Facial Emotion Recognition

Project 3: Group 6

Mengyao He, Yuwei Tong, Yiqi Lei, Jiaqi Yuan, Qinzhe Hu

Table of Contents

Introduction

Simple emotion:

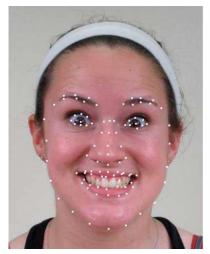Happy, sad, angry, disgusted, …

Compound emotion:

Happily surprised, angrily surprised, sadly disgusted, …

Introduction

- Input data: 6006 features extracted from 78 fiducial points
    - 3003 horizontal distances and 3003 vertical distances

Baseline Model – Gradient Boosting Machine (GBM)

- Parameters: n.trees=60, shrinkage=0.15, others=default
- Accuracy: 73.7%
- AUC: 74.6%
- Training time: 28.958s
- Testing time: 0.338s

Models We Tried

- Linear discriminant analysis (LDA)

- Support Vector Machine (SVM)

- Random Forest (RF)

- Neural Network (NN)

- Naive Bayes

- XGBoost

Summary of Performance

Model Name	Accuracy	AUC	Training Time	Testing Time
GBM - Baseline	73.7%	74.6%	28.958s	0.338s
Neural Network (oversample)	66.7%	84.4%	34.1s	331ms
LDA (with PCA)	82.7%	61.2%	0.090s	0.346s
Naive Bayes	80.0%	62.8%	0.13s	0.12s
SVM	79.5%	65.5%	132.069s	8.947s
Random Forest	80.3%	69.27%	940.659s	1.159s
XGBoost - Advanced	86.8%	70.1%	212.727s	0.129s

Advanced Model - XGBoost

- Parameters: eta=0.05, colsample_bytree=0.2, max_depth=6, min_child_weight=1, nrounds=300
- Accuracy: 86.8%
- AUC: 70.1%
- Training time: 212.727s
- Testing time: 0.129s

Feature Selection

- PCA
  - threshold = 0.95
  - 65 features
- Forward Stepwise Selection
  - 219 features
- Lasso
  - 117 features

Performance with Selected Features

Model Name	Accuracy	AUC	Training Time	Testing Time
XGBoost	**86.8%**	**70.1%**	**212.727s**	**0.129s**
XGBoost (with PCA)	82.7%	60.1%	5.988s	0.022s
XGBoost (with forward)	82.8%	63.7%	19.705s	0.031s
XGBoost (with lasso)	80.8%	51.3%	5.472s	0.027s
Random Forest	**80.3%**	**69.27%**	**940.659s**	**1.159s**
Random Forest (with PCA)	81.8%	78.4%	19.263s	0.076s
Random Forest (with forward)	74.2%	53.9%	112.774s	0.085s
Random Forest (with lasso)	73.3%	54.56%	44.972s	0.066s

Summary

Our Final Choice:

- Model: XGBoost (eta=0.05, colsample_bytree=0.2, max_depth=6, min_child_weight=1, nrounds=300)
- Input data: 6006 features
- Result: accuracy=86.8%, AUC=70.1%, training time=212.727s, testing time=0.129s
- Computer configuration: system: macOS Mojave version 10.14.6 | processor: 2.7 GHz Intel Core i5 | memory: 8 GB 1867 MHz DDR3

Reference

Github Link: https://github.com/TZstatsADS/Fall2020-Project3-group6