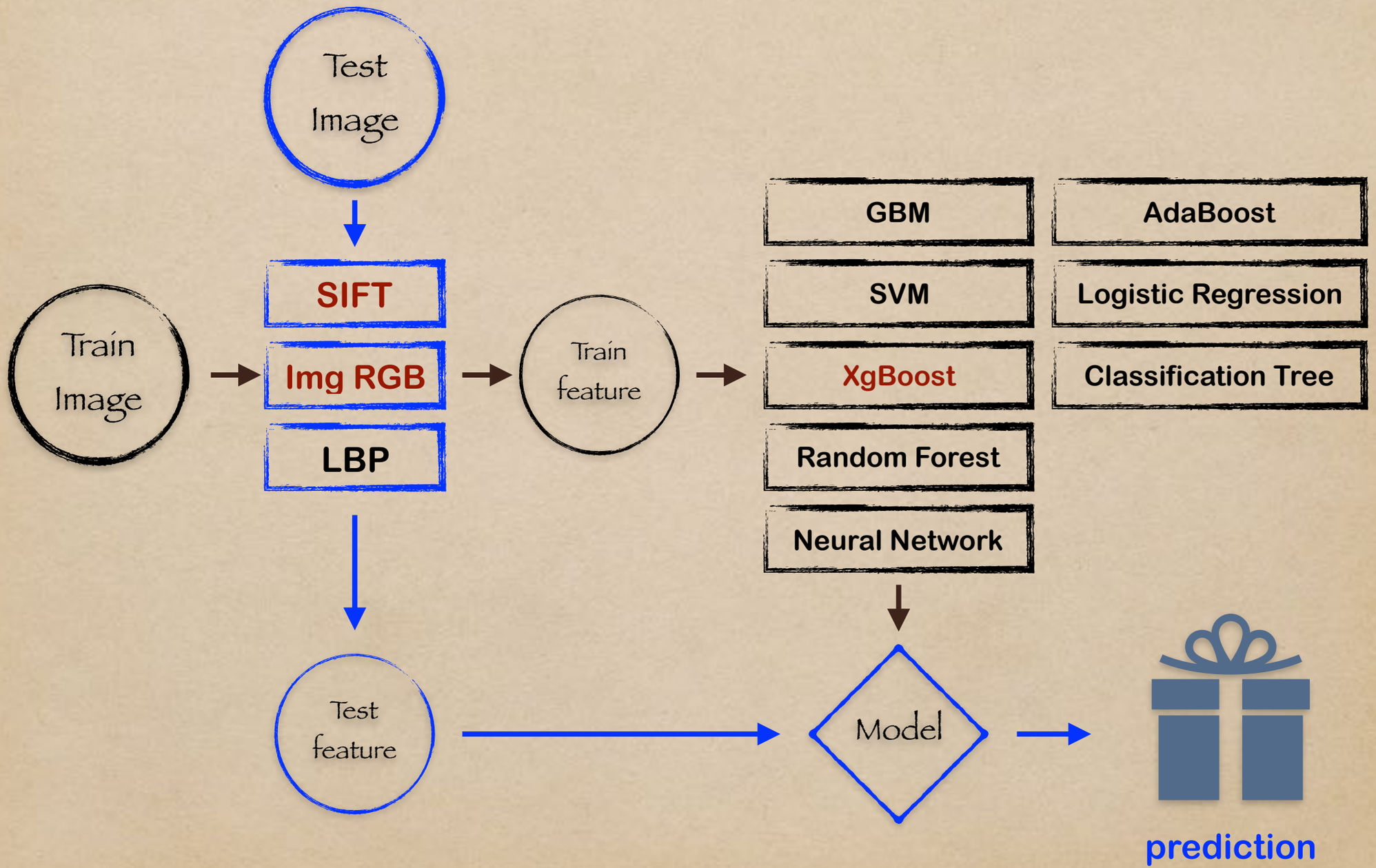
Project3-Group3

Dogs, Fried Chicken, or Blueberry Muffins?

Keran Li, Mingming Liu, Zhongxing Xue, Yuhan Zha, Junkai Zhang

SIX Model



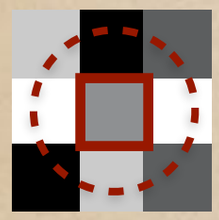
Dogs, Fried Chicken, or Blueberry Muffins?

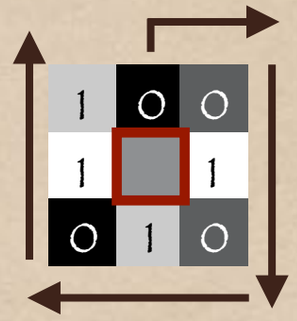
SIX Model

- Feature Selection — SIFT, LBP, Image RGB
- Classification Model
- Result

LBP Feature

(Local Binary Patterns)





$(00101011)_2$

LBP#43

Calculate frequency

■ < ▢ : 0

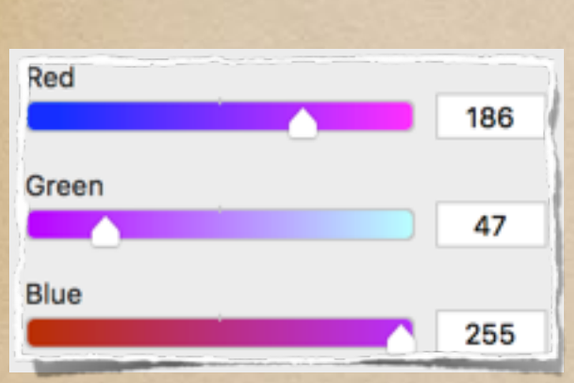
□ > ▢ : 1

- ✗ lack of color change
- ✗ Scale & spin Variance
- ✗ Huge time cost!

Dogs, Fried Chicken, or Blueberry Muffins?

Image RGB Feature

RGB position → (186,47,255) → Subgroup → (9,3,12) → f(a,b,c) = 144(a-1) + 12(b-1)+c → Color#1188



Range	Subgrp
0-21	1
22-42	2
43-64	3
65-85	4
86-106	5
107-128	6
129-149	7
150-170	8
171-192	9
193-213	10
214-234	11
235-255	12

Calculate frequency of each number as a RGB feature.

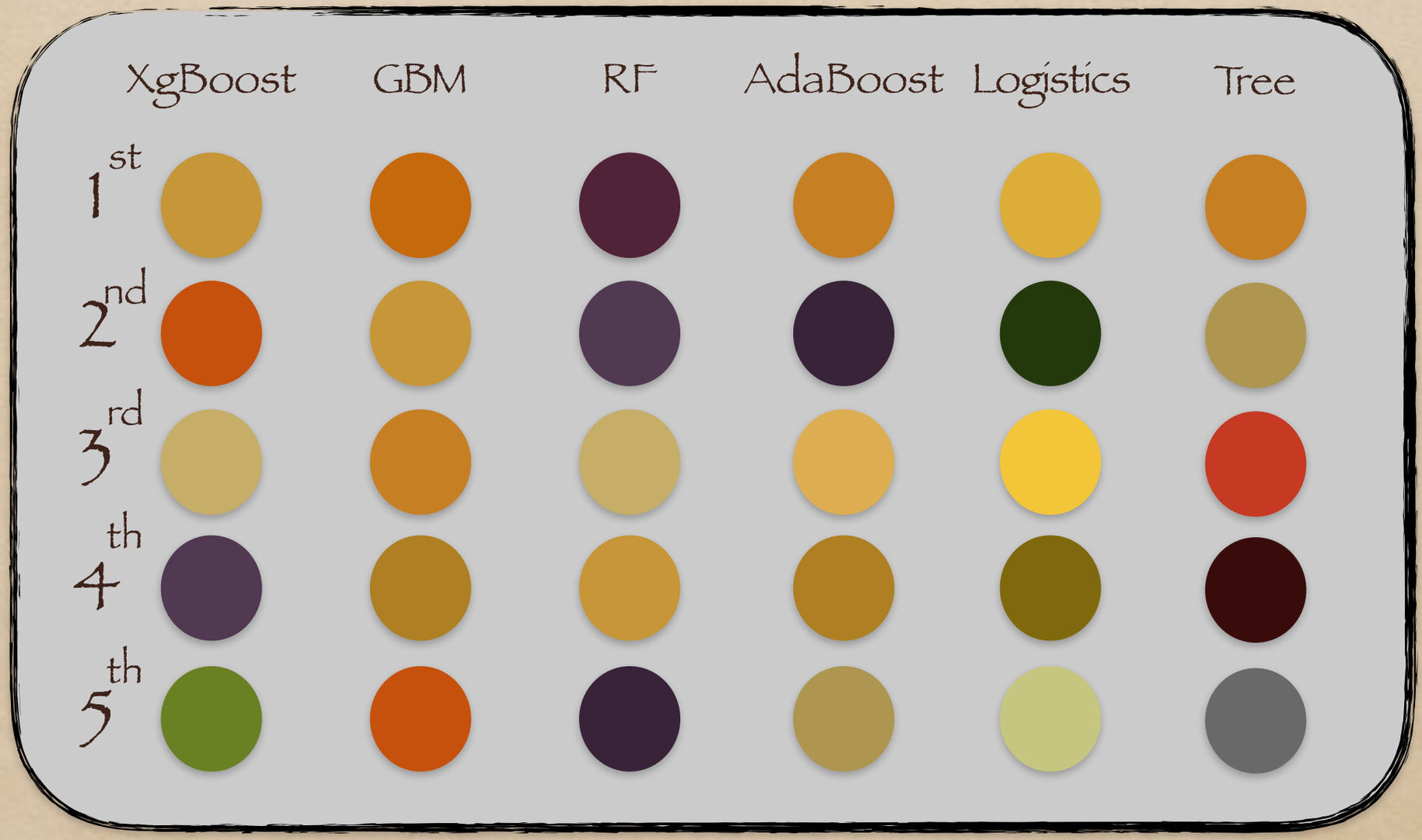
Now we have 1728 new features:
Color#1 ~ Color#1728

- ✓ Time cost: 15~20 mins
- ✓ Huge increase of accuracy

Dogs, Fried Chicken, or Blueberry Muffins?

Image RGB Visualization

(With feature: only ImageRGB)



Dogs, Fried Chicken, or Blueberry Muffins?

SIX Model

- Feature Selection
- Classification Model — XgBoost
- Result

Classification Model

(Up to 2018.03.21)

92.06%

XgBoost

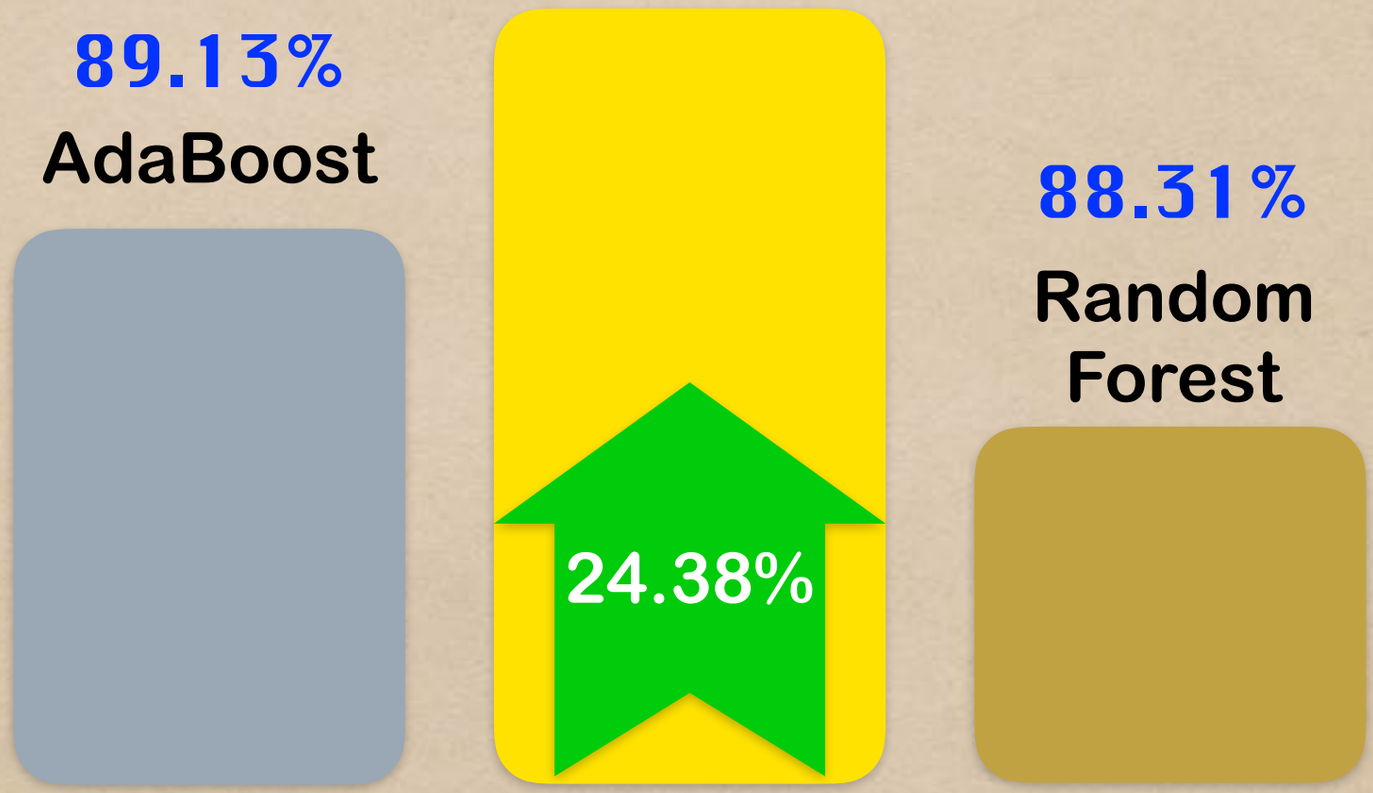
(With feature: SIFT+ImageRGB)

89.13%

AdaBoost

88.31%

Random Forest



GBM Base model

67.68%

(With feature: SIFT)

Dogs, Fried Chicken, or Blueberry Muffins?

SIX Model

- Feature Selection
- Classification Model
- Result

Accuracy(%) + Time(s)

(Up to 2018.03.21)

Features \ Algorithm	GBM	XgBoost	Adaboost	RF	Logistic	SVM	Tree	NN
S 2000d	67.68	70.35	65.17	64.64	68.34	71.66	53.77	72.74
S+I 2512d	83.75	88.69	85.57	86.09	79.45	83.36	78.02	49.37
S+I 3728d	86.24	92.06	89.13	88.31	82.80	83.77	79.50	48.83
I 1717d	84.30	90.65	87.48	87.21	71.11	74.97	71.11	30.12
S+I+L 3949d	86.73	87.40	86.46	84.85	79.62	54.96	79.62	58.44

Dogs, Fried Chicken, or Blueberry Muffins?

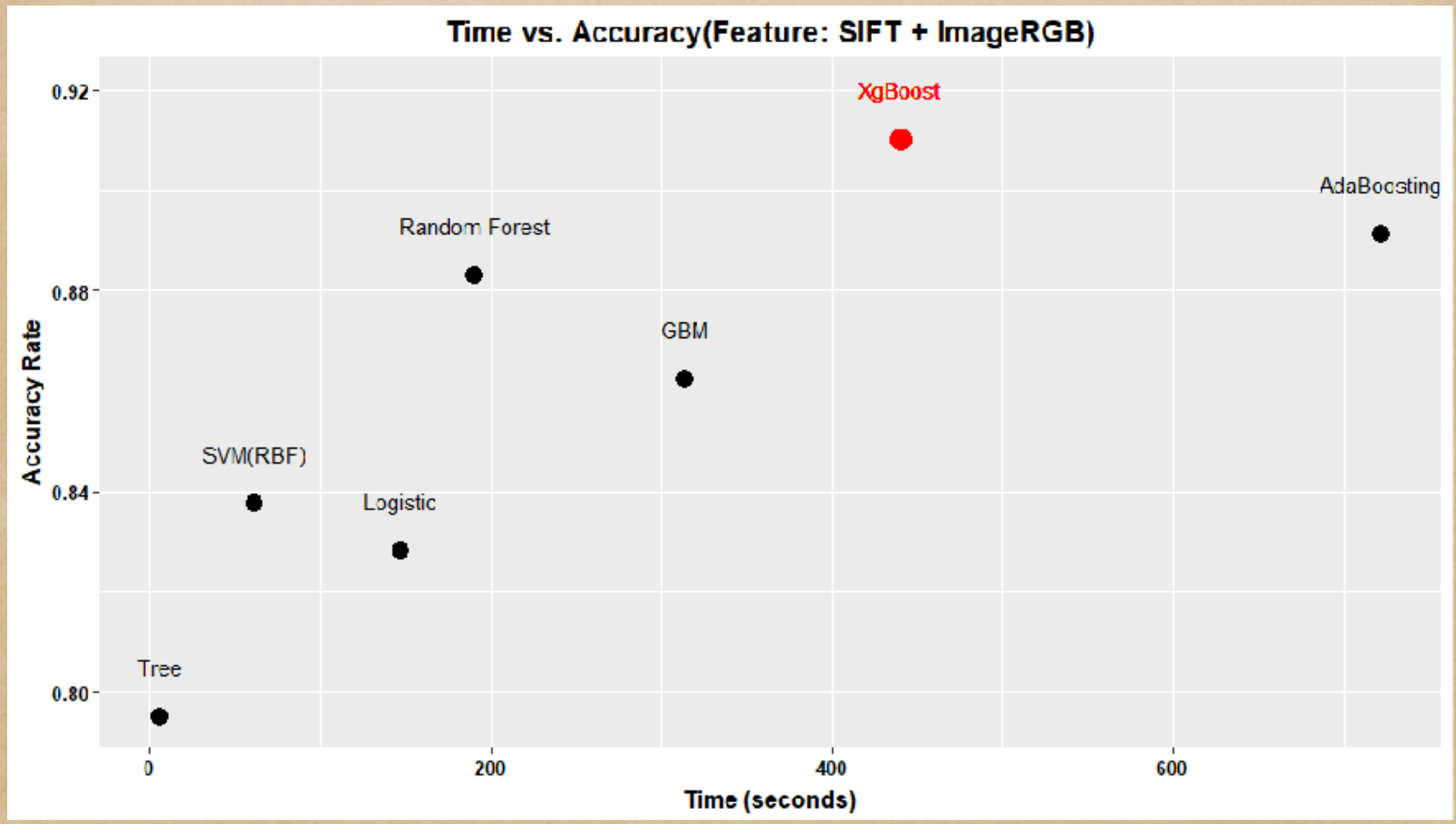
Accuracy(%) + Time(s)

(Up to 2018.03.21)

Features \ Algorithm	GBM	XgBoost	Adaboost	RF	Logistic	SVM	Tree	NN
S 2000d	67.68 178.34	70.35 375.78	65.17 5.96	64.64 188.09	68.34 55.54	71.66 43.11	53.77 2.23	72.74 >999
S+I 2512d	83.75 219.04	88.69 389.36	85.57 503.74	86.09 122.98	79.45 50.97	83.36 36.58	78.02 2.79	49.37 >999
S+I 3728d	86.24 313.56	92.06 440.56	89.13 721.95	88.31 190.51	82.80 146.64	83.77 61.21	79.50 5.64	48.83 >999
I 1717d	84.30 146.32	90.65 453.77	87.48 284.36	87.21 84.90	71.11 87.21	74.97 25.44	71.11 1.79	30.12 >999
S+I+L 3949d	86.73 333.12	87.40 463.39	86.46 932.91	84.85 227.16	79.62 153.2	54.96 27.16	79.62 3.83	58.44 687.00

Dogs, Fried Chicken, or Blueberry Muffins?

Result



Dogs, Fried Chicken, or Blueberry Muffins?

About train data

0784.jpg: Chicken, 0/8

Dogs, Fried Chicken, or Blueberry Muffins?

Project3-Group3

Dogs, Fried Chicken, or Blueberry Muffins.

Keran Li: Neural Network, Logistic
Mingming Liu: AdaBoost, Random Forest
Zhongxing Xue: GBM
Yuhan Zha: SVM, Classification Tree
Junkai Zhang: XgBoost

ppt: Zhongxing Xue
presenter: Zhongxing Xue

Good evening every one,I'm Zhongxing Xue, the presenter of Group3. On behalf of our entire group, I'd like To share some interesting ideas and wonderful result of this project to you all.

Let's introduce our model, SIX model. In our model, we transfer
Original picture to SIFT features and ImageRGB features, and use XgBoost to train and predict.
SIX is the abbreviation of those method.

There are hundreds of ways to select feature from a picture. After discussions and trails, we focus on three Candidate features: SIFT, LBP, and Image RGB. SIFT is quite good, and was selected in our model.
LBP features shows local texture very well. But there are some disadvantages:
It couldn't differentiate color change, and it is not stable with zooming and rotating. And the worst point is that it cost huge of times. In a trade-off decesion, we throwed LBP away. 30s

Image RGB is a type of color feature. For each pixel we pick it's RGB color position, then we use a 1 - to 1 function to transform that 3 dimensions Vector to a number, with range of 1 to 1728. By calculating the frequency of each number, we got those RGB features. We chose Image RGB features because it made a huge increase in our model and it's time cost is acceptable.

Besides, RGB features provide intuitional and concise visualization report. This plot shows the top 5 important color in each Classification model. These critical colors can be divided into three groups: first is dark yellow and orange, these are The main color of chickens, muffins, and most of dogs. So it makes sense that the algorithm thought these are important,
Second grouo: black and dark purple, main color of blueberry, we believe these method could be more accurate classifying muffins and others. Third: green, after we check the picture, we found that for the most of dog's picture, the background is grassland.
Green help algorithm to classify dog with other objects better
This corresponds with our expectation, and indicating that RGB feature would offer a reasonable and accurate classification standard . 2min

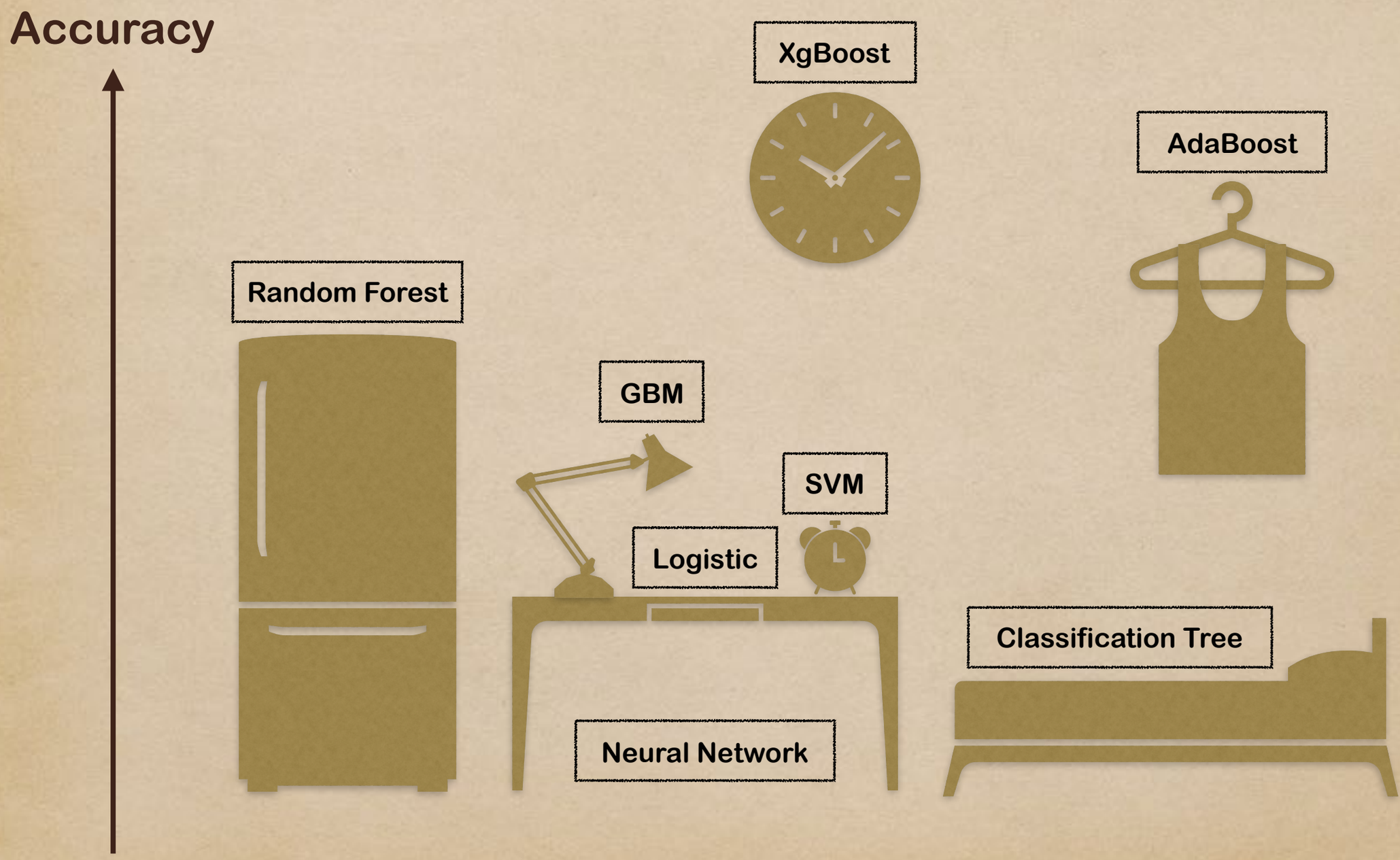
Now let's move to classification model part. In this step,. we got 8 algorithms, and select the
Best one as our classification method in the SIX model. The top algorithm is Xgboost, with 23.38% of increase.
All of the algorithm included the CV step. XgBoost is no exception.

Now for the result part, we paired one different feature set with one algorithm, and made 40 compared result. The top left is Baseline model, red one is our SIX model. Plus time cost, it shows that the SIX model did training part within 10minutes. Check
The time-accuracy plot, SIX model has the highest accuracy position, with time-position not too much.
So, We believe that SIX model is an efficient and accurate classification model for this project.

At last, we did some analysis about the training data. We found a funny picture that belongs to chicken class, but none of our Algorithm returned true answer. Probably we all consider it as a lovely dog eating muffins.

That's all of out presentations, thank you very much.

Classification Model



Dogs, Fried Chicken, or Blueberry Muffins?