A hybrid deep learning-random forest model for predicting pet adoption

Anik Shikarpuri, Daniel Lee, Keteng Li, Peter Wu

Introduction:

With the growing interest in Artificial Intelligence, it is crucial to identify which algorithms are most effective for specific applications. This paper focuses on the pet adoption sector, exploring the utility of Machine Learning (ML) versus Deep Learning (DL) models for predicting how quickly a shelter animal will be adopted. Given the dual nature of data in this field—tabular data representing the animals' attributes and image data capturing their physical appearance—this study investigates whether these models should be used independently or in conjunction. We aim to determine the optimal strategy to enhance the predictability of adoption rate, which is vital for improving the outcomes for shelter animals and optimizing shelter resources. To answer this question we will use a Pet Adoption dataset. This dataset will be used to predict whether or not a pet will be adopted. In other words, we will be using traditional machine learning and deep learning to develop a binary classification model. We will then show how a simplistic approach can be surprisingly robust. The reason for choosing our model to access animal adoptions is for its ethical gratification. Every year millions of stray animals and pets suffer on the streets or are euthanized in shelters every day around the world. A lot of pets may be undernourished or require a new home due to personal reasons. With this being said, many families are left wondering if their pet would even have a chance of being adopted or if they would have to undergo the unfortunate event of euthanization. Our goal is to predict if a pet or stray animal would be adopted. This could help adoption shelters understand which animals their marketing should be focused towards. As well as helping families understand the best future actions regarding their own pet. So now we ask, what model would help save these pets?

Related Work:

To begin we will discuss previous approaches to this question. Janae Bradley and Suchithra Rajendran's research focused on predicting the duration of animals' stays in shelters to boost adoption rates and reduce euthanasia. Their paper introduced a two-phase approach to animal relocation, balancing adoption speed and relocation costs using various machine learning models, including logistic regression, random forest, and gradient boosting. The results indicated that the gradient boosting algorithm surpassed other models, with age, size, and color being key factors affecting stay duration. Conversely, a study by Pin Wang and colleagues compared the Support Vector Machine (SVM) and Convolutional Neural Network (CNN), finding that performance varied with dataset size. Traditional machine learning achieved higher accuracy with smaller datasets, while deep learning excelled with larger ones. Although the first study employed multiple traditional algorithms to predict shelter stay durations, it highlighted the limitations of simpler models. The second study focused on dataset size impacts rather than exploring different machine learning and deep learning techniques. Recognizing these limitations, our project will investigate diverse methods to address these challenges.

Methods:

It is crucial that we first discuss the structure of the dataset before we can move to the intricate development and progress we made in Artificial Intelligence. Our dataset consisted of both tabular and image data. Within the tabular data, we were given a variety of features such as age, breed, color, gender, sterilization, and much more. In total we were given 24 covariate and 15,000 pets in total. During the NLP process, Textblob was used to create two new covariates which are the description score of each data value along with the word counts. In regards to the image data, we were given multiple images that all correspond to a pet ID which allowed for a connection between the images and the tabular data.

Deep Learning:

In the deep learning process, we conducted two preprocessing steps before moving on to model evaluation and result generation. First, we split the tabular training data into 80% for training and 20% for testing. Both the tabular data and the image data contain the same amount of pet IDs, which allows us to link these two datasets. Additionally, we balanced the data's adoption speed, then converted the adoption speed into binary data as shown in Figure 1 and Figure 2. Instead of using the original multi-class format, we transformed the data into a binary classification where classes 0-3 were combined into one class called 'Adopted' and class 4 was labeled 'Not Adopted'. To ensure each class within 'Adopted' was balanced, we performed a data balancing step to prevent significant influence from any single class. After balancing the tabular data and before splitting the image data, we proceeded to allocate the images. For instance, if a pet ID was found in the 80% training set, we moved all corresponding images into the training image set. The same process was applied for the test set. This resulted in separate training and testing datasets for the images.

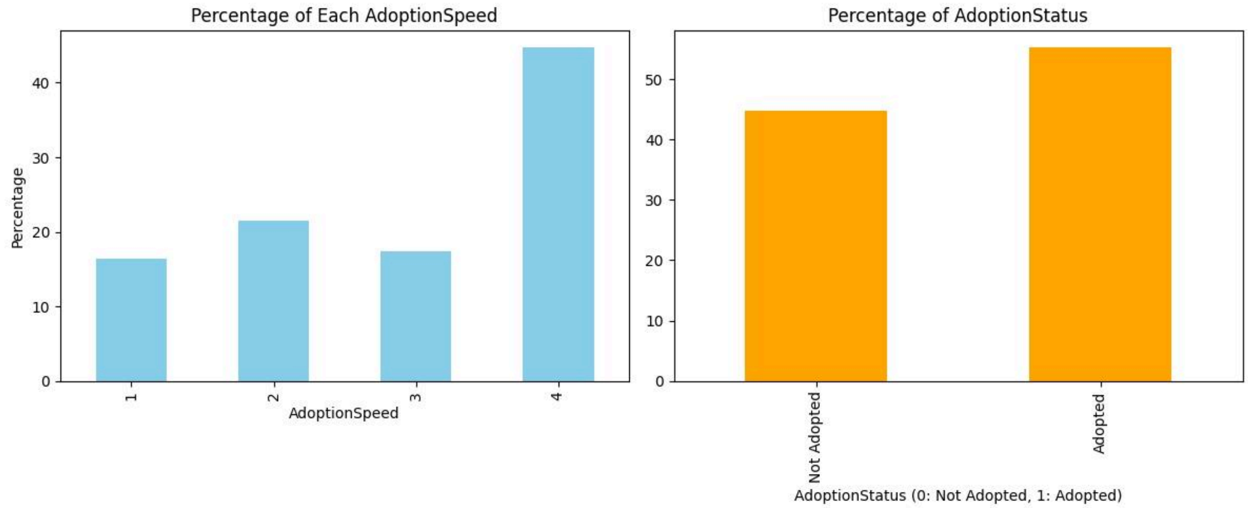

Figure 1. Percentage of Multiclass to Binary Classification labels for train set

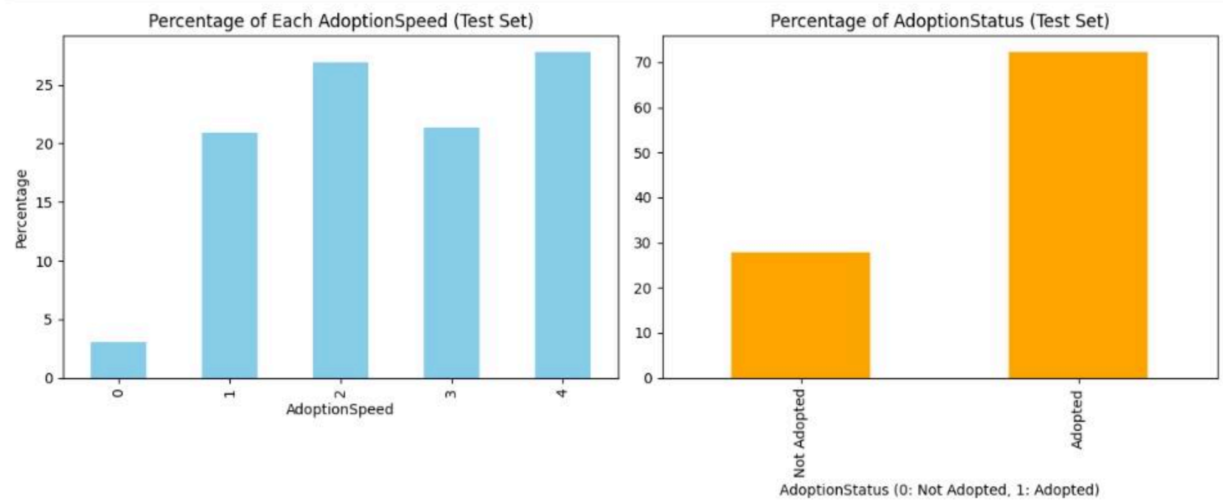

Figure 2. Percentages of Multiclass to Binary Classification Labels for test set

In the image preprocessing phase of our deep learning pipeline, a critical step involved the standardization of input image sizes. Initially, we performed an image exploration to understand the variety of sizes and shapes of the images in our dataset. The average image size, as illustrated, was 399.75*422.25. However, rather than resizing all images to this average size, we chose to resize them to the most common size found in the dataset, which was 400*300. This decision was made to maintain consistency across the dataset, as using the most common size helps avoid distortion and potential loss of information that could occur if the average size was used as the standard. The image resizing step ensures that our model receives uniformly-sized inputs, which is crucial for the consistency of feature extraction.

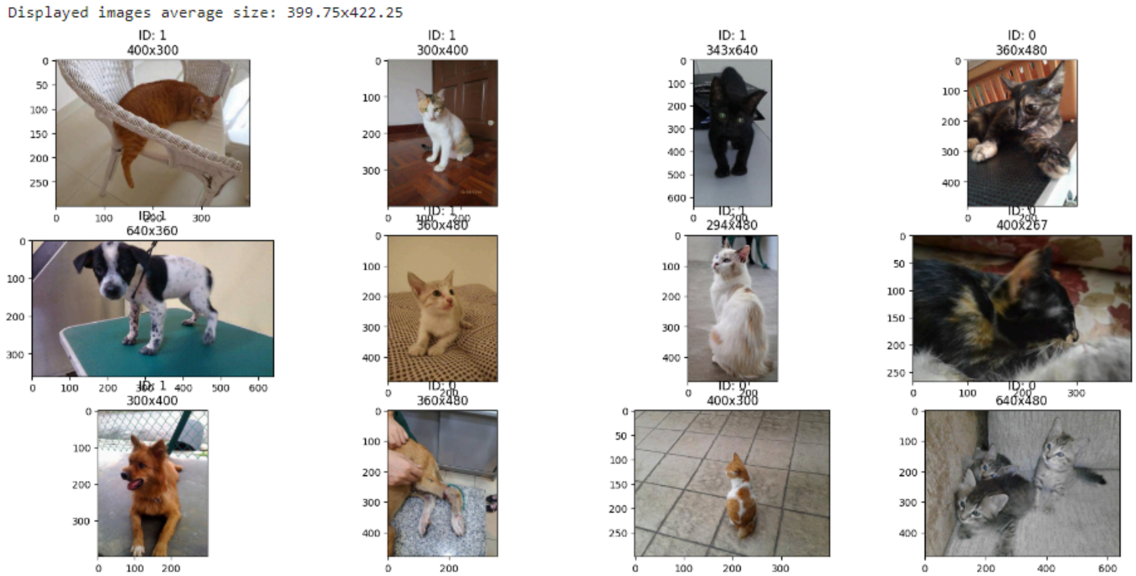

Figure 3. Image Size Pre-Processing

Instead of building a deep learning model from scratch, we employed a technique called transfer learning. Our model, built on the robust foundation of DenseNet121, showcased its proficiency in image classification by achieving 87.5% accuracy during training and 83.33% accuracy on the test image set. By using pretrained weights from ImageNet, the model started with a solid base for feature detection that was then fine-tuned with our specific dataset of pet images. Thanks to the implementation of Global Average Pooling and a fully connected dense layer, we avoided overfitting. The fine-tuning process, together with the strategic use of a sigmoid output layer for binary classification and model checkpoints to save the best iterations, resulted in a well-adjusted model. This model is capable of robustly and consistently distinguishing between 'Adopted' and 'Not Adopted' pets. After evaluating the model, we proceeded to run a prediction with the model on the entire set of training images. The output contained two features: PetID and the prediction results from the deep learning models.

	Train	Test
Keras Densenet 121	87.5%	83.33%

Figure 4. Keras Densenet 121 Evaluation

Traditional Machine Learning:

	Train	Test
KNN	86.31%	66.66%
Logistic Regression	67.91%	66.63%
Random Forest	95.75%	80.93%

Figure 5. ML training and testing results

The following section will shift focus from Deep Learning to Machine Learning. We started with creating multiple simple machine learning models. These models were a K Neighbor Classifier (KNN), Logistic Regression, and Random Forest. The KNN model had a test accuracy of 66.66% while the Logistic Regression was 66.63%. Showing us that the difference between the two were marginal. This then led to the development of the random forest model with an accuracy of 80.93% (refer to *Figure 5*). This was much more powerful and was the reason we focused our attention to a random forest approach.

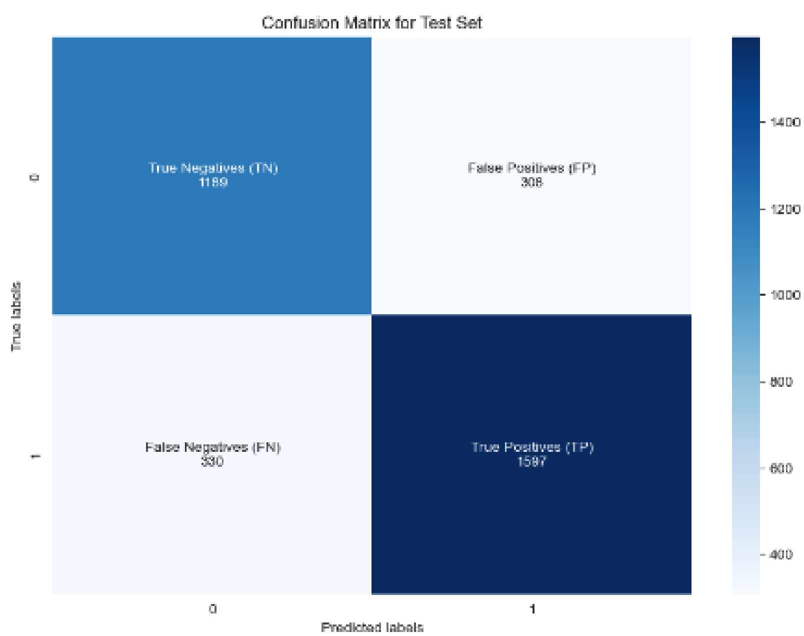

Figure 6. Hyperparameter Random Forest Results

When comparing these results to the deep learning models, the results are very similar, not to mention the much smaller thirst for computational power when deploying machine learning compared to deep learning. However, once we came to this result, it made us question if the most accurate results required for a hybrid and stacked approach. In other words, should we combine the deep learning and machine learning models together? The reason for this approach is we hope to find that where machine learning falls short, deep learning would prove to be accurate and vice versa. As we can see in *Figure 6*, the traditional Machine learning model was accurate but lacked robustness on true negative predictions relative to true positives. We believe this is due to an imbalance in the dataset. With more values of 1 being provided than 0. To alleviate or reduce the imbalance as well as creating a more robust model we created a Hybrid Model.

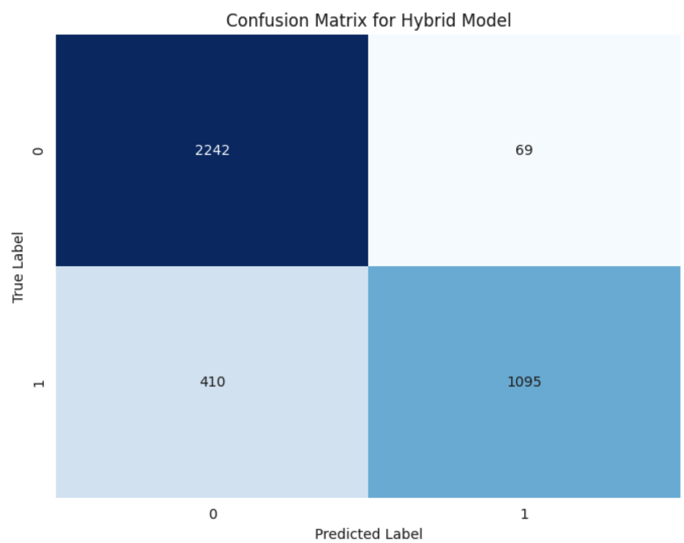

Figure 7. Confusion Matrix for Hybrid Model

The hybrid model incorporates the identical dataset utilized in the traditional machine learning approach, along with the inclusion of deep learning predictions as a covariate. Using deep learning predictions as a covariate, random forest with hyperparameter tuning was utilized which resulted in approximately 92% accuracy. Compared to the accuracy of the hyperparameter tuning model without the deep learning predictions as a covariate, the accuracy of the hybrid model increased by 12 percent. As we can see in *Figure 7*, the hybrid model was accurate and showed a reasonable amount of true positives and true negatives. However, there are some false negatives that could be improved.

Now that we have discussed results provided from Deep Learning, Traditional Machine Learning, and now the Hybrid model; let us now delve deeper into the results. We can say with pretty high certainty that we can predict whether or not a pet would be adopted with the hybrid model having over 90% accuracy. This is a testament to its robustness and the synergistic effect of combining machine learning and deep learning techniques. This high level of accuracy is particularly noteworthy because it significantly improves the predictive capabilities over the individual models. By harnessing the strengths of both model types, the hybrid approach not only enhanced the overall performance but also simplified the analytical process, providing clear insights with less complexity. We use both image and tabular data which means we are able to use the full scope of data effectively and efficiently.

This simple yet powerful hybrid model approach underscores a critical lesson in predictive modeling: complexity does not always equate to efficacy. The ability to derive profound insights from a straightforward model configuration demonstrates that an intelligently designed hybrid system can meet the analytical needs efficiently. For stakeholders in animal welfare, this means more informed decision-making and better allocation of resources, ultimately leading to increased adoption rates and improved animal welfare. This model serves as a compelling example of how innovative, simplified hybrid modeling can address complex problems effectively.

The original question we were trying to answer is whether or not we can predict if an animal is adopted. This is our main objective but if we step back what are we really trying to achieve? We want to be able to help adoption centers and pet owners with their decision to adopt or even put up their own animal up for adoption. We want to find insight as to what makes it easier to get a pet adopted and what limitations should you consider when you are putting your beloved pet up for adoption. Through this extensive research and development we were able to produce the key attributes that help with animal adoption.

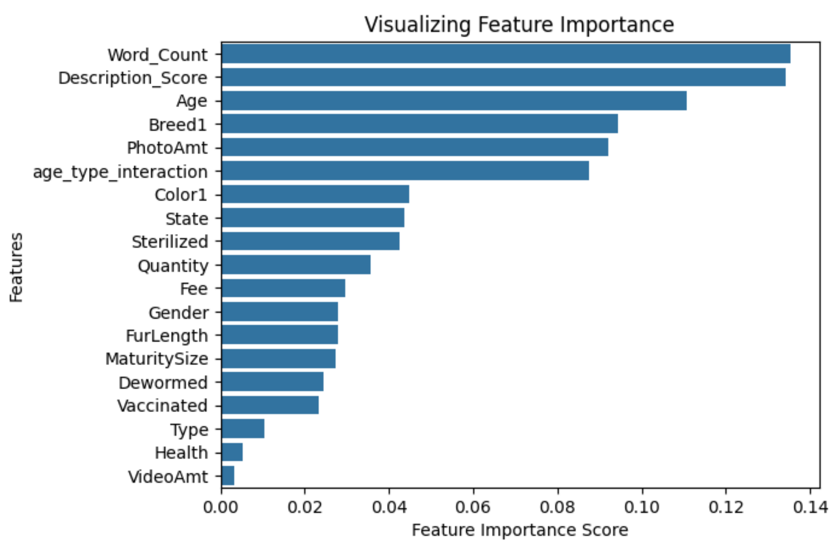

Figure 8. Feature Importance in Hybrid Model

As we can see from *figure 8*, the top 6 features are very influential to determining if a pet is adopted or not. We see the most important features are part of the NLP model, meaning descriptive characteristics are extremely helpful with pet adoption. We also have the insight to conclude that ‘Age’ and ‘Breed’ play a crucial role in animal adoption. These insights are how we know this research accomplished its goals as this can be very intriguing information to shelters. Most importantly, the pipeline and systematic approach has shown us that simplicity is key to finding solutions. Understanding your data and finding the best way to optimize results with the given information is the largest factor in creating astounding results.

Future Work:

Given more time, a deeper exploration into the dynamics between the random forest and the convolutional neural network (CNN) components of our hybrid model would be immensely valuable. Specifically, understanding in which situations the random forest model learns to trust the predictions made by the deep learning model could enhance our approach significantly. This investigation would involve analyzing the decision boundaries and feature importance metrics from the random forest in conjunction with the feature maps generated by the CNN. By identifying the scenarios where the random forest either defers to or diverges from the CNN's insights, we could refine the model's accuracy and reliability. Additionally, this nuanced understanding would allow us to tailor the hybrid model more effectively to different datasets or potentially similar applications in other domains, ensuring optimal performance and broader applicability. This kind of analytical depth could pave the way for more sophisticated, context-aware predictive models in the future. Another limitation we had was computational power. All of our work was done on our personal computers. We believe if we were given stronger computational capacity, such as access to a super computer, results could have been drastically different. And as stated previously if given more time in parallel to stronger computational capacity we could have achieved results that are much higher and faster than the ones we produced. But then that could lead to the question of is all that energy and power really necessary to increase an already substantially high accuracy. We believe when it comes to the intersection of efficiency and accuracy our model has hit the nail on the head.

Conclusion:

Applying machine learning and deep learning to our hybrid model has not only helped us make better predictions about adoptions, but has also shed light on a path toward stronger, clearer solutions for future predictive models. The evolution of our models from traditional to a hybrid approach advocates a key factor that increasing complexity often does not, in fact, lead to improved results. Instead, as our work suggests, employing a hybrid approach that leverages the strengths of traditional and advanced techniques can significantly improve performance. The model continues to offer insights that can help animal shelters and pet owners make critical decisions that could increase adoption rates and improve the lives of animals.

Work cited

Bradley, Janae, and Suchithra Rajendran. “Increasing Adoption Rates at Animal Shelters: A Two-Phase

Approach to Predict Length of Stay and Optimal Shelter Allocation - BMC Veterinary Research." *BioMed Central*, BioMed Central, 5 Feb. 2021, bmcvetres.biomedcentral.com/articles/10.1186/s12917-020-02728-2.

https://github.com/qyefearth/PTW_CNN_Pet-Adoption

https://github.com/DanielLee0226/Pet-Adoption/blob/main/Traditional_ML.ipynb

https://github.com/DanielLee0226/Pet-Adoption/blob/main/Traditional_ML_DeepLearningPredictions.ipynb