



Aalto University  
School of Science  
and Technology

# Multilabel classification through structured output learning

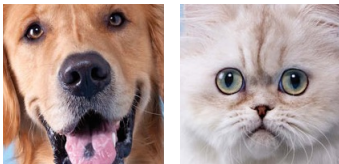
Hongyu Su

Department of Computer Science  
School of Science, Aalto University  
[hongyu.su@aalto.fi](mailto:hongyu.su@aalto.fi)

March 24, 2015

# Example: dog vs. cat?

- ▶ We have 5000 pictures of dog and 5000 pictures of cat.



- ▶ Computer digitalize each picture into  $100 \times 100$  pixels.
- ▶ Given a new picture, we want to answer: is it a dog or a cat?
- ▶ Simple task for human, dog, or cat.
- ▶ Golle (2008) claimed this is a difficult task for machines with only 82.7% accuracy.
- ▶ In 2013, 98.5% accuracy was reported in a Kaggle competition (<https://www.kaggle.com/c/dogs-vs-cats>).

# In human verification system

- ▶ Human verification system is a program that protects website from robots by generating and grading test that human can pass but machine cannot.
- ▶ CAPTCHA system (Ahn et al., 2003) uses distorted text.



- ▶ ASIRRA system (Elson et al., 2007) uses images.

Asirra

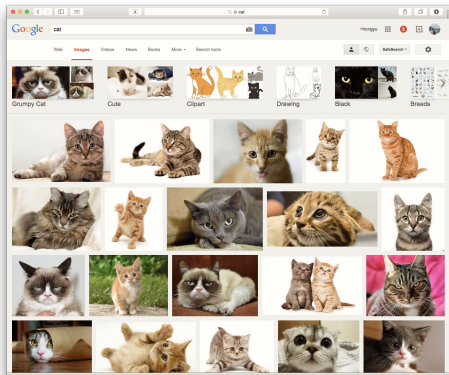
Asirra is a human interactive proof that asks users to identify photos of cats and dogs. It's powered by over two million photos from our unique partnership with [iStockphoto.com](#). Protect your web site with Asirra -- free!



- ▶ To test if the ASIRRA system is safe from machine learning attack.
  - ▶ One should get all 12 pictures right!
  - ▶ Accuracy for machine is  $(98.5\%)^{12} \approx 83.4\%$ .

# In search engine

- ▶ If machine can assign cat/dog to all pictures correctly, we can search pictures with keywords.
- ▶ Search all cat pictures.



# In search engine

- ▶ If machine can assign cat/dog to all pictures correctly, we can search pictures with keywords.
- ▶ Search all dog pictures.



# Single label classification

- ▶ In machine learning, the problem is known as *single label classification*.
  - ▶ Input is an object  
e.g., an image.
  - ▶ Output is an attribute of the object called *label*  
e.g., dog or cat?
  - ▶ Explore a set of known object and label pairs called *Training data*  
e.g.,  $\{(\text{image\#1}, \text{dog}), \dots, (\text{image\#5001}, \text{cat}), \dots\}$ .
  - ▶ Learn a *mapping function* that predict the label of a new object  
e.g., (new image, dog or cat?)
- ▶ Mathematically, we define the single label classification problem
  - ▶ Data come in pairs  $(\mathbf{x}, y) \in \mathcal{X} \times \mathcal{Y}$ , sampled from some unknown distribution  $P(\mathbf{x}, y)$ .
  - ▶  $\mathcal{X} = \mathbb{R}^d$  is a domain of input,  $\mathcal{Y} = \{+1, -1\}$  is a domain of output.

# Future work



# To get benefit?

- ▶ Fingerprint identification
- ▶ Voice recognition
- ▶ Information assistant



# To contribute?

- ▶ SETI@home
- ▶ Rosetta@home
- ▶ Foldit

# Bibliography

- Ahn, L. V., Blum, M., Hopper, N. J., and Langford, J. (2003). Captcha: Using hard ai problems for security. In *Proceedings of the 22Nd International Conference on Theory and Applications of Cryptographic Techniques, EUROCRYPT'03*, pages 294–311, Berlin, Heidelberg. Springer-Verlag.
- Elson, J., Douceur, J. R., Howell, J., and Saul, J. (2007). Asirra: A captcha that exploits interest-aligned manual image categorization. In *Proceedings of 14th ACM Conference on Computer and Communications Security (CCS)*. Association for Computing Machinery, Inc.
- Golle, P. (2008). Machine learning attacks against the asirra captcha. In *Proceedings of the 15th ACM Conference on Computer and Communications Security, CCS '08*, pages 535–542, New York, NY, USA. ACM.