Model Selection and Evaluation

Dogs, Fried Chicken or Blueberry Muffins?

Group4: Lin Yanjun; Jiang Chenfei;
Tao Wesley; Wan Qianhui; Yao Jingtian

Model Selection and Evaluation

1. Problem description
2. Model selection
3. Model assessment and comparison
4. Model improvement and prospect
5. Reference

1. Problem description

Dogs, Fried Chicken or Blueberry Muffins?

- **image classification**: the task of extracting information classes from a multiband raster image.
- **multiclass classification**: classifying instances into one of three or more classes.

2. Model selection

2.1 Terminology

- **Hypothesis/model**: a certain function that we believe (or hope) is similar to the true function, the target function that we want to model.
- **Learning algorithm**: a set of instructions that tries to model the target function using our training dataset.
- **Classifier**: a hypothesis or discrete-valued function which is used to assign (categorical) class labels to particular data points.
- **Hyperparameters**: tuning parameters of a machine learning algorithm.

(while **model parameters** are the parameters that a learning algorithm fits to the training data)

2. Model selection

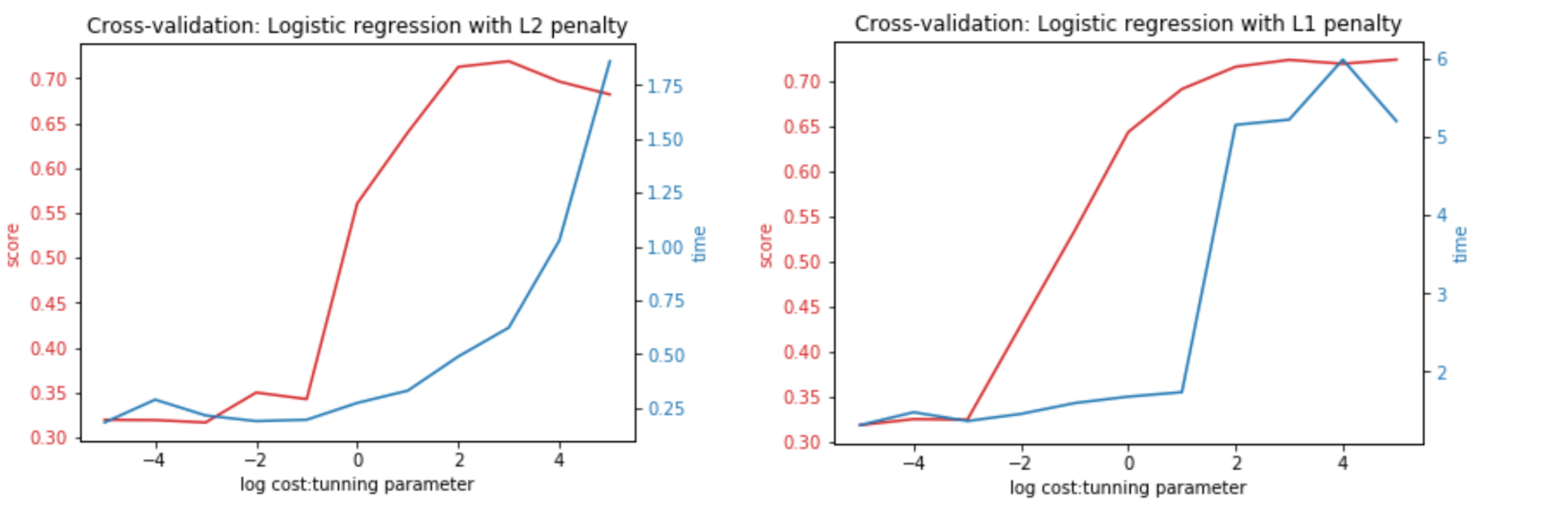
2.2 feature extraction

- **ORB(Oriented FAST and Rotated BRIEF):** a fusion of FAST keypoint detector and BRIEF descriptor with many modifications to enhance the performance.
- **SIFT(Scale Invariant Feature Transform):** local and based on the appearance of the object at particular interest points, and are invariant to image scale and rotation.
- **SURF(speeded up robust features):** based on the similar principles and steps as SIFT; algorithm contains three main steps: interest point detection, local neighborhood description and matching.

2. Model selection

2.3 basic model comparision

feature	model		
	Logistic	QDA	LDA
SIFT	0.5827	0.3327	0.491
RGB	0.6407	0.553	0.579



Optimal Logistic Model:

- ORG feature
- L2 penalty
- Linear method
- Cost = 1000
- Performance: 5-fold-CV score: 0.8147 ; running time per round: 1.5828s.

2. Model selection

2.4 Alternative advanced model

2.4.1 SVM (Linear/RBF kernel)

- **Description**

Supervised learning models which constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space.

- **Advantage**

Effective in high-dimensional spaces, even when dim(n) > p

Memory effective: only depend on support vectors

Versatile: different Kernel functions can be specified for different decision functions

- **Weakness**

Risk of overfitting when n>>p

Calculation expensiveness: calculated using 5-fold CV

2. Model selection

2.4 Alternative advanced model

2.4.2 XGBoost (Extreme Gradient Boosting)

- **Description**

An advanced implementation of gradient boosting algorithm by adding new models sequentially until no further improvement is achieved.

- **Advantage**

Regularization: ‘regularized boosting’ technique, helps to reduce overfit.

Parallel Processing: fast computations

High flexibility: allow to define custom optimization objectives and evaluation criteria.

Have in-built routine to handle missing values

- **Weakness**

Risk of overfitting when not having enough data

2. Model selection

2.4 Alternative advanced model

2.4.3 CNN (MobileNet)

- **Description**

"Vision begins with eyes, but truly takes place in the brain."

Mostly based on an artificial neural network; using a cascade of multiple layers of nonlinear processing units for feature extraction and transformation .

- **Advantage**

High performance with enough data

Time efficient: reduces the need for feature engineering

Universality: can be adapted to new problems relatively easily

- **Weakness**

Extremely computationally expensive to train

Without strong theoretical foundation, hard to comprehend

3. Model assessment and comparison

3.1. Cost evaluation

- feature dimensions

RGB: 5*5*5   SIFT: 2000   MobileNet: 256*256*3

- model running time (cross validation/training/prediction time)

XGBoost (9675.41s/1.46s/0.04s)

SVM (Linear 0.41/3.18/0.40s   RBF 2.59/2.45/0.33s)

CNN (resize 118.69/548.78/22.68s)

3.2 Performance evaluation

- Accuracy

Baseline	SVM(linear)	SVM(RBF kernal)	XGBoost(RGB feature)	XGBoost(SIFT feature)	MobileNet
0.883	0.589	0.742	0.889	0.678	0.993

4. Model improvement and prospect

4.1 Model improvement

4.1.1 Advanced feature

- **SIFT**

Invariant to image scale and rotation.

Robust to changes in illumination, noise, and minor changes in viewpoint.

Highly distinctive, relatively easy to extract and allow for correct object identification with low probability of mismatch.

- RGB (more relevant to the classification problem characteristics)

Gradient-based features: makes the scheme robust to illumination variations whereas use of orientation information to define features provides robustness against contrast variations.

4. Model improvement and prospect

4.1 Model improvement

4.1.2 Parameter tuning

- **SVM model**

  Linear: cost: (0.0001,0.001,0.1,1)

  RBF: cost: (0.0001,0.001,0.1,1) ; gamma: (0.01,0.1,1,10,100)

- **XGBoost model**

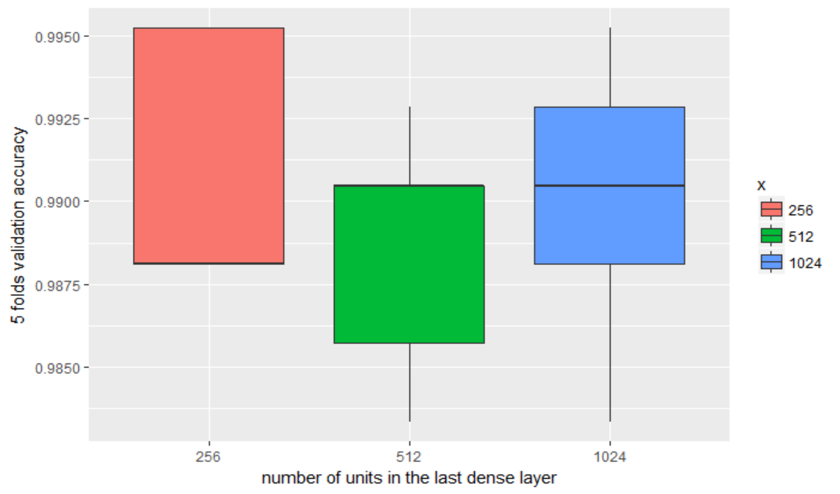
General Parameters

Booster Parameters (eta = 0.15, max_depth = 4)

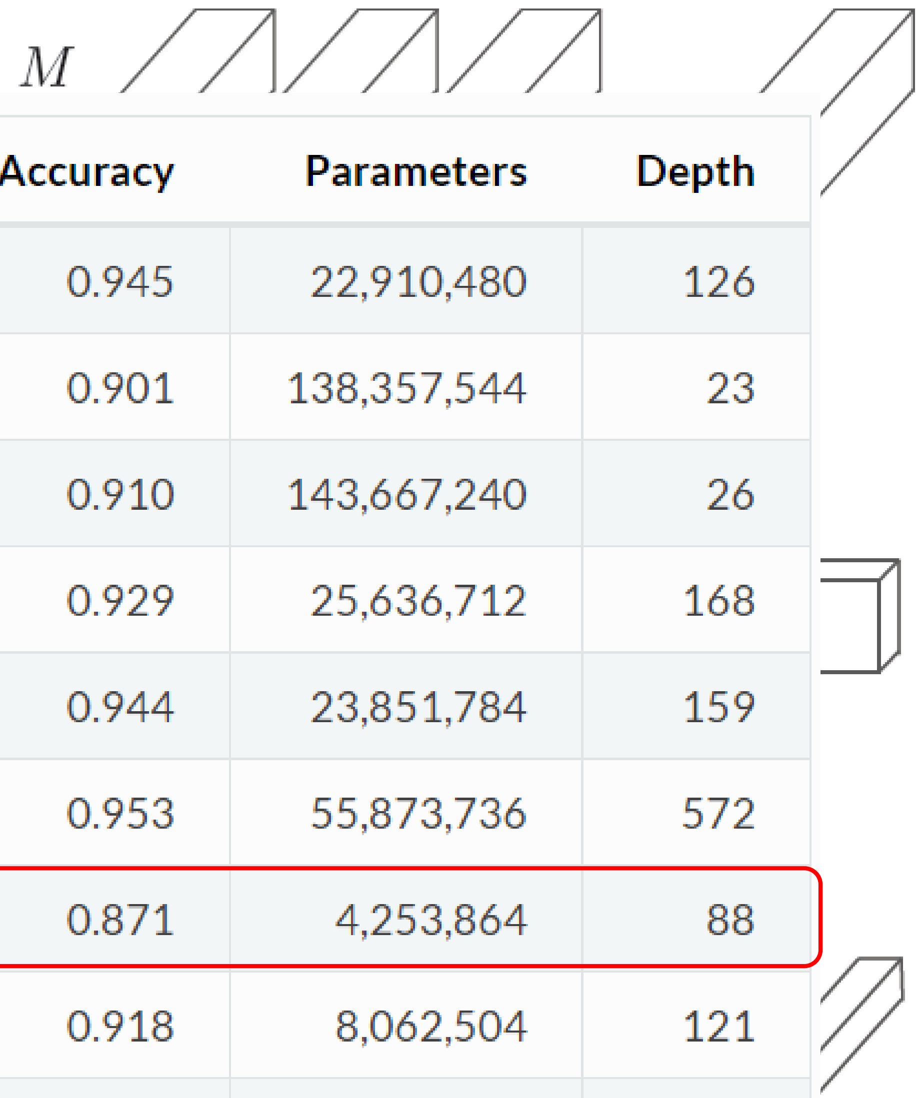
Learning Task Parameters

- **CNN(mobile net)**

Number of units in the last dense layer: 256



Model	Size	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth
Xception	88 MB	0.790	0.945	22,910,480	126
VGG16	528 MB	0.715	0.901	138,357,544	23
VGG19	549 MB	0.727	0.910	143,667,240	26
ResNet50	99 MB	0.759	0.929	25,636,712	168
InceptionV3	92 MB	0.788	0.944	23,851,784	159
InceptionResNetV2	215 MB	0.804	0.953	55,873,736	572
MobileNet	17 MB	0.665	0.871	4,253,864	88
DenseNet121	33 MB	0.745	0.918	8,062,504	121
DenseNet169	57 MB	0.759	0.928	14,307,880	169
DenseNet201	80 MB	0.770	0.933	20,242,984	201



(c) 1×1 Convolutional Filters called Pointwise Convolution in the context of Depthwise Separable Convolution

4. Model improvement and prospect

4.2 Prospect

4.2.1 Potential problem

Independence of SIFT feature contradicts the requirement of cross validation.

4.2.2 Further improvement

Standardization

More indicators besides accuracy

(ROC curve, gams & lift charts etc)

5. Reference

[1] Model evaluation, model selection, and algorithm selection in machine learning Sebastian Raschka

[2] Model Evaluation (Classification & Regression) Noureddin Sadawi

[3] Complete Guide to Parameter Tuning in XGBoost (with codes in Python) AARSHAY JAIN

[4] MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications

Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam

[5] Texture Feature Extraction of RGB, HSV, YIQ and Dithered Images using GLCM, Wavelet Decomposition Techniques Manisha Lumb, Poonam Sethi

[6] Xception: Deep Learning with Depthwise Separable Convolutions Franc¸ois Chollet

[7] The Effectiveness of Data Augmentation in Image Classification using Deep Learning Jason Wang, Luis Perez

[8] http://scikit-learn.org/stable/model_selection.html

[9] http://scikit-learn.org/stable/modules/svm.html

[10] https://keras.io/applications/#documentation-for-individual-models

[11] http://blog.csdn.net/u014380165/article/details/75142710

Q&A

Thanks for listening.