

Primera entrega de proyecto

Javier Mauricio Suárez Monsalve

Curso: Introducción a la inteligencia artificial

Profesor: Raul Ramos Pollan

1. Descripción del problema

Millones de animales callejeros sufren en todo el mundo y encontrarles un hogar puede salvar vidas y crear familias felices. PetFinder.my es la principal plataforma de bienestar animal de Malasia y colabora con amantes de los animales, medios de comunicación, corporaciones y organizaciones globales para mejorar el bienestar animal. Las tasas de adopción de animales están fuertemente correlacionadas con los metadatos asociados con sus perfiles en línea, como el texto descriptivo y las características de las fotos. Por lo anterior, el problema predictivo a resolver utilizando el dataset de Petfinder Adoption Prediction es predecir la velocidad de adopción de un animal en función de su perfil, incluyendo información como edad, raza, género, salud, entre otros.

2. Dataset a utilizar

El dataset utilizado es Petfinder.my - Adoption Prediction, disponible en la sección de competencias de la plataforma Kaggle, en el siguiente enlace: <https://www.kaggle.com/competitions/petfinder-adoption-prediction/> . Contiene información detallada de más de 14,000 animales en adopción en Malasia, recopilada a través del sitio web de Petfinder.my. Incluye características del animal, como su edad, raza, género, tamaño, nivel de salud, así como información sobre el entorno del animal, como su ubicación geográfica y la cantidad de fotos y descripciones en su perfil de adopción. El dataset tiene 24 columnas, pero usando algunas variables continuas como la edad y la tarifa por adopción se pueden discretizar, obteniendo así las 30 columnas solicitadas.

3. Métricas de desempeño requeridas

La competición sugiere utilizar el valor de kapa ponderado. En este caso se va a utilizar como métrica de machine learning el RMSLE (Raíz del error logarítmico cuadrático medio) la cual se puede aplicar en la competición Petfinder.my Adoption Prediction para evaluar el desempeño de los modelos de predicción de adopción de mascotas en función de la velocidad de adopción.

Por ejemplo, si un modelo predice una velocidad de adopción de 2 para una mascota, pero la velocidad real es de 10, la RMSLE sería mayor que si el modelo predice una velocidad de adopción de 8 para una mascota cuya velocidad real es 10. Un modelo que tenga un RMSLE bajo se consideraría más preciso en la predicción de la velocidad de adopción y, por lo tanto, se consideraría un mejor modelo en términos de desempeño en la competición.

4. Desempeño

Un modelo que pueda predecir con precisión la velocidad de adopción podría ser utilizado por los refugios y organizaciones de rescate para mejorar sus prácticas de adopción y reducir el tiempo que los animales pasan en refugios. Es decir, una vez identificado con el modelo aquellas mascotas que toman mas tiempo en ser adoptados, esta información será utilizada para guiar a los refugios y rescatistas de todo el mundo para mejorar el atractivo de los perfiles de sus mascotas, reducir el sufrimiento de los animales y la eutanasia. Si la velocidad de adopción de los animales que mas demoran en los albergues, es mejorada, quiere decir que el modelo en producción es bueno.