

Medida de Pawpularity con Deep Learning

Juan Camillo Avendaño Rodríguez
camilo.avendano1@udea.edu.co
Ingeniería de sistemas

Universidad de Antioquia
Medellín, Colombia

Contexto- Cada año, millones de animales son abandonados o sacrificados en refugios de todo el mundo. Las plataformas de adopción de mascotas juegan un papel crucial para conectar a estos animales con posibles adoptantes, incrementando así sus probabilidades de encontrar un hogar permanente. Sin embargo, la competencia es alta: los refugios y rescatistas tienen que mostrar decenas, a veces cientos, de animales en sus listados, por lo que captar la atención de los usuarios se vuelve un reto.

Las fotografías juegan un papel determinante en este proceso. Estudios y observaciones indican que las mascotas con fotos atractivas tienen más probabilidades de ser adoptadas que aquellas con imágenes de baja calidad[1]. En plataformas como PetFinder.my, las fotos son el principal medio para captar el interés de los adoptantes potenciales. Sin embargo, lo que define una "buena" foto puede variar y no siempre es fácil de cuantificar.

PetFinder.my, la principal plataforma de adopción de animales en Malasia, ha implementado un Cuteness Meter, que analiza fotos de mascotas basándose en ciertos parámetros como la composición de la imagen. Aunque útil, esta herramienta está en una fase experimental y depende de un algoritmo sencillo que tiene mucho margen para mejoras. Aquí es donde el Deep Learning puede aportar un valor significativo, permitiendo desarrollar modelos que analicen de manera más precisa la calidad de las fotos y su "Pawpularity", optimizando la visibilidad de los animales en adopción.

El proyecto que se plantea tiene un impacto real en el bienestar animal, ya que un modelo eficaz puede aumentar la tasa de adopción y, como consecuencia, mejorar la vida de miles de animales. Este contexto también es relevante para demostrar habilidades avanzadas en Ciencia de Datos y Deep Learning, ya que combina el análisis de imágenes y datos tabulares para abordar un problema de clasificación y regresión en un entorno de la vida real.

medido a través de métricas como las visitas a la página y las interacciones con el perfil del animal.

En términos técnicos, este proyecto tiene dos componentes principales:

1. **Predicción basada en imágenes:** Utilizando un modelo de redes neuronales convolucionales (CNN), entrenaremos el sistema para analizar la composición y calidad visual de las fotos. Las redes CNN son ideales para este tipo de problemas porque pueden captar patrones visuales en las imágenes (como enfoque, iluminación, etc.) que influyen en la percepción del atractivo, así extraer estas características para llegar a un modelo objetivo de regresión basada en metadatos.
2. **Predicción basada en metadatos:** el modelo también considerará un conjunto de características tabulares (metadatos) que incluyen información clave sobre la calidad visual, presencia de elementos adicionales (como humanos o accesorios), y otras características, el cual apunta a sacar la medida de pawpularity. Este enfoque híbrido permitirá al modelo aprender tanto de la imagen como de los descriptores manuales.

El objetivo es no solo predecir con precisión el puntaje de Pawpularity, sino también interpretar el impacto de ciertos atributos en la calidad de las fotos. A nivel de negocio, un modelo exitoso permitirá hacer recomendaciones automáticas sobre cómo mejorar la calidad de las fotos, por ejemplo, sugiriendo ajustes en el ángulo, iluminación o presencia de accesorios.

Además, este reto es parte de un curso donde se espera demostrar conocimientos en Deep Learning. El objetivo es no solo entrenar un modelo predictivo, sino también aplicar técnicas de preprocesamiento de imágenes, análisis de datos y evaluación de modelos, utilizando métricas relevantes como el **Root Mean Squared Error (RMSE)**[2] para medir el rendimiento del modelo.

I. Objetivo

El objetivo principal de este proyecto es desarrollar un modelo de **Machine Learning** que sea capaz de predecir la "Pawpularity" de una imagen de mascota. Este puntaje refleja el atractivo de la foto y está correlacionado con el interés que genera entre los usuarios,

II. Dataset

- **Tipo de datos:**
 - Imágenes de mascotas en formato .jpg.

- Metadatos en formato tabular, que incluyen características como el enfoque, la presencia de humanos, accesorios, y más.
- **Tamaño:**
 - Aproximadamente **9,912** imágenes de entrenamiento.
 - El archivo **train.csv** contiene metadatos sobre cada foto y su respectivo puntaje de "Pawpularity".
 - Tamaño en disco: Aproximadamente **1 GB** (para imágenes y metadatos combinados).
- **Distribución de las clases:**
 - El objetivo es predecir un puntaje de "Pawpularity", que es una variable continua (no categórica). La distribución del puntaje varía entre 0 y 100 la cual tiene la siguiente distribución

Distribuciones de los datos

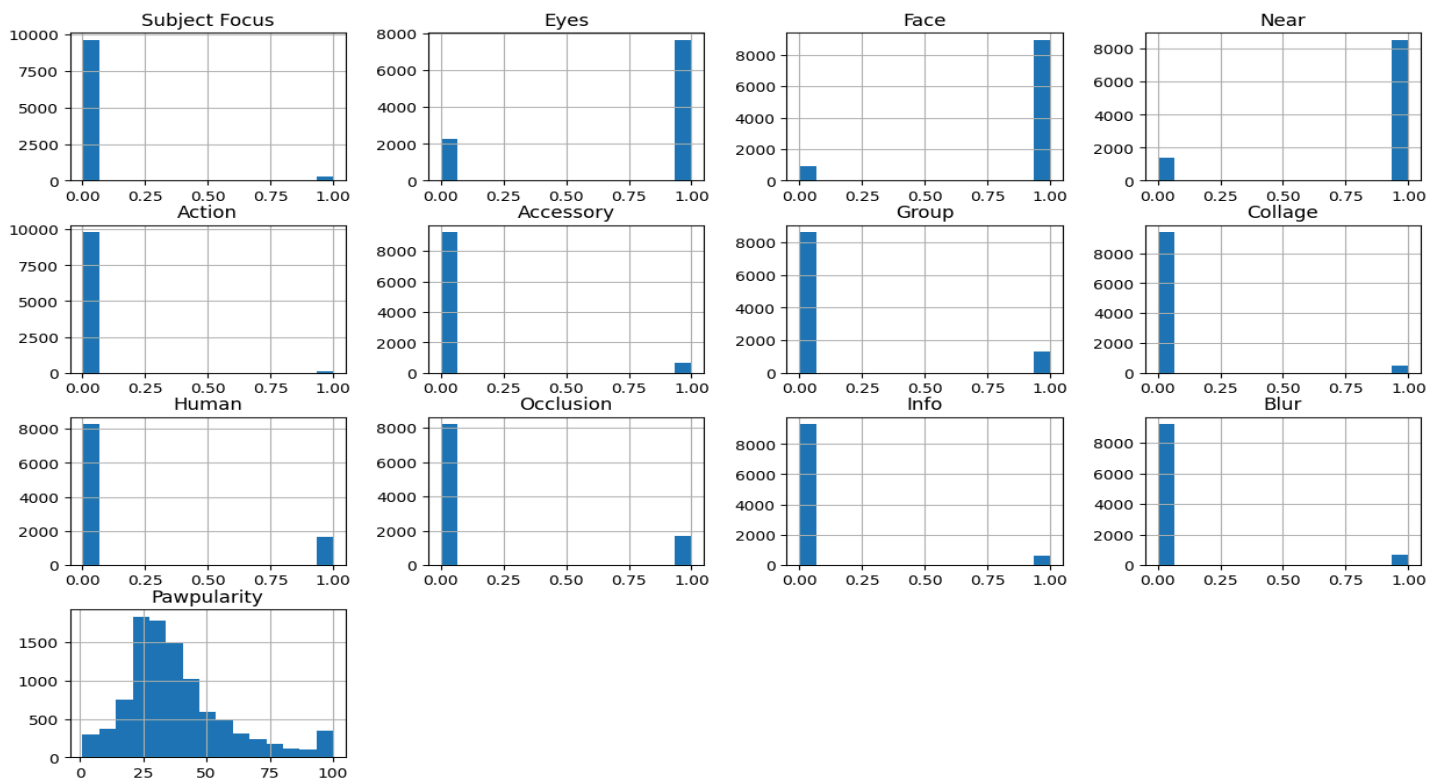


Figura 1

Distribución de los metadatos de las imágenes

La métrica principal de evaluación será el **Root Mean Squared Error (RMSE)**, ya que estamos tratando con un problema de regresión. RMSE mide la diferencia entre los valores reales y los valores predichos por el modelo, y minimizaremos esta métrica para optimizar el desempeño del modelo.

III. Métricas de desempeño

- **Métricas de Negocio:**

A largo plazo, el éxito del modelo se medirá por su capacidad de mejorar la tasa de adopción de mascotas, reflejada en la reducción del tiempo promedio que una mascota permanece en adopción. Un modelo efectivo ayudaría a optimizar los perfiles de las mascotas, haciéndolas más atractivas para los posibles adoptantes.

IV. RESULTADOS PREVIOS

La plataforma PetFinder.my actualmente utiliza un Cuteness Meter básico que analiza la composición de las fotos de mascotas, pero este sistema es experimental y tiene espacio para mejoras sustanciales.

Actualmente, estoy definiendo la metodología para construir un **modelo intermedio** que transforme las imágenes en sus correspondientes características de metadatos. Este modelo actuará como un puente, extrayendo información clave de las imágenes (como el Subject Focus, Eyes, Face, Near, Action, Accessory, Group, Collage, Human, Occlusion, Info, Blur, Pawpularity) que luego se utilizará como entrada para el modelo final de predicción de **Pawpularity**. De esta manera, se busca automatizar el proceso de generación de metadatos a partir de las imágenes, facilitando la integración de estas características en el modelo de regresión.

V. REFERENCIAS

[1] "Efecto de superioridad de la imagen: ¿qué es y cómo nos afecta?"

<https://psicologiaymente.com/psicologia/efecto-superioridad-imagen> (accessed Sept 7, 2024).

[2] RMSE, https://docs.oracle.com/cloud/help/es/pbcs_common/PFUSU/insights_metrics_RMSE.htm#PFUSU-GUID-FD9381A1-81E1-4F6D-8EC4-82A6CE2A6E74 (accessed Sept 7, 2024)

[3] Challenge Kaggle, "PetFinder.my - Pawpularity Contest", <https://www.kaggle.com/competitions/petfinder-pawpularity-score/overview> (accessed Aug 24, 2024)