

E6893 Big Data Analytics

NICE DRAWING

- *A tool that reproduces similar looking line drawings from sketches or photos*

Project ID: 201812-22

Team Members: Xueyao Li (xl2719) Yiyi Zhang (yz3280)

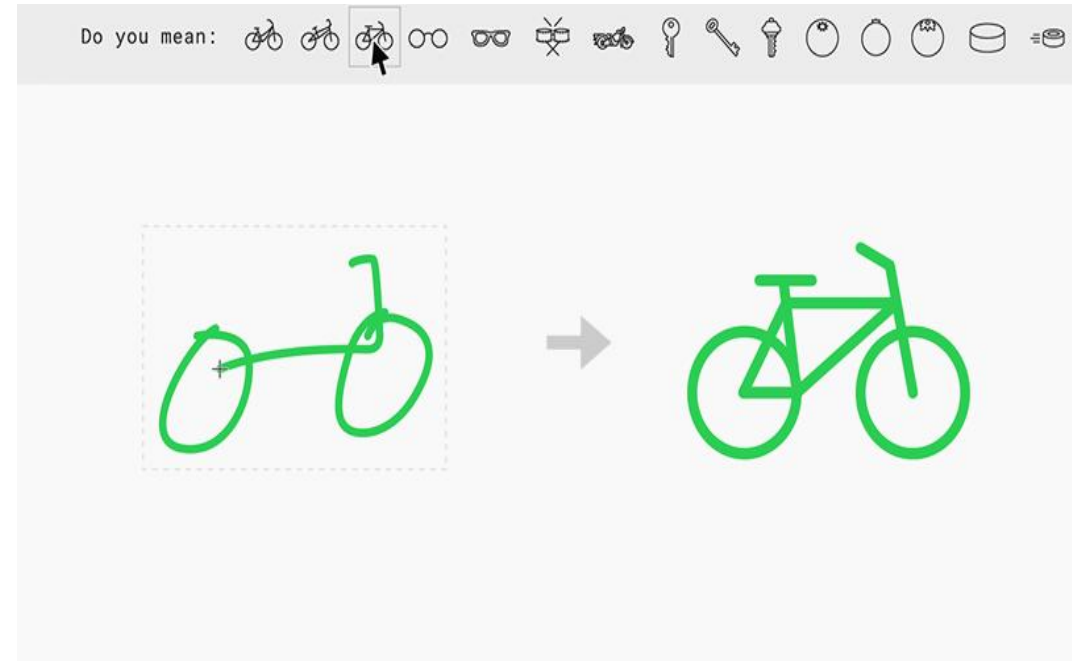
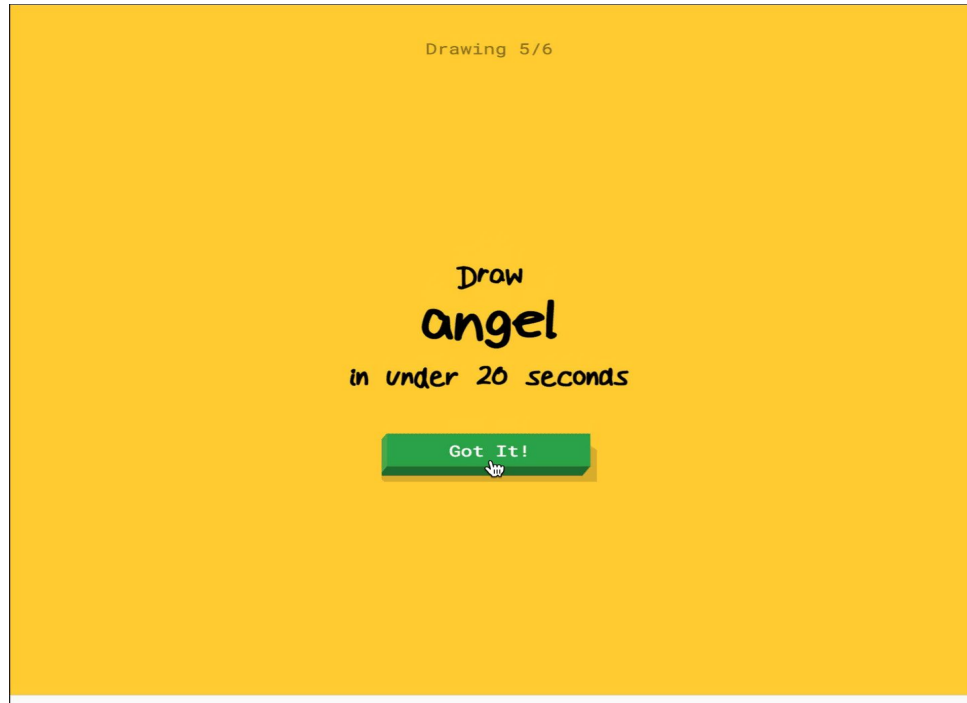


Agenda

- ❖ **Introduction & Motivation**
- ❖ **Dataset**
- ❖ **Algorithms & Tools**
- ❖ **Demo**
- ❖ **Contributions & References**

Introduction & Motivation

- ❖ Inspired by [Quick, Draw!](#) and [AutoDraw](#) developed by Google
- ❖ Explore the application of the state-of-the-art technologies toward visual art
- ❖ Make the drawing experiences easier, faster, and more fun to the general public
- ❖ Assist the creative process of professional artists and designers and help them expand imagination



Dataset

The Quick, Draw! Dataset: 50 millions of drawings across 345 categories

❖ **Categories: 15 animal categories**

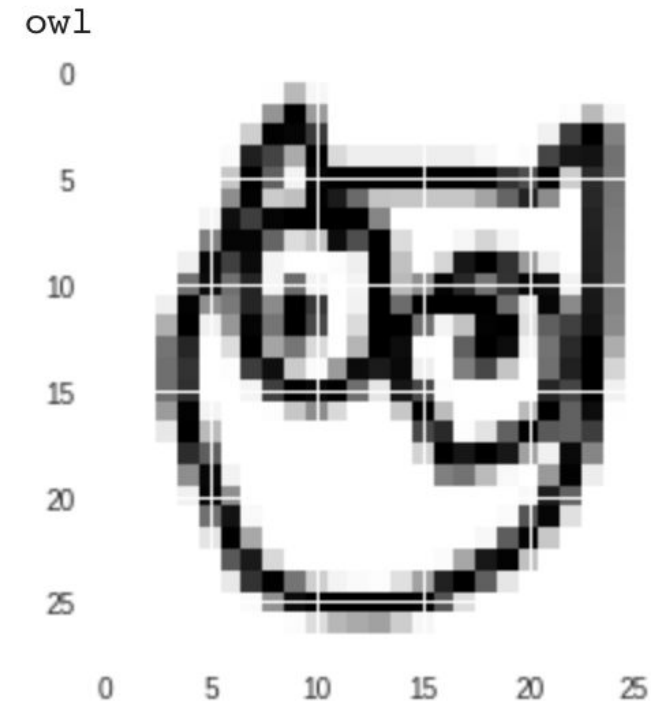
ant	124,612	dog	152,159	owl	169,632
bear	120,890	duck	135,480	penguin	253,791
bee	134,762	flamingo	124,569	pig	186,770
bird	133,572	frog	159,047	snail	133,757
cat	123,202	octopus	150,152	tiger	121,067

❖ **Data Size: 1.85G**

Classification: 100K samples (80K training & 20K test) per category

❖ **Data Format: Numpy Bitmaps (.npy)**

The drawings have been rendered into a 28x28 grayscale bitmap in numpy .npy format.



Algorithms & Tools

Algorithms:

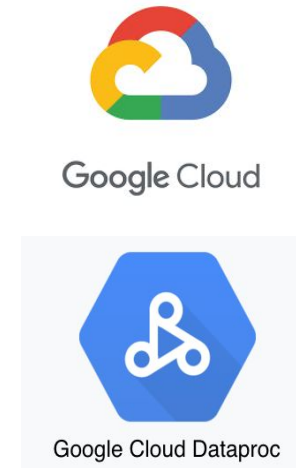
❖ Classification:

- **Drawings:** Logistic Regression, CNN
- **Photographs:** TensorFlow Image Recognition and Object Detection API

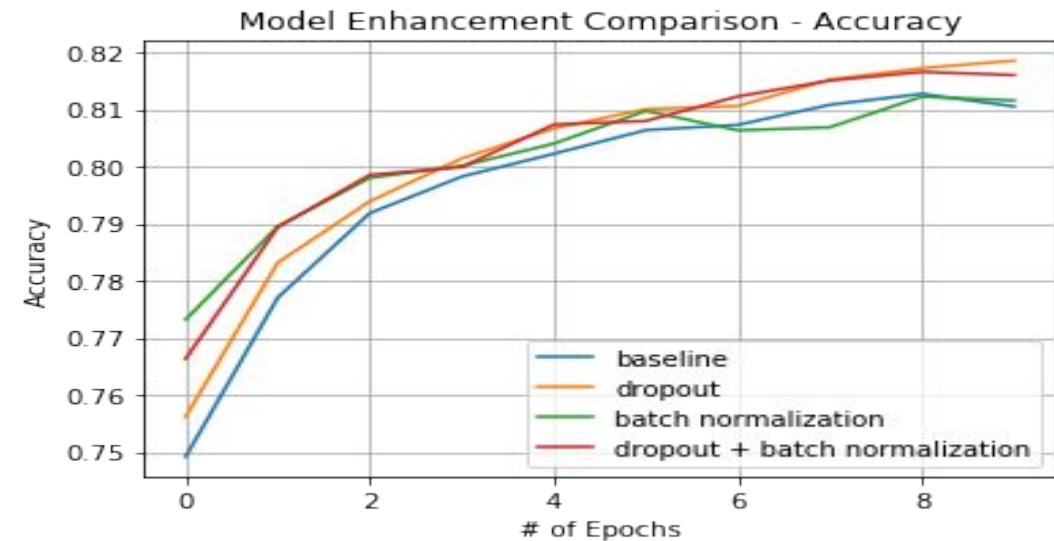
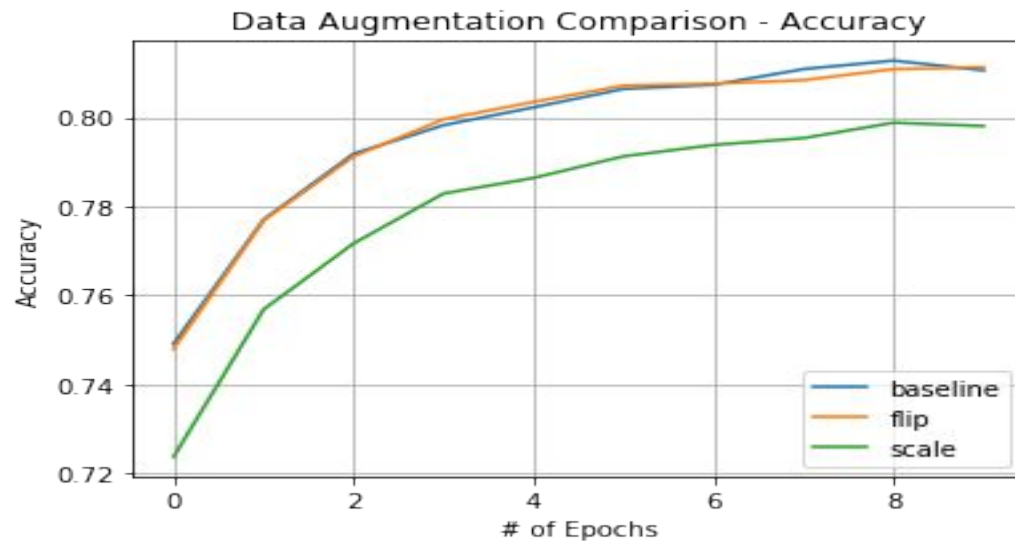
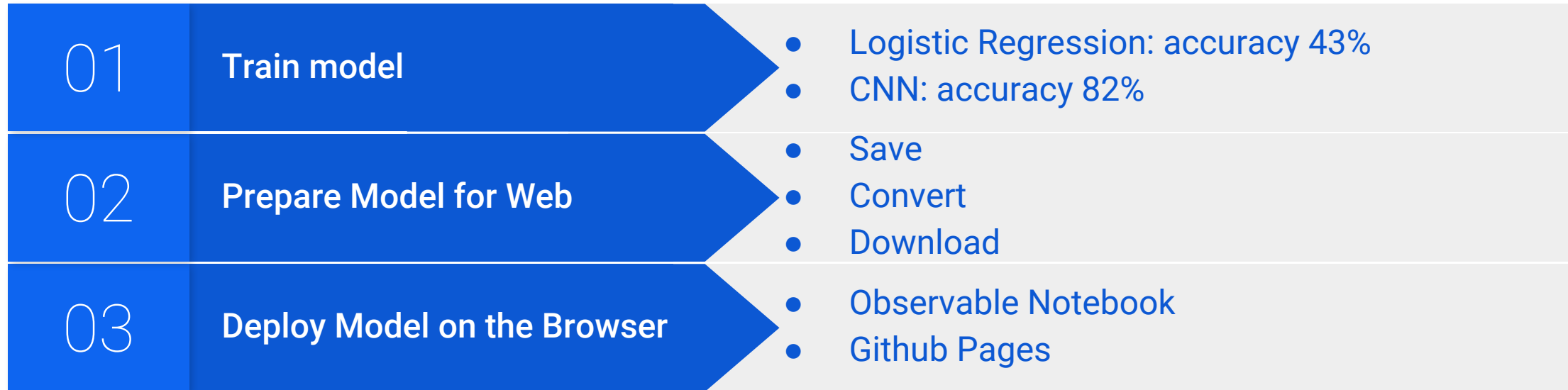
❖ Visualization: Use t-SNE to visualize similarities between drawings.

Tools:

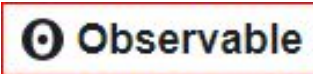
Google Cloud Platform, Google Cloud Dataproc, Spark
TensorFlow, TensorFlow.js,
Jupyter Notebook, Google Colab, Observable Notebook, GitHub Pages



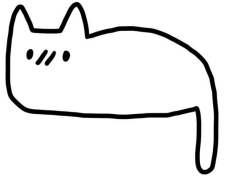
Pipeline



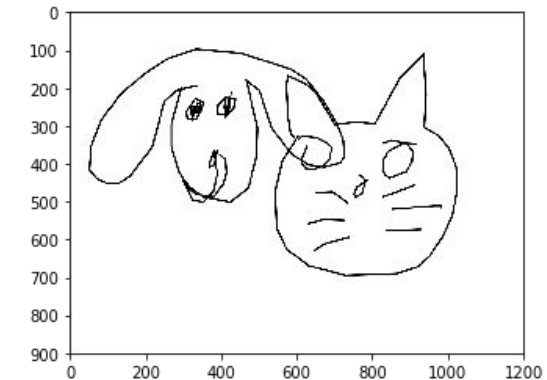
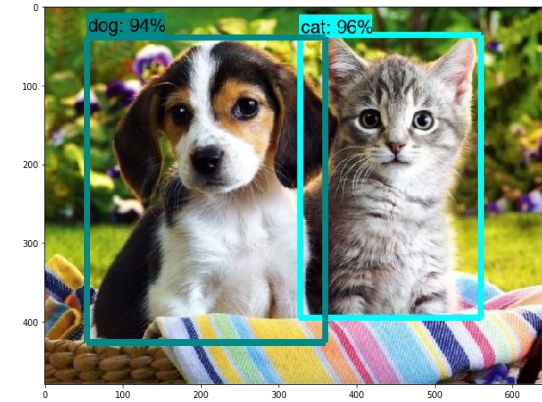
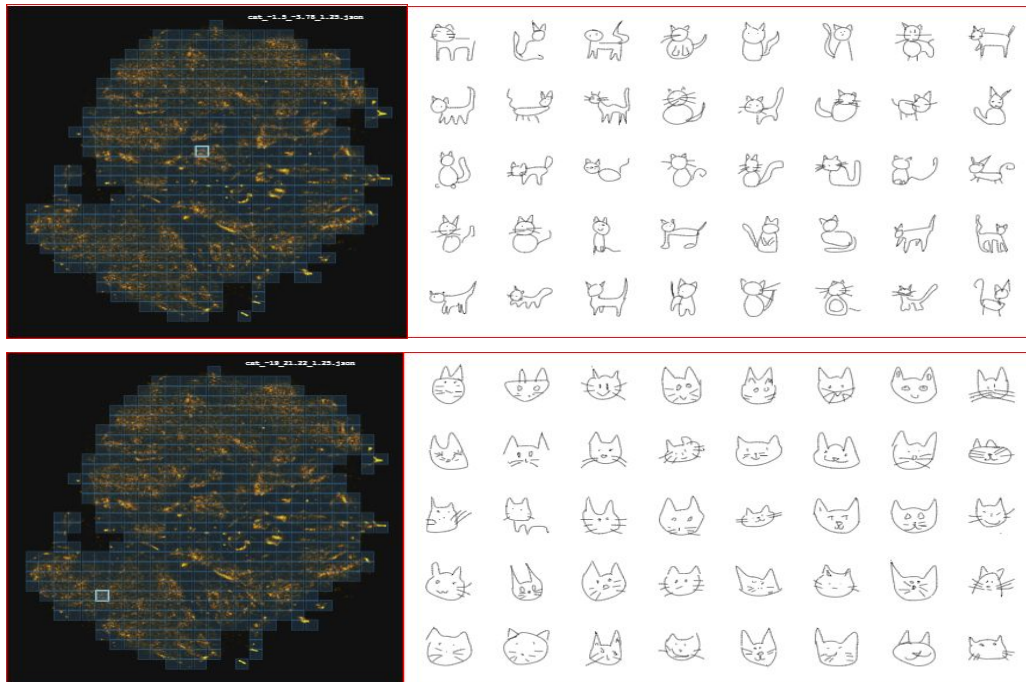
Demo



❖ Classification: drawings or photographs



❖ Visualization: all drawings in the recognized categories



Contributions & References

Contributions:

- ❖ **Xueyao Li:** Classification (Drawings), Big Data Frameworks, Website (Github Pages)
- ❖ **Yiyi Zhang:** Classification (Photographs), t-SNE Visualization, Website (Observable), Final Report, and Presentation Videos

References:

- ❖ The Quick, Draw! Dataset <https://github.com/googlecreativelab/quickdraw-dataset>
- ❖ TensorFlow Image Recognition and Object Detection API <https://github.com/tensorflow/tfjs-models>
- ❖ Machine Learning for Visualization <https://medium.com/@enjalot/machine-learning-for-visualization-927a9dff1cab>
- ❖ Deploy model on the browser using TensorFlow.js <https://github.com/zaidalyafeai/zaidalyafeai.github.io/tree/master/sketcher>

Thank You!