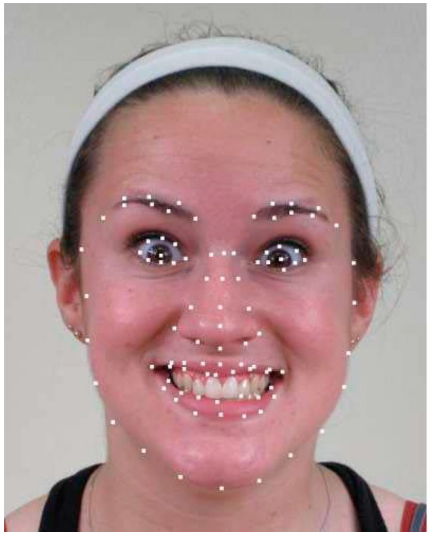
# Facial Emotion Recognition

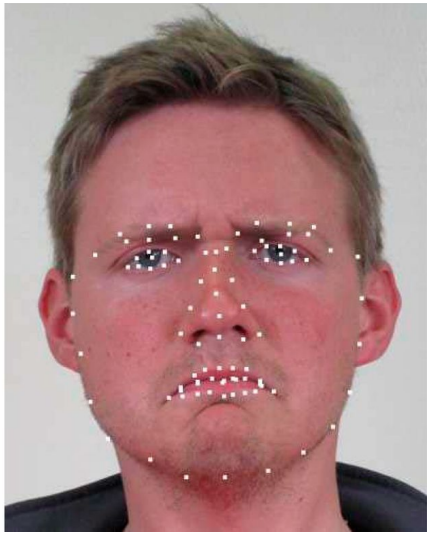
Group 1

Kristen Akey, Levi Lee, Yiran Lin,
Hanyi Wang, Wen Yin

# Introduction

## Can you recognize the emotion from an image of a face?





Categorizing two emotion types:

- Basic emotions
- Compound emotions

Goal: accurately recognize the emotion from facial images

Challenge: imbalanced data

# Baseline Model - Gradient Boosting Machine

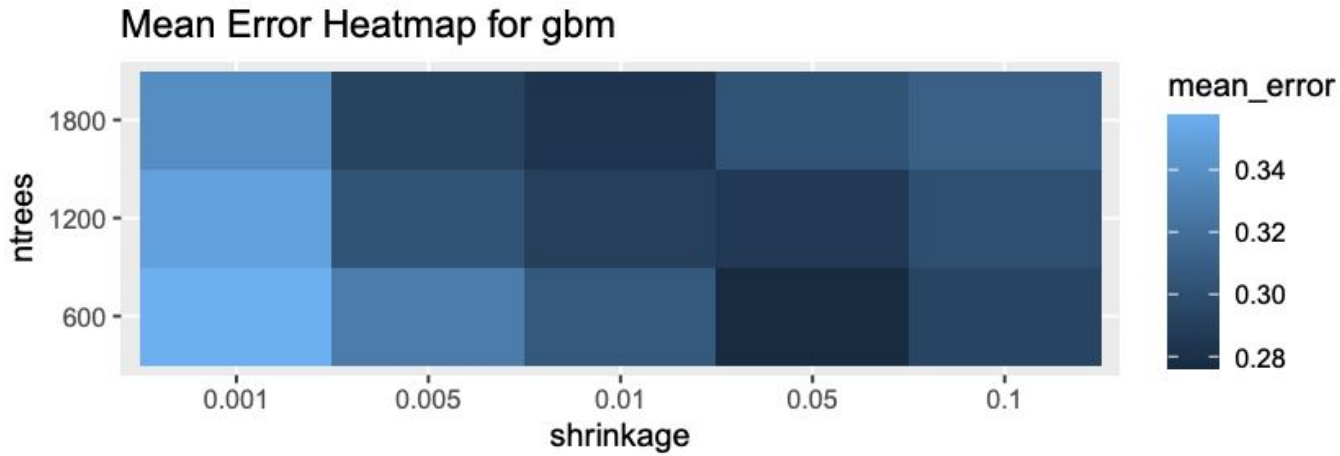
- Trained 15 models with cross-validation.
- Cross-validated ntrees and shrinkage:
  - ntrees: 600, 1200, and 1800
  - shrinkage: 0.001, 0.005, 0.010, 0.050, 0.100

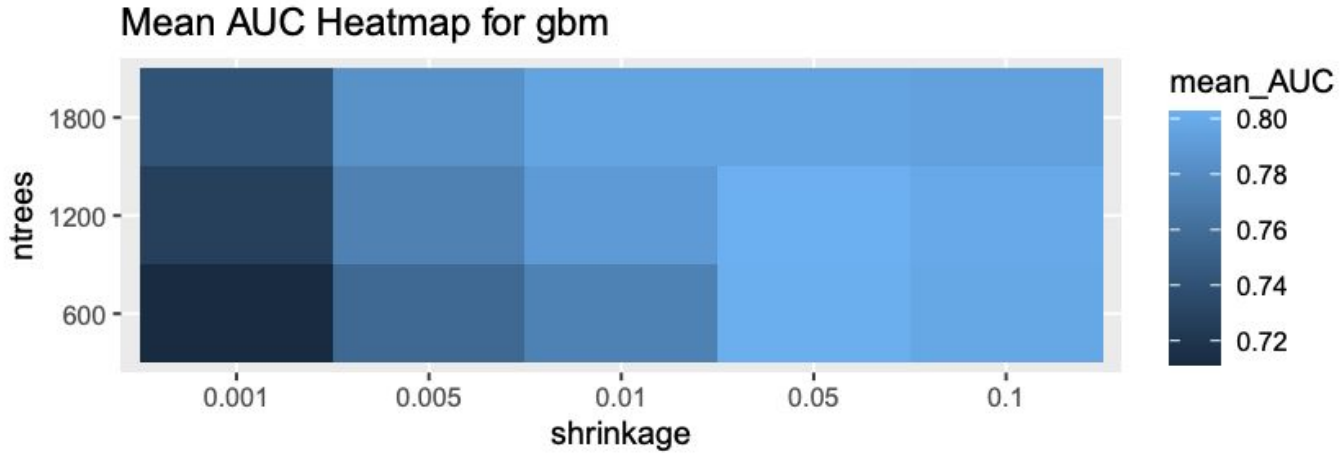
# Baseline Model - Gradient Boosting Machine

- Trained 15 models with cross-validation.
- Cross-validated ntrees and shrinkage:
  - ntrees: 600, 1200, and 1800
  - shrinkage: 0.001, 0.005, 0.010, 0.050, 0.100

subset of cross-validated gbm models ordered by mean AUC (15 total)

| shrinkage | n.trees | mean_error | sd_error | mean_AUC | sd_AUC |
|---|---|---|---|---|---|
| 0.05 | 1200 | 0.2861576 | 0.026049925 | 0.8022183 | 0.018723658 |
| 0.05 | 600 | 0.2764453 | 0.008045855 | 0.8015074 | 0.009653744 |
| 0.10 | 1200 | 0.3004216 | 0.031100034 | 0.7986702 | 0.024253701 |
| 0.10 | 600 | 0.2941406 | 0.026467746 | 0.7965060 | 0.021828465 |
| 0.05 | 1800 | 0.3031199 | 0.027850857 | 0.7949472 | 0.019940282 |


Mean Error Heatmap for gbm
ntrees
1800
1200
600
0.001
0.005
0.01
0.05
0.1
shrinkage
mean_error
0.34
0.32
0.30
0.28


Mean AUC Heatmap for gbm
ntrees
1800
1200
600
0.001
0.005
0.01
0.05
0.1
shrinkage
mean_AUC
0.80
0.78
0.76
0.74
0.72

# Baseline Model - Gradient Boosting Machine

- Trained 15 models with cross-validation.
- Cross-validated ntrees and shrinkage
  - ntrees: 600, 1200, and 1800
  - shrinkage: 0.001, 0.005, 0.010, 0.050, 0.100

subset of cross-validated gbm models ordered by mean AUC (15 total)

| shrinkage | n.trees | mean_error | sd_error | mean_AUC | sd_AUC |
|---|---|---|---|---|---|
| 0.05 | 1200 | 0.2861576 | 0.026049925 | 0.8022183 | 0.018723658 |
| 0.05 | 600 | 0.2764453 | 0.008045855 | 0.8015074 | 0.009653744 |
| 0.10 | 1200 | 0.3004216 | 0.031100034 | 0.7986702 | 0.024253701 |
| 0.10 | 600 | 0.2941406 | 0.026467746 | 0.7965060 | 0.021828465 |
| 0.05 | 1800 | 0.3031199 | 0.027850857 | 0.7949472 | 0.019940282 |

# Baseline Model - Gradient Boosting Machine

Accuracy: 70.48%
AUC: 0.79

Train time: 132.18 seconds
Test time: 13.38 seconds

# Advanced Model - Extreme Gradient Boosting (XGBoost)

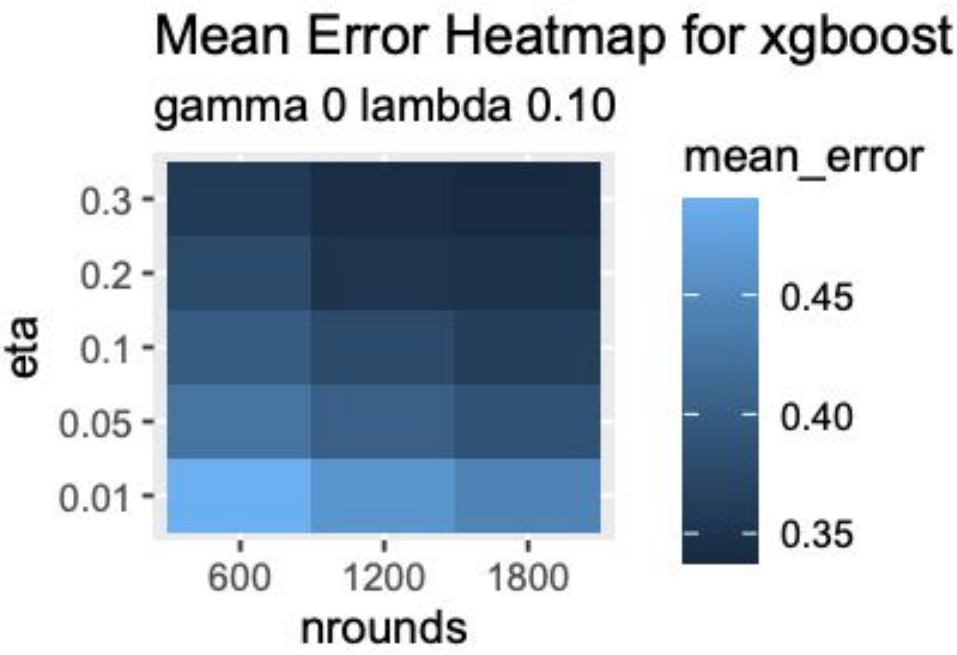
- Trained 150 models with cross-validation.
- Cross-validated nrounds, eta, gamma, and lambda:
  - nrounds: 600, 1200, 1800
  - eta: 0.01, 0.05, 0.1, 0.2, 0.3
  - gamma: 0, 5
  - lambda: .001, 0.005, 0.010, 0.050, 0.100

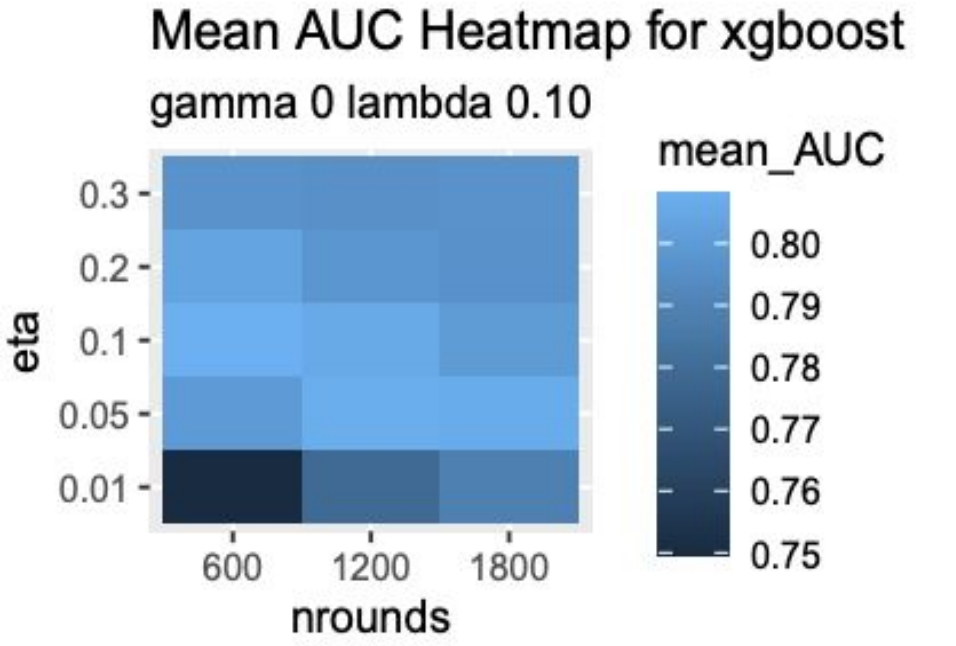
# Advanced Model - XGBoost

- Trained 150 models with cross-validation.
- Cross-validated nrounds, eta, gamma, and lambda:
  - nrounds: 600, 1200, 1800
  - eta: 0.01, 0.05, 0.1, 0.2, 0.3
  - gamma: 0, 5
  - lambda: .001, 0.005, 0.010, 0.050, 0.100

subset of cross-validated xgboost models ordered by mean AUC (150 total)

| eta | lambda | gamma | nrounds | mean_error | sd_error | mean_AUC | sd_AUC |
|---|---|---|---|---|---|---|---|
| 0.05 | 0.050 | 0 | 1200 | 0.4054330 | 0.01838352 | 0.8083090 | 0.01479400 |
| 0.10 | 0.100 | 0 | 600 | 0.3987534 | 0.01527476 | 0.8081563 | 0.01500276 |
| 0.10 | 0.001 | 0 | 600 | 0.3967334 | 0.01385476 | 0.8080448 | 0.01802835 |
| 0.10 | 0.050 | 0 | 600 | 0.4014677 | 0.01679825 | 0.8078401 | 0.01590874 |
| 0.05 | 0.001 | 0 | 1200 | 0.4048485 | 0.01376101 | 0.8074209 | 0.01509062 |


Mean Error Heatmap for xgboost
gamma 0 lambda 0.10
mean_error
0.45
0.40
0.35
eta
0.3
0.2
0.1
0.05
0.01
600
1200
1800
nrounds


Mean AUC Heatmap for xgboost
gamma 0 lambda 0.10
mean_AUC
0.80
0.79
0.78
0.77
0.76
0.75
eta
0.3
0.2
0.1
0.05
0.01
600
1200
1800
nrounds

# Advanced Model - XGBoost

- Trained 150 models with cross-validation.
- Cross-validated nrounds, eta, gamma, and lambda:
  - nrounds: 600, 1200, 1800
  - eta: 0.01, 0.05, 0.1, 0.2, 0.3
  - gamma: 0, 5
  - lambda: .001, 0.005, 0.010, 0.050, 0.100

subset of cross-validated xgboost models ordered by mean AUC (150 total)

| eta | lambda | gamma | nrounds | mean_error | sd_error | mean_AUC | sd_AUC |
|---|---|---|---|---|---|---|---|
| 0.05 | 0.050 | 0 | 1200 | 0.4054330 | 0.01838352 | 0.8083090 | 0.01479400 |
| 0.10 | 0.100 | 0 | 600 | 0.3987534 | 0.01527476 | 0.8081563 | 0.01500276 |
| 0.10 | 0.001 | 0 | 600 | 0.3967334 | 0.01385476 | 0.8080448 | 0.01802835 |
| 0.10 | 0.050 | 0 | 600 | 0.4014677 | 0.01679825 | 0.8078401 | 0.01590874 |
| 0.05 | 0.001 | 0 | 1200 | 0.4048485 | 0.01376101 | 0.8074209 | 0.01509062 |

# Advanced Model - XGBoost

Accuracy: 70.98%
AUC: 0.80

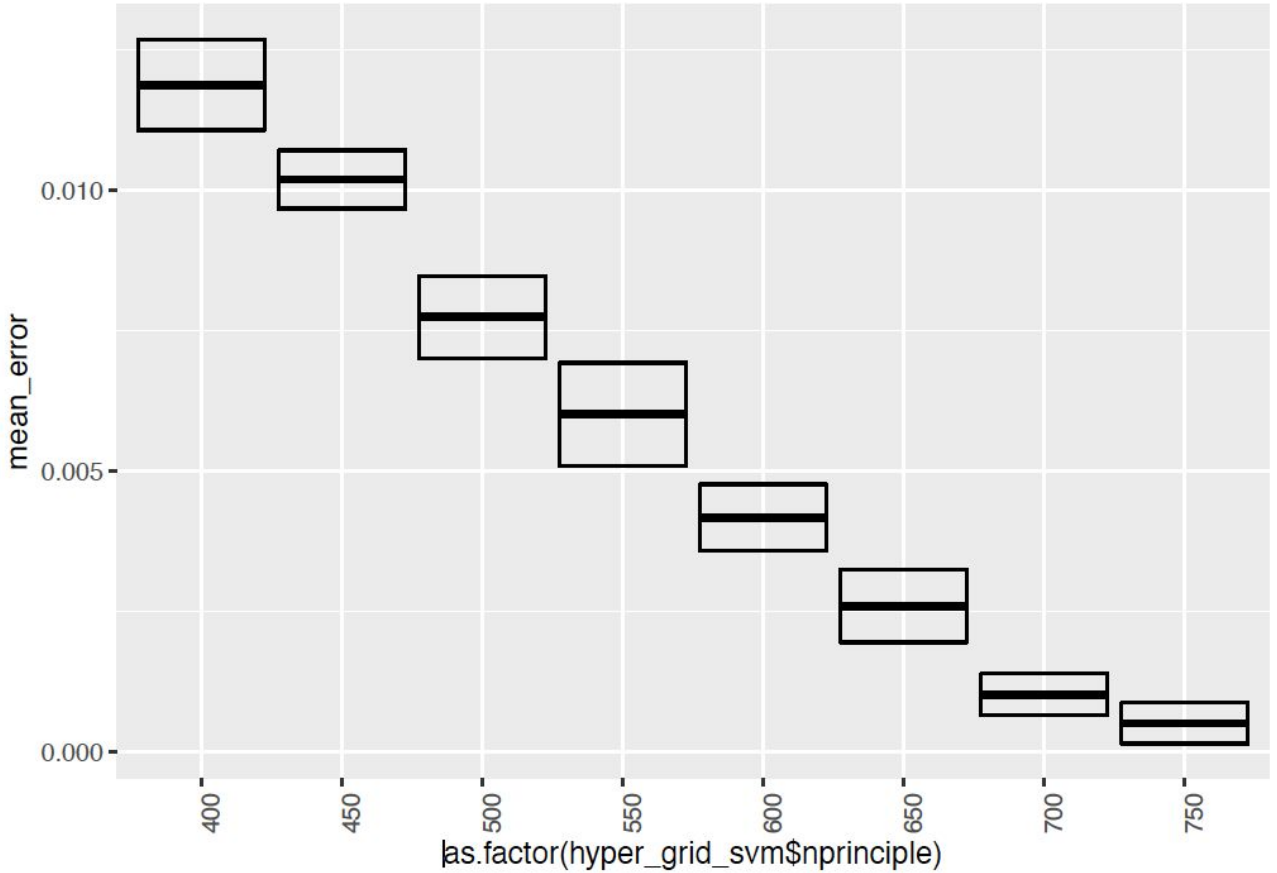
Train time: 74.03 seconds
Test time: 0.11 seconds

# Other Models

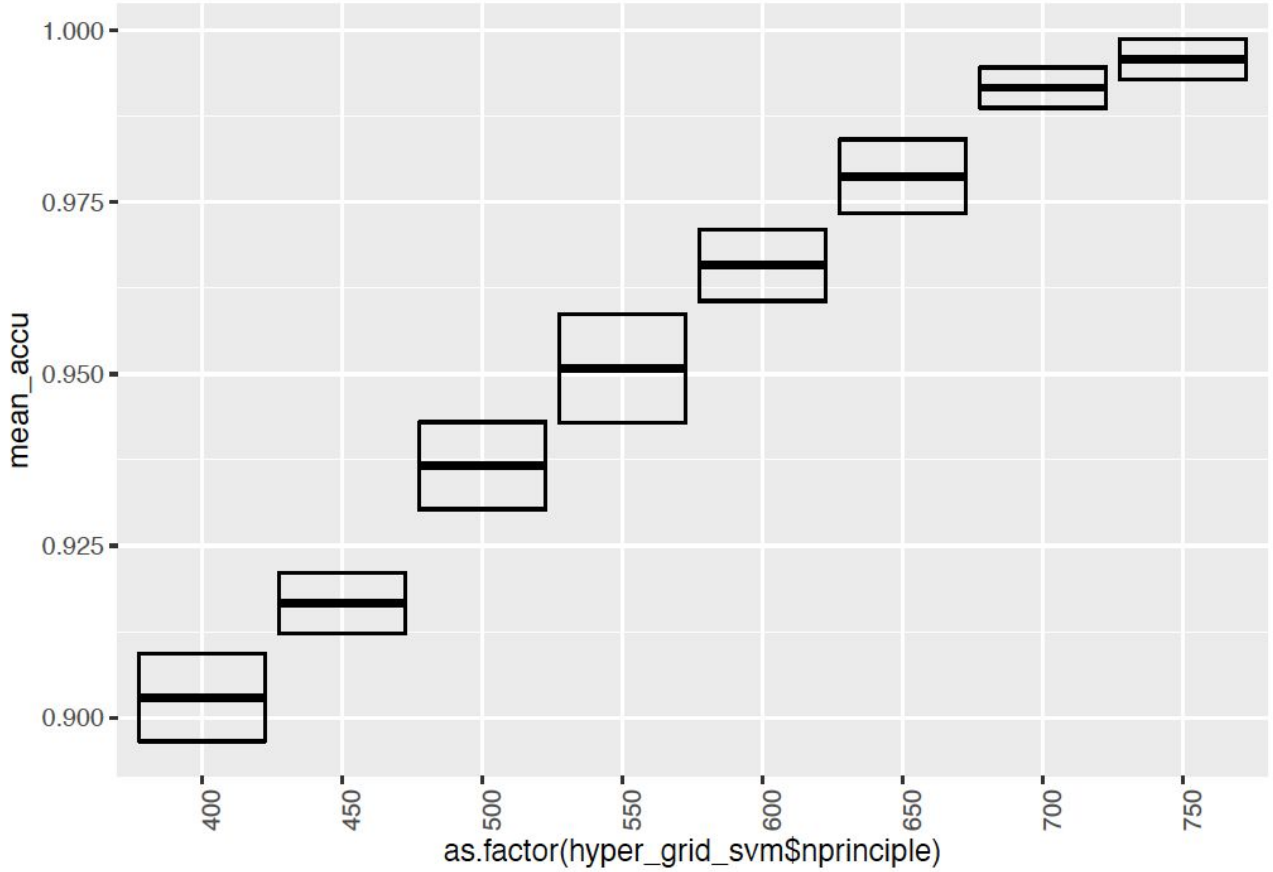
# Principal Components Analysis (PCA) + Support Vector Machine (SVM)

- Cross-validated principle component (pc):
  - pc: 400, 450, 500, 550, 600, 650, 700, 750

# UPD


mean_error
0.010
0.005
0.000
400
450
500
550
600
650
700
750
as.factor(hyper_grid_svm$nprinciple)

# UPDAT


1.000
0.975
0.950
0.925
0.900
mean_accu
400
450
500
550
600
650
700
750
as.factor(hyper_grid_svm$nprinciple)

# Principal Components Analysis (PCA) + Support Vector Machine (SVM)

Accuracy: 78.62%
AUC: 0.933

Train time: 13.285 seconds
Test time: 8.271 seconds

Reason why SVM is not our advanced model:

- The accuracy in the final test round was about 20% off from the cross validation session which is too much of a difference.
- Skeptical
- Less interpretable

# Random Forest

- Cross-validated ntrees and max depth:
  - ntrees: 100, 300, 500, 800, 1000
  - mtry: 100, 500

Accuracy: 81.66%
Unweighted AUC: 0.82

Train time: 531.55 seconds
Test time: 0.419 seconds

Reason why Random Forest is not our advanced model:

- Unweighted model training imbalanced data

# Random Forest with weights

- Cross-validated ntrees and max depth:
  - ntrees: 1500, 3000, 6000
  - max depth: 0, 5, 10, 15, 20, 25

Accuracy: 66.12%
AUC: 0.66

Train time: 9.486 seconds
Test time: 0.274 seconds

Reason why Random Forest with weights is not our advanced model:

- Accuracy and AUC not as good as XGBoost model

# Conclusion

- XGBoost is our proposed model
  - Slightly better accuracy and AUC than GBM (baseline model) and much faster running time

Thank you!