**Motivation**

- Inspired by a recent kaggle challenge to predict pet adoption speed from pet profiles. The long-term goal of the project is to guide shelters and rescue on improving their pet profiles’ appeal, reducing animal suffering and euthanization.
- The input is the pet profile, including structured data columns and image data. The output is the adoption speed categorized from 0 to 4 (0 being the fastest and 4 being the slowest).
- We use pre-trained ResNet to extract image features, concatenate, with structured features at certain stage, and feed the combined vectors to another network.

**Dataset**

- 14993 pet profiles with labels from 0 to 4 indicating the adoption speed (the smaller, the faster the pet is adopted).
- Structured data: 24 columns, including numerical feature, nominal feature, and ordinal feature (11 out of 24 are shown on the right).
- Image data: 58311 pet images
  - One pet profile can have more than one photo, and one photo can include multiple pets.

**Feature Engineering**

- Structured data:
  - Keep numerical features as what they are (age, Adoption Fee, Number of photos, etc)
  - convert nominal data and ordinal data into One-Hot Vectors
  - In this way, every sample has a feature vector size of 5970 derived from structured data
- Image data:
  - Pick and scale the first image for every sample profile, for those that don’t have a image, we will insert a dummy black image. Preprocessed into uniform 224*224*3 array.
  - Transfer learning: use ResNet50 to process the images
- Structured data + Image data:
  - In a simple saying, at certain stage of the network, we will merge two set of feature data together into one.

**Models, Results, Discussion**

<table>
<thead>
<tr>
<th>DATASET</th>
<th>NETWORK</th>
<th>LOSS FUNCTION</th>
<th>ACC (O.C.)</th>
<th>O.P.RATE (O.C.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>Data</td>
<td>3-layers NN</td>
<td>Crossentropy</td>
<td>26.40%</td>
</tr>
<tr>
<td>Structure</td>
<td>Data</td>
<td>Pre-trained ResNet + 3-layers NN</td>
<td>Crossentropy</td>
<td>27.50%</td>
</tr>
<tr>
<td>Structure</td>
<td>Data</td>
<td>6-layers NN</td>
<td>Crossentropy</td>
<td>31.10%</td>
</tr>
<tr>
<td>Structure</td>
<td>Data</td>
<td>SAG: Pre-trained ResNet for image feature extraction, then feed into 3 layers NN</td>
<td>Crossentropy</td>
<td>32.13%</td>
</tr>
<tr>
<td>Structure</td>
<td>Data</td>
<td>SAG: Concatenate the results from 3 and 2, then run through softmax layer</td>
<td>Crossentropy</td>
<td>40.40%</td>
</tr>
</tbody>
</table>

**Future Work and Citations**

- Clean up description text data, translate non-English sentences into English, and extract important information like “reason for pet relinquishment” – “Break the 0.5”
- Enrich the infrastructure and improve the running efficiency, to allow more users seamless feeding in their own data to see the results

**Error Analysis**

- Rescuer Effect: Star rescuer vs single pet rescuer.
- Quantity effect: each profile can include multiple pets.
- Information in description not able to use yet - personality, reason for relinquishment.
- Profile quality: Low quality - Low prediction.

**Compare with other teams**

- Current Kaggile Public Leaderboard top 0.4920 vs. Our highest of 0.4801

---

1. Liwei Wang, Yin Li, Jing Huang, Svetlana Lazebnik (2018). “Learning Two-Branch Neural Networks for Image-Text Matching Tasks”
2. Classifying e-commerce products based on images and text.
3. Adopter-dog interactions at the shelter: Behavioral and contextual predictors of adoption

Stats table for full at the full report.
Link to Video:
https://youtu.be/QP6BnCuSXSk