Furry Futures: Predictive Analytics for Adoption Outcomes in Animal Shelters

Tess Anderson, Dr. Robert Kelley

Bellarmine University Data Science Program

This project delves into the realm of predictive analytics to enhance the operational efficiency and welfare outcomes of a no-kill animal shelter. Leveraging a dataset encompassing over **28,000 records of cats from the Austin Animal Shelter**, this study employs a **Random Forest model** to predict the outcomes of sheltered cats with a commendable **accuracy of nearly 73%.** By scrutinizing various features such as age, breed, sex, and spay/neuter status, the model aids in foreseeing potential outcomes including adoption and transfer. The insights gleaned from this predictive model empower shelter administrators to allocate resources judiciously, optimize strategic decision-making processes, and ultimately bolster the prospects of finding permanent homes for cats, thereby **fostering a more humane and sustainable approach to animal welfare.**

OBJECTIVES

- **Predictive Model Development:** develop a robust predictive model using Random Forest algorithms to forecast the outcomes of cats in a no-kill animal shelter.
- **Evaluation of Model Performance:** evaluate the performance of the developed predictive model, including measures such as accuracy, precision, recall, and F1-score
- **Enhanced Resource Allocation:** facilitate more efficient resource allocation within the animal shelter by leveraging insights provided by the predictive model.
- **Strategic Decision Support:** provide shelter management with valuable decision support tools based on the predictive model's outputs. These tools enable administrators to make informed decisions regarding intervention strategies, adoption campaigns, foster programs, and community outreach initiatives, ultimately leading to improved animal welfare outcomes and a more sustainable operation.

MATERIALS & METHODS

- **Dataset Acquisition:** Acquiring the dataset from Austin Animal Shelter, via Kaggle.com, consisting of over 28,000 records of cats. This dataset contains information on various features such as age, breed, sex, spay/neuter status, and outcome types.
- **Data Preprocessing:** Dataset underwent preprocessing steps in Jupyter Notebook using Pandas library.This included handling missing values, encoding categorical variables, standardizing numerical features, and splitting the data into training and testing sets using various tools from SciKit-Learn package in Python.
- **Feature Selection:** This involved identifying influential factors affecting outcomes, such as age group, weekday and time of adoption, spay/neuter status.
- **Model Development:** The Random Forest algorithm was implemented to develop the predictive model. The model was trained on the training dataset and developed using tools from Scikit-Learn.
- **Model Evaluation:** The performance of the developed model was evaluated using accuracy scores. Cross-validation techniques were employed. Confusion matrices were analyzed to understand the model's behavior across different outcome classes.
- **Results Interpretation:** Feature importance plots were generated to facilitate interpretation and aid in decision-making processes for shelter management. Plots were generated using Matplotlib in Python.
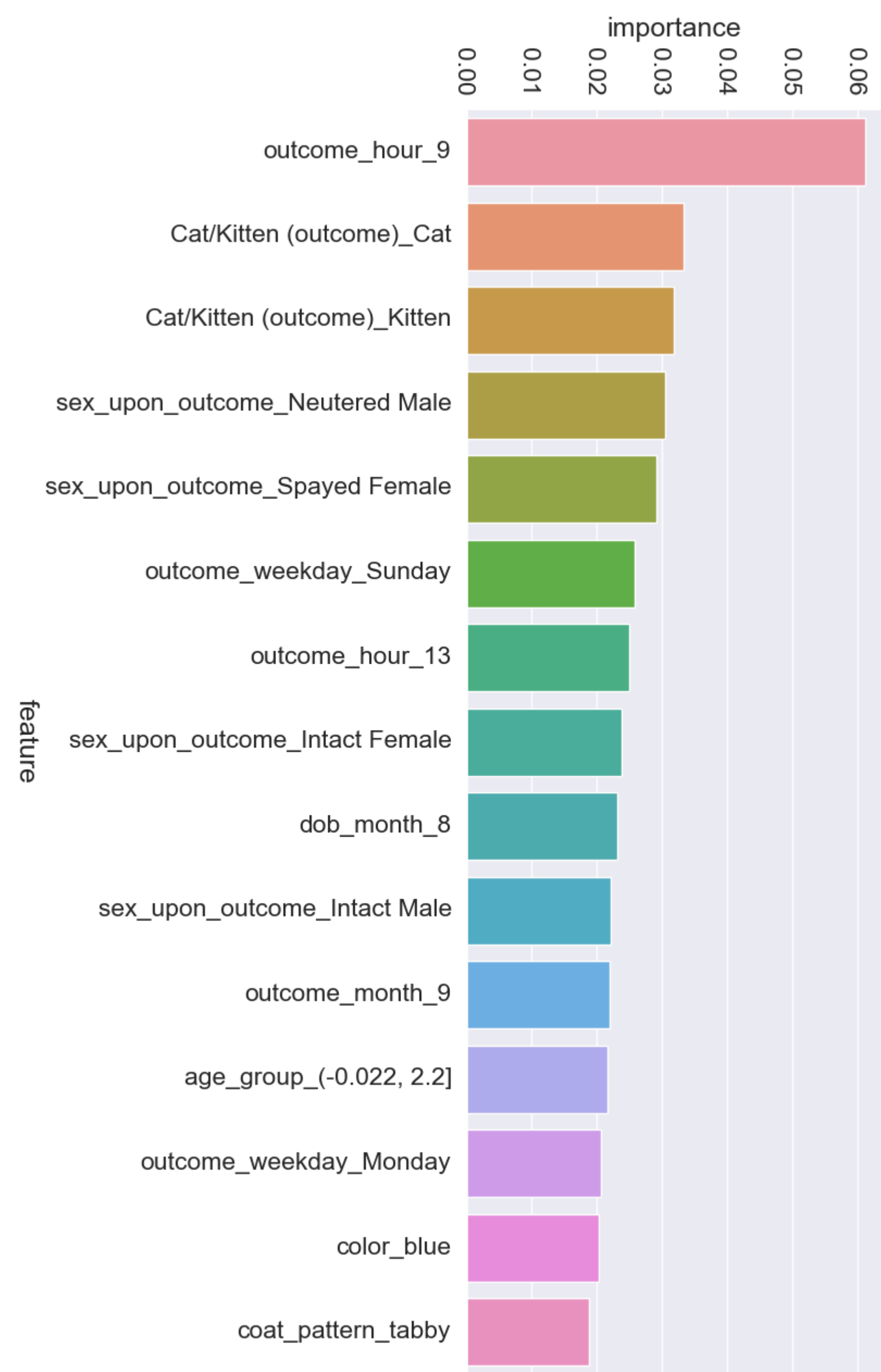
RESULTS

- **Accuracy:** The preliminary results of the random forest model show an accuracy rate of approximately 73%, **indicating moderate success** in predicting outcomes for cats in the Austin Animal Shelter.
- **Challenges:** Throughout the model development process, several challenges were encountered, particularly in feature selection and data preprocessing. The **dataset's diverse nature presented difficulties** in identifying the most influential features and handling missing or inconsistent data.
- **Lessons Learned:** Despite these challenges, the project provided valuable insights into the complexities of predictive analytics in animal shelter management. It underscored the importance of thorough data exploration, feature engineering, and model tuning for improving predictive performance. Moving forward, further refinement and optimization of the model are necessary to enhance its accuracy and applicability in real-world scenarios.
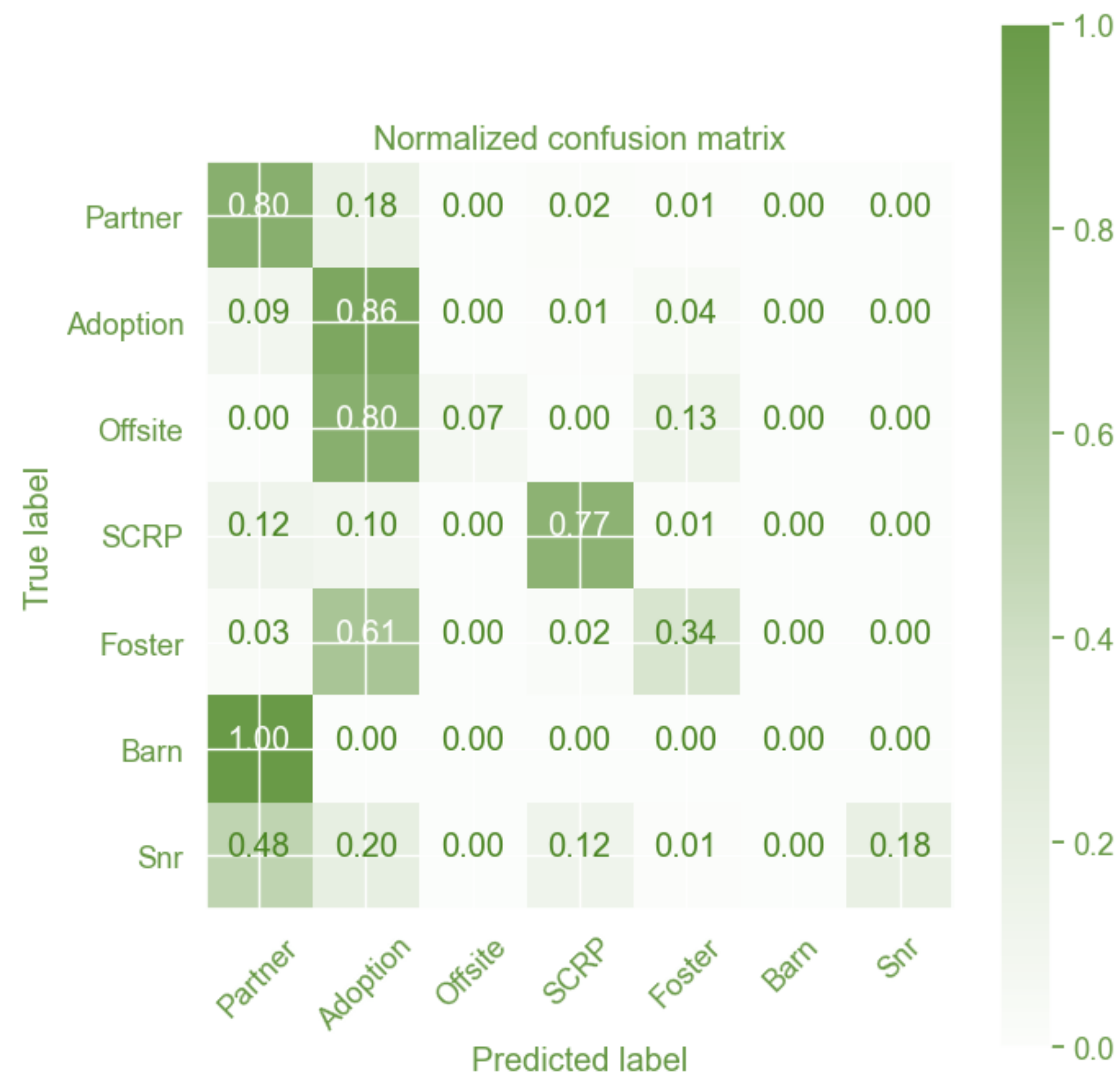
FUTURE WORK

Comparative Analysis: investigate alternative predictive models to assess their performance in predicting outcomes for cats in animal shelters.

Hyperparameter Tuning: Further refinement of the random forest model will be pursued through extensive hyperparameter tuning.

Evaluation Metrics: In addition to accuracy, the evaluation of model performance will be expanded to include a comprehensive set of metrics such as precision, recall, and F1-score.

Real-World Deployment: Ultimately, the goal is to deploy the refined predictive model in real-world animal shelter settings to assist shelter staff in making data-driven decisions regarding the care and placement of cats.

CONCLUSIONS

The preliminary results of our predictive analytics project demonstrate the potential of using machine learning algorithms to forecast outcomes for cats in animal shelters. Despite the model's current state of refinement, the random forest algorithm shows promising performance with an accuracy of nearly 73%. However, it is essential to acknowledge the limitations encountered during this initial phase, including the need for further model refinement and the exploration of alternative algorithms. Moving forward, continued efforts will focus on enhancing the model's predictive accuracy through hyperparameter tuning and comparative analysis with other machine learning techniques. By leveraging data-driven insights, we aim to empower animal shelters with the tools necessary to improve decision-making processes and ultimately enhance the welfare of shelter animals.

REFERENCES

Dataset: Austin Animal Center Shelter Outcomes by Aaron Schlegel, https://www.kaggle.com/datasets/aaronschlegel/austin-animal-center-shelter-outcomes-and?resource=download

Tools: Python, Jupyter Notebooks, Pandas, Matplotlib, Scikit-Learn, Tableau

CONTACT INFORMATION

Tess Anderson: tanderson4@bellarmine.edu
Dr. Robert Kelley: rkelley@bellarmine.edu