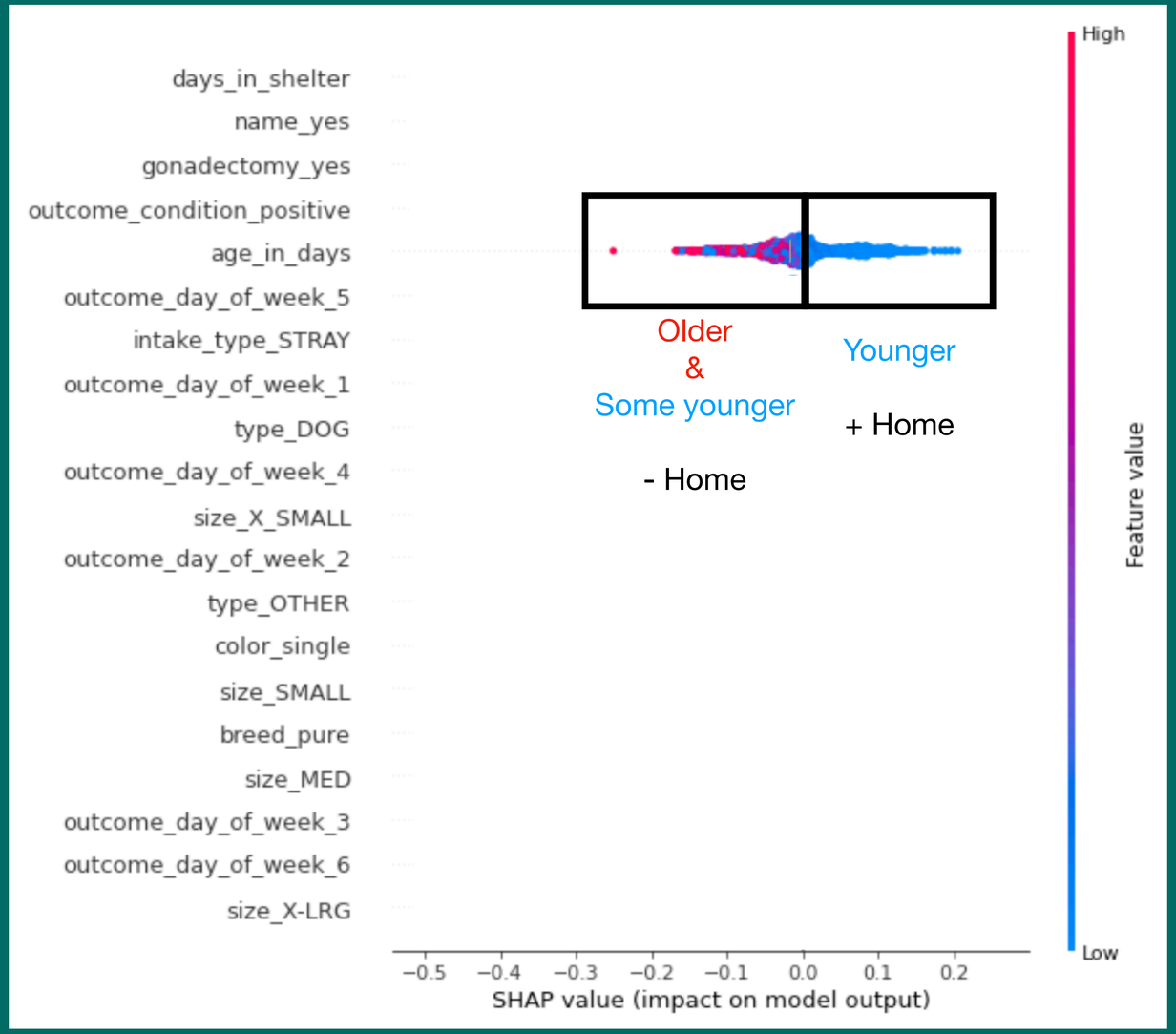
Feature Importance



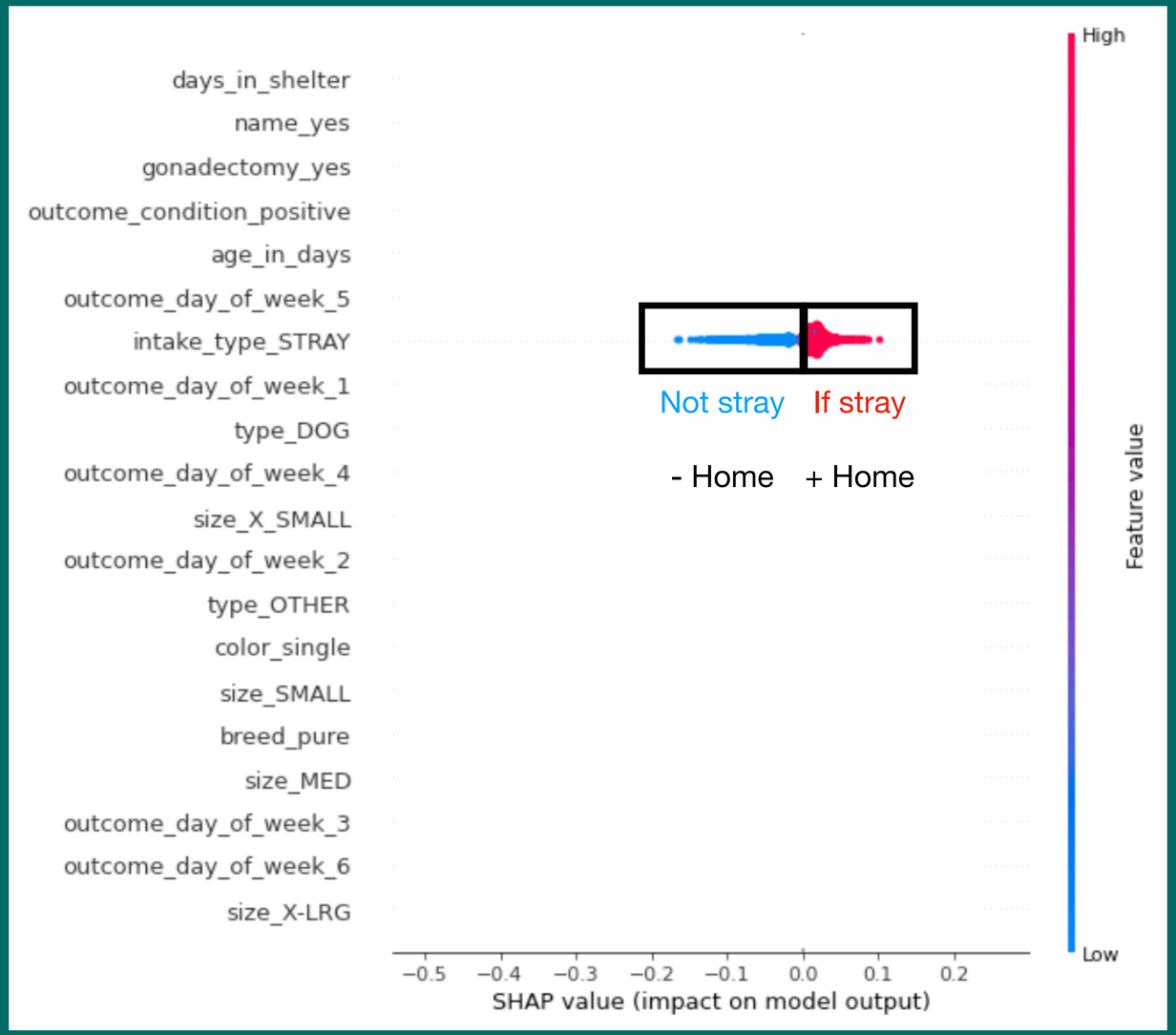
Feature Importance



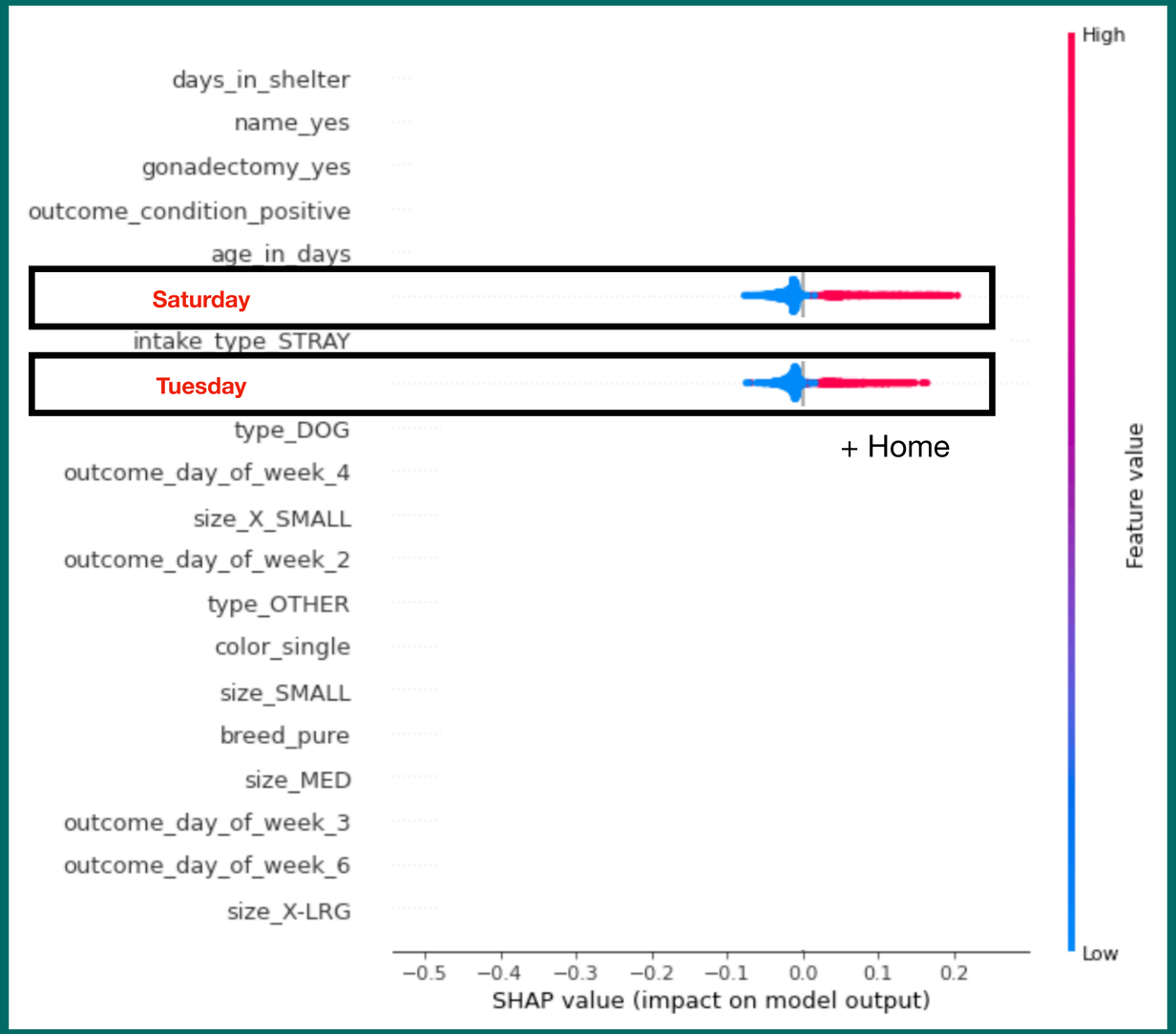
Feature Importance



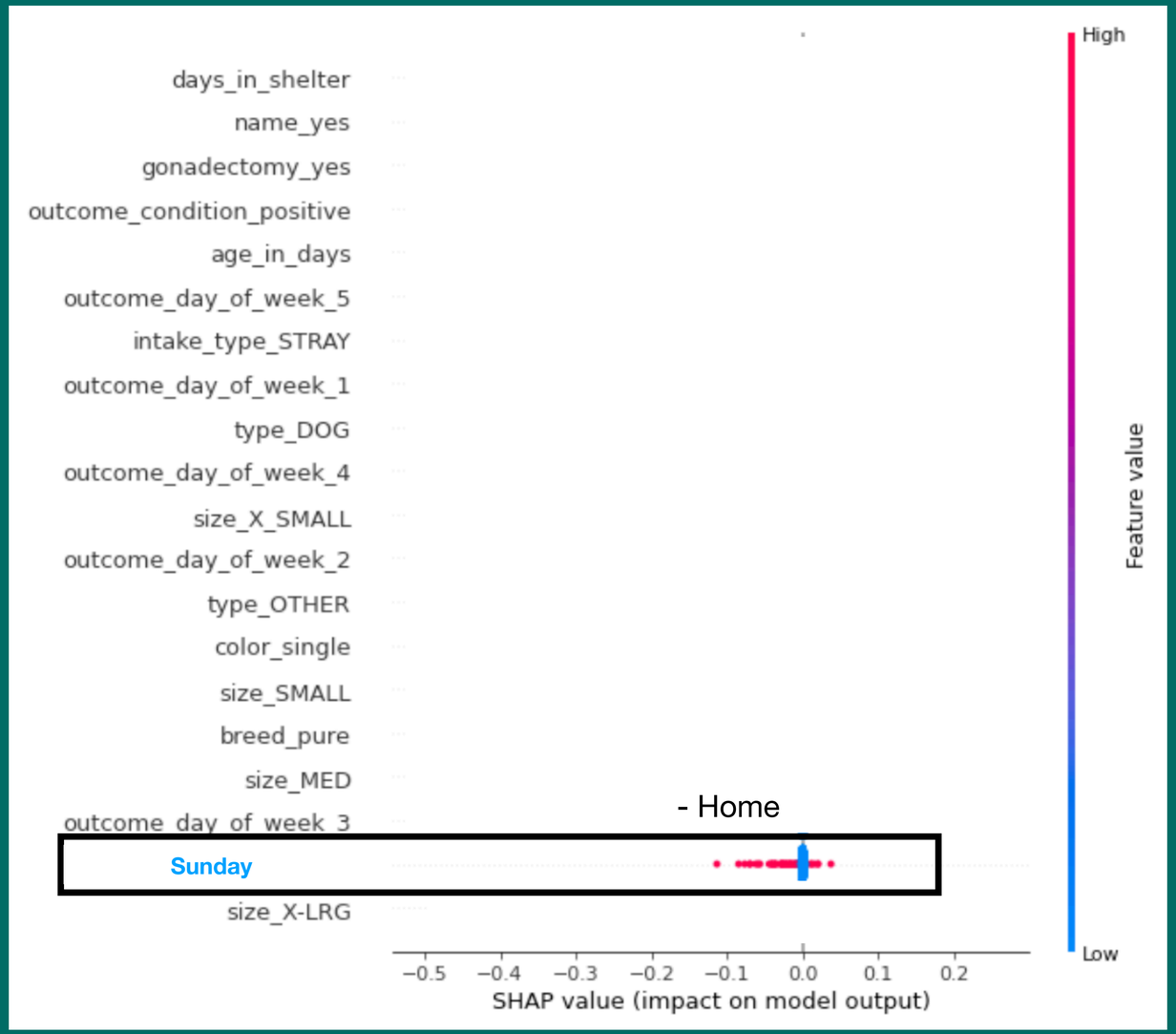
Feature Importance



Feature Importance



Feature Importance



Conclusion

❖Features that increase chance of positive outcome:

- Shorter duration in shelter
- Animals the are named, neutered/spayed, in good health condition, younger, and stray
- Tuesday and Saturday

❖Features that decrease chance of positive outcome:

- Longer duration in shelter
- Animals that are not named, neutered/spayed, and stray, and are in bad health condition and older
- Sunday

Thank you

Appendix

Baseline vs Final Model

	Baseline (Logistic regression)	**Final (Random forest)**
Accuracy	0.85	0.88
AUC	0.87	0.93