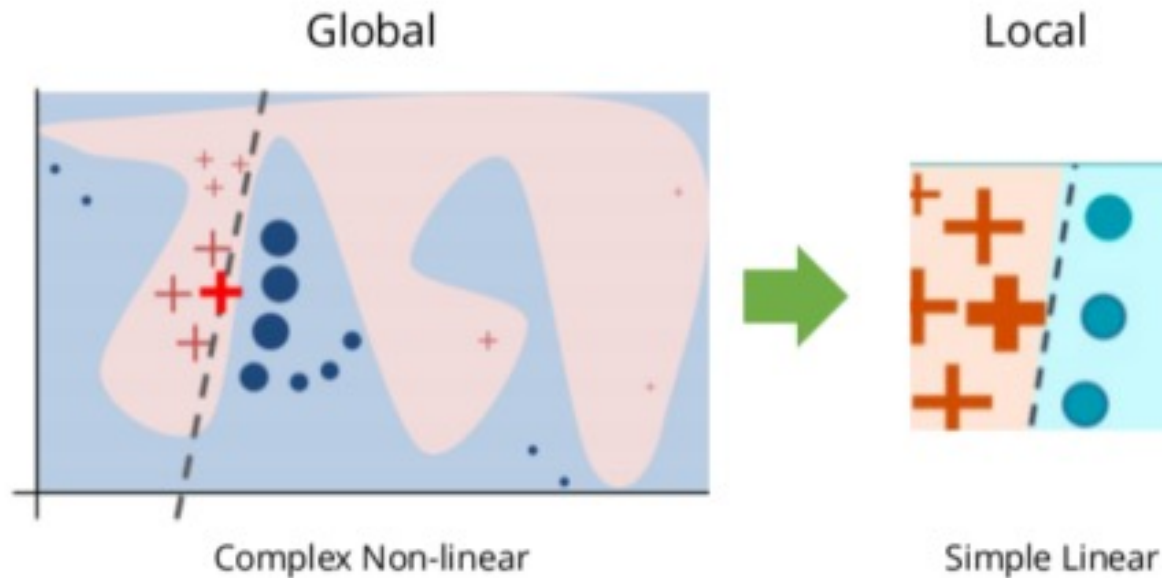


# Explainability

Eric Wong

9/29/2022

# Local Linearity



Marco Tulio Ribeiro "Local Interpretable Model-Agnostic Explanations (LIME): An Introduction"

# Superpixels for "interpretable" features

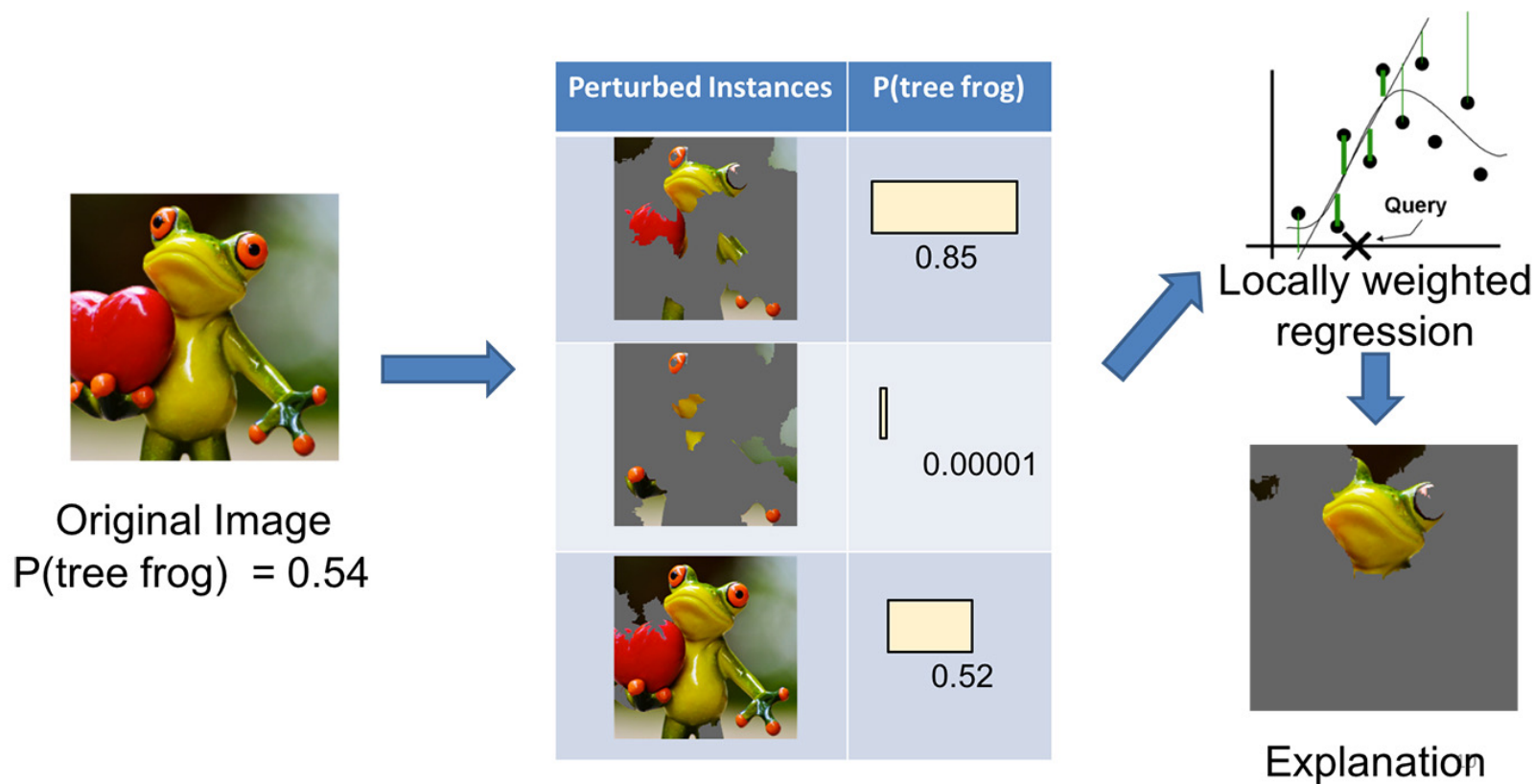


Original Image



Interpretable  
Components

# Perturb superpixels



# Explaining images



Marco Tulio Ribeiro "Local Interpretable Model-Agnostic Explanations (LIME): An Introduction"

# Explaining words

Prediction probabilities



atheism

christian



## Text with highlighted words

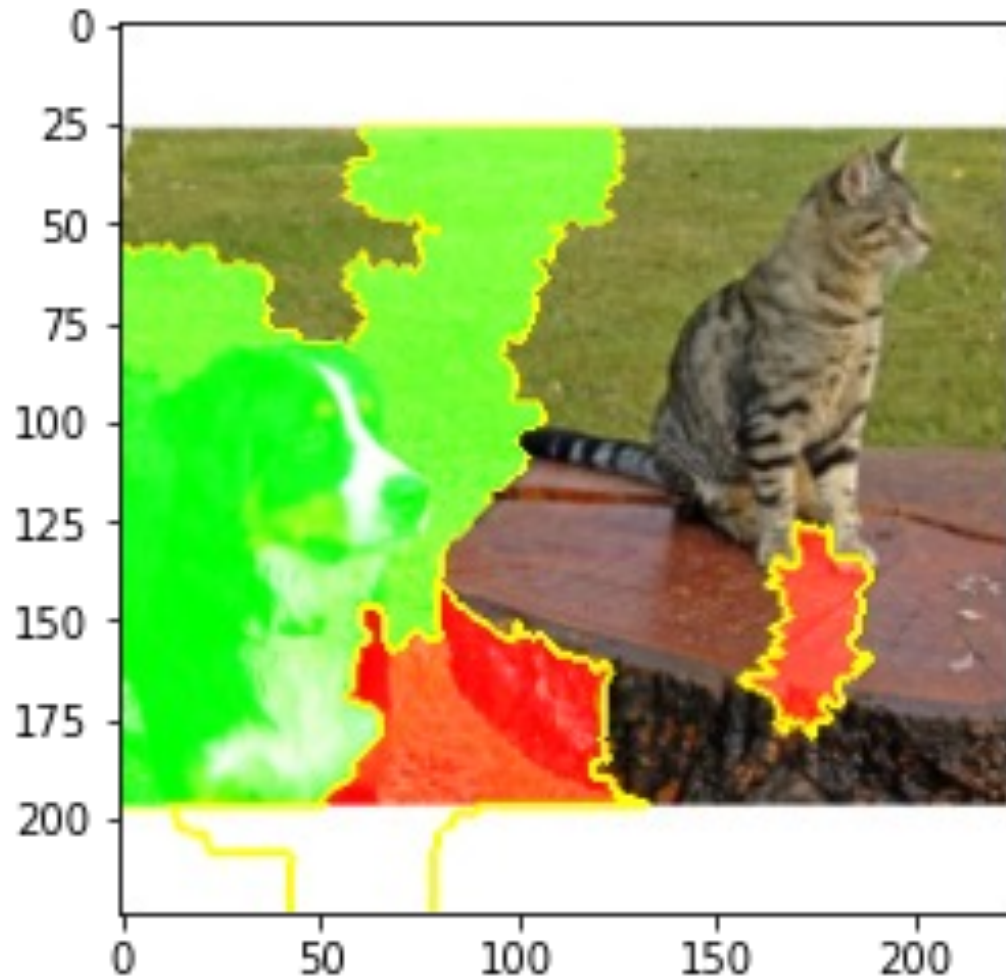
From: johnchad@triton.unm.edu (jchadwic)  
Subject: Another request for Darwin Fish  
Organization: University of New Mexico, Albuquerque  
Lines: 11  
NNTP-Posting-Host: triton.unm.edu

Hello Gang,

There have been some notes recently asking where to obtain the DARWIN fish.

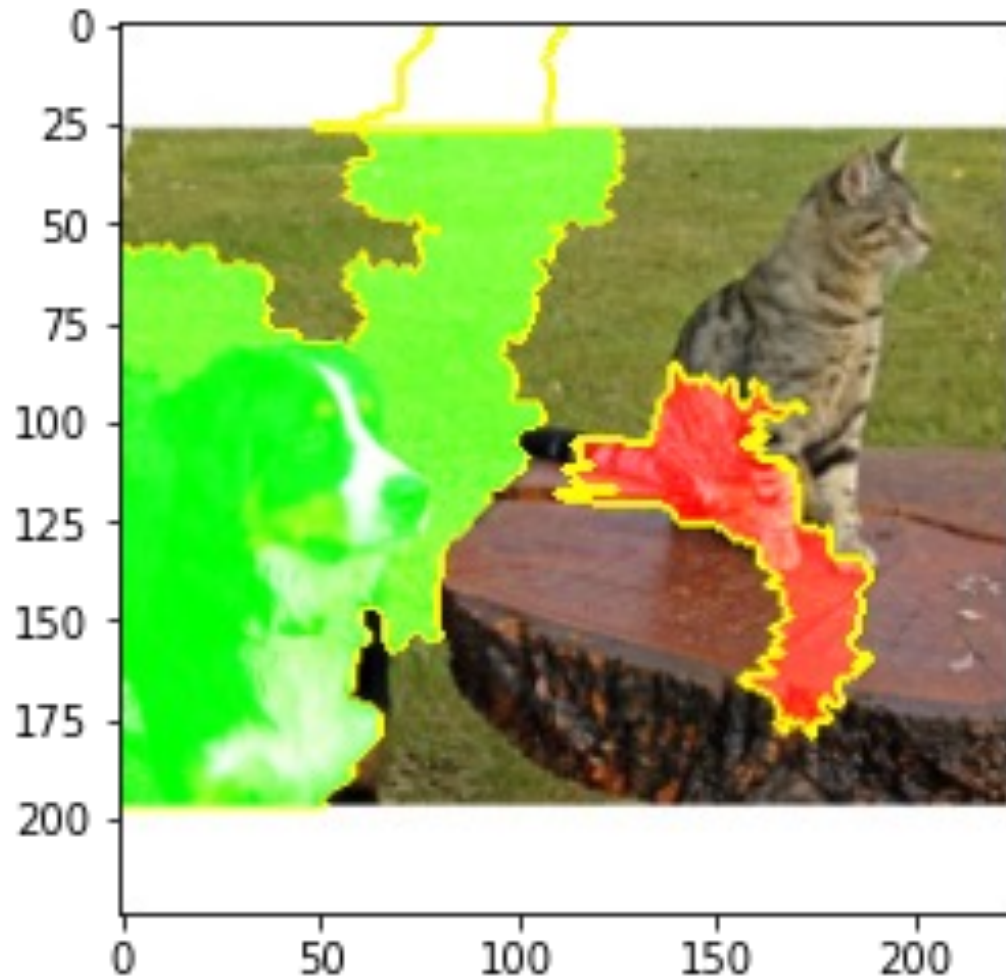
This is the same question I have and I have not seen an answer on the net. If anyone has a contact please post on the net or email me.

# A closer look at dog



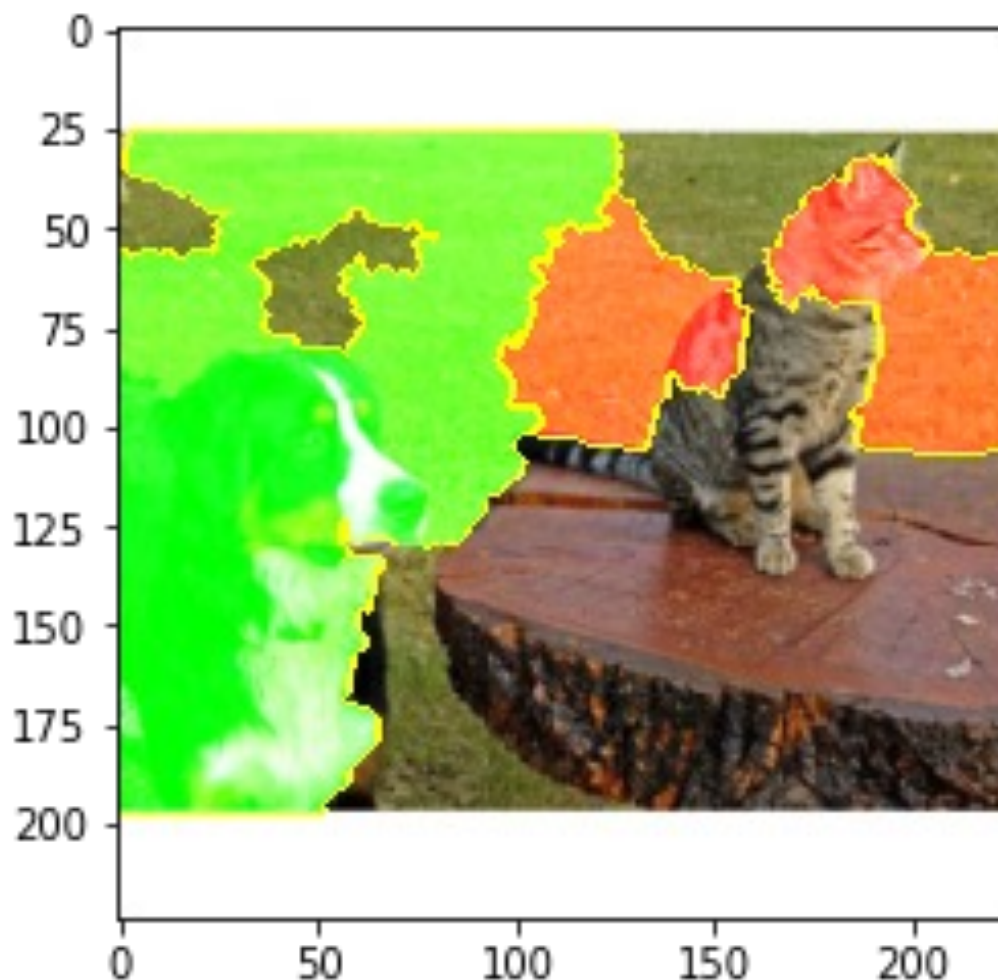


# A closer look at dog

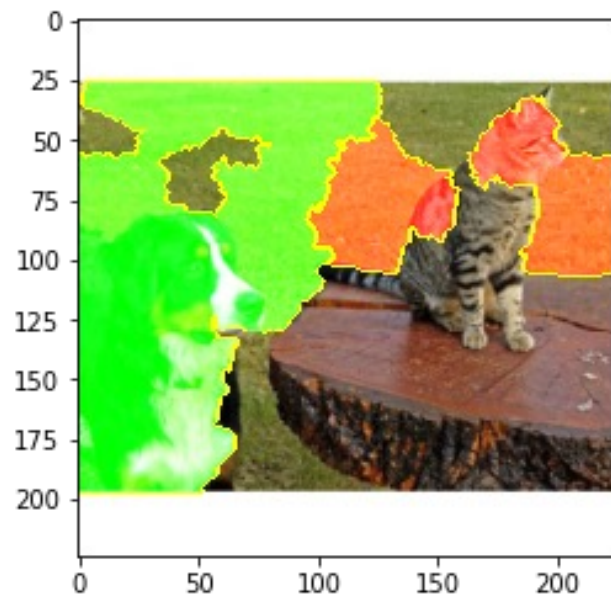
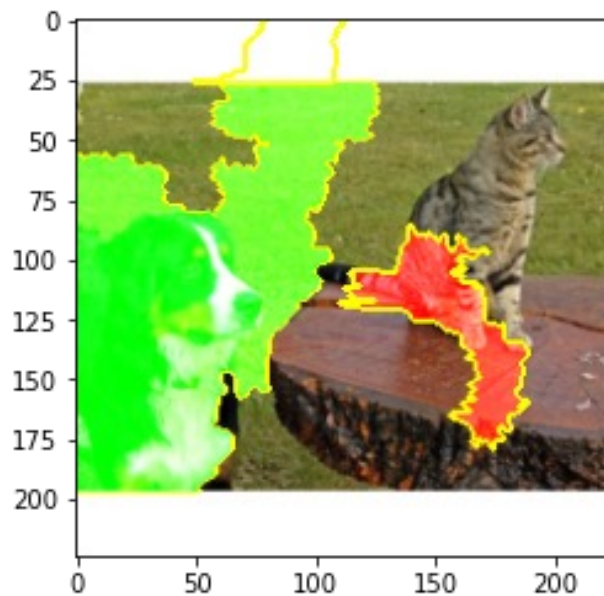
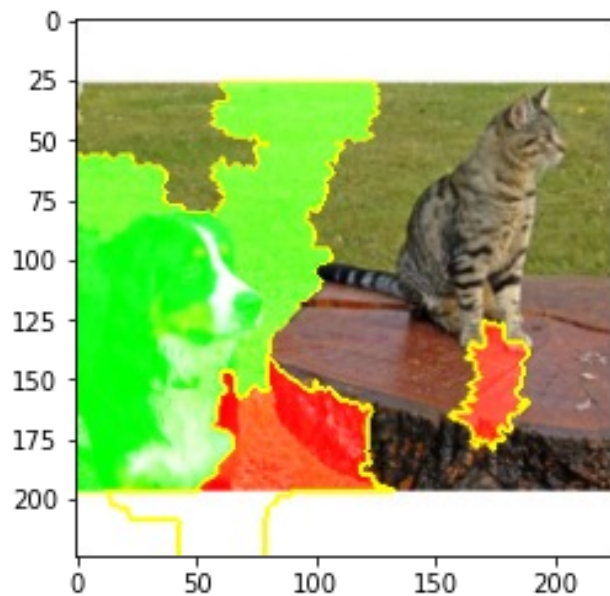




# A closer look at dog



# A closer look at dog



# Local linearity?



Mask=0



Mask=0.5



Mask=1

# Local linearity?

