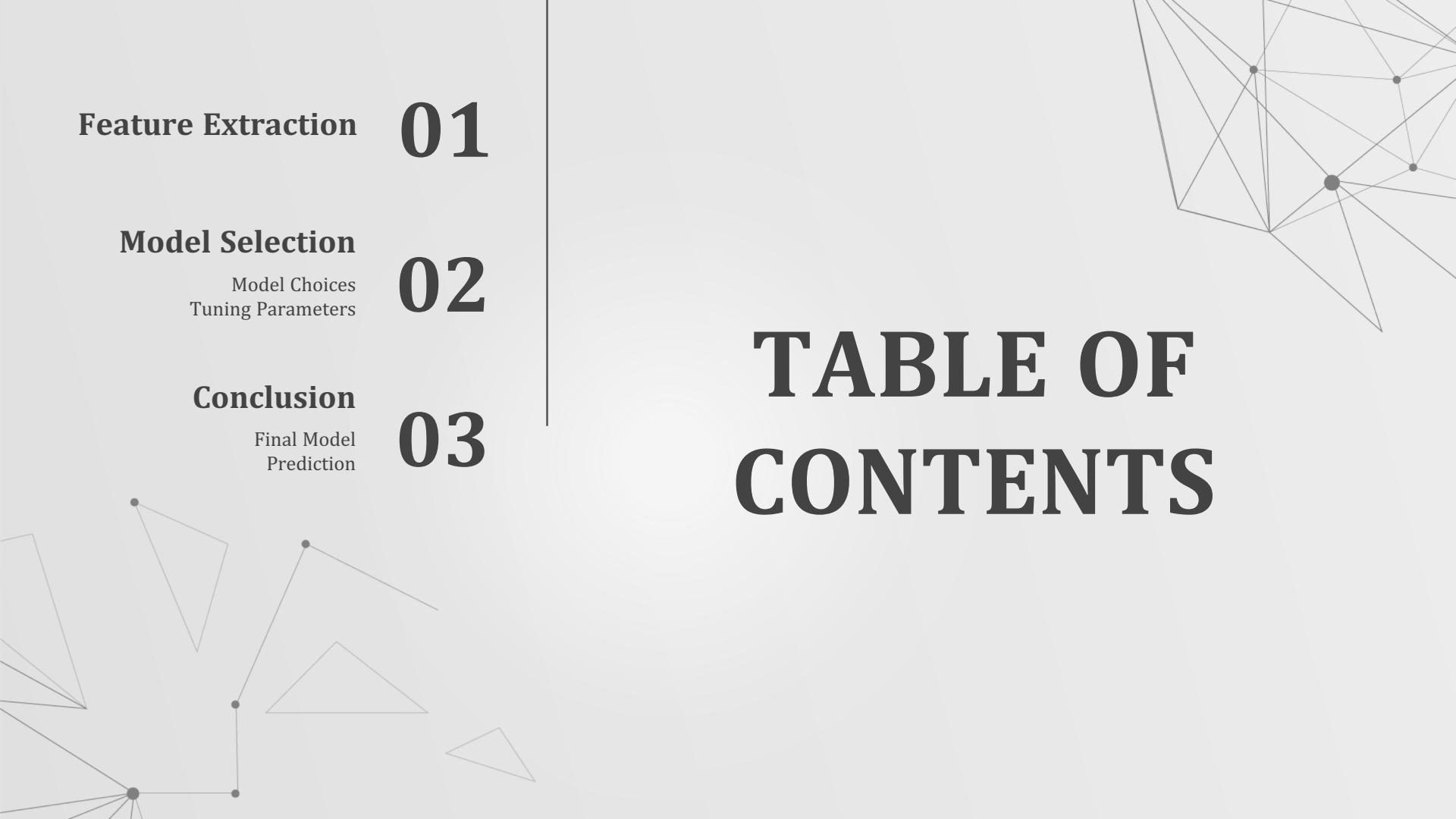
A faint background network graph consisting of numerous small, semi-transparent grey dots connected by thin grey lines. Interspersed among these dots are several larger, hollow grey triangles of varying sizes.

Applied Data Science

Project 3 : Predictive Modeling Facial Emotion Recognition

Group 6:
Syed Ahsan Bukhari
Xiwen Chen
Sung In Cho
Feng Qiu
Xuanhong Ye



Feature Extraction

01

Model Selection

Model Choices
Tuning Parameters

02

Conclusion

Final Model
Prediction

03

TABLE OF CONTENTS

01. Feature Extraction



- Used **pairwise distances** between all the fiducial points as a feature for facial emotion recognition.
- A model that used pairwise distances as a feature showed much higher prediction accuracy than a model that used just a location of fiducial points.

Image source: <https://link.springer.com/article/10.1186/1687-5281-2013-8>

02. Model Selection

Baseline
GBM

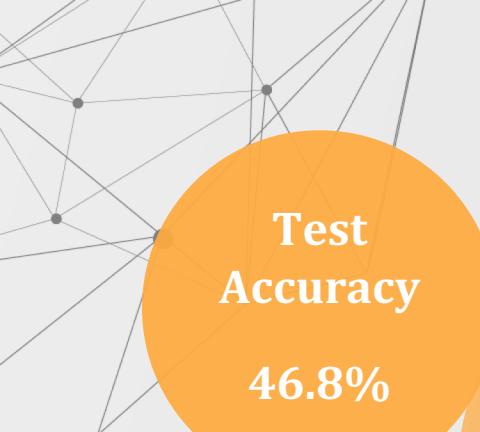
XGBOOST

SVM

Classification Models

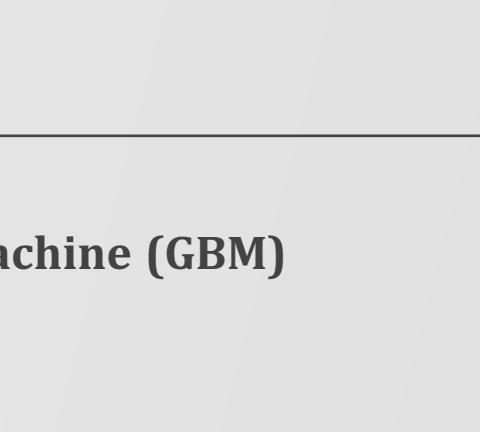
- Baseline model
 - GBM
- Advanced Models
 - XGBOOST
 - SVM

02. Model Selection



Test Accuracy

46.8%



Train Accuracy

96.5 %

Predicting Time

0.0312 s

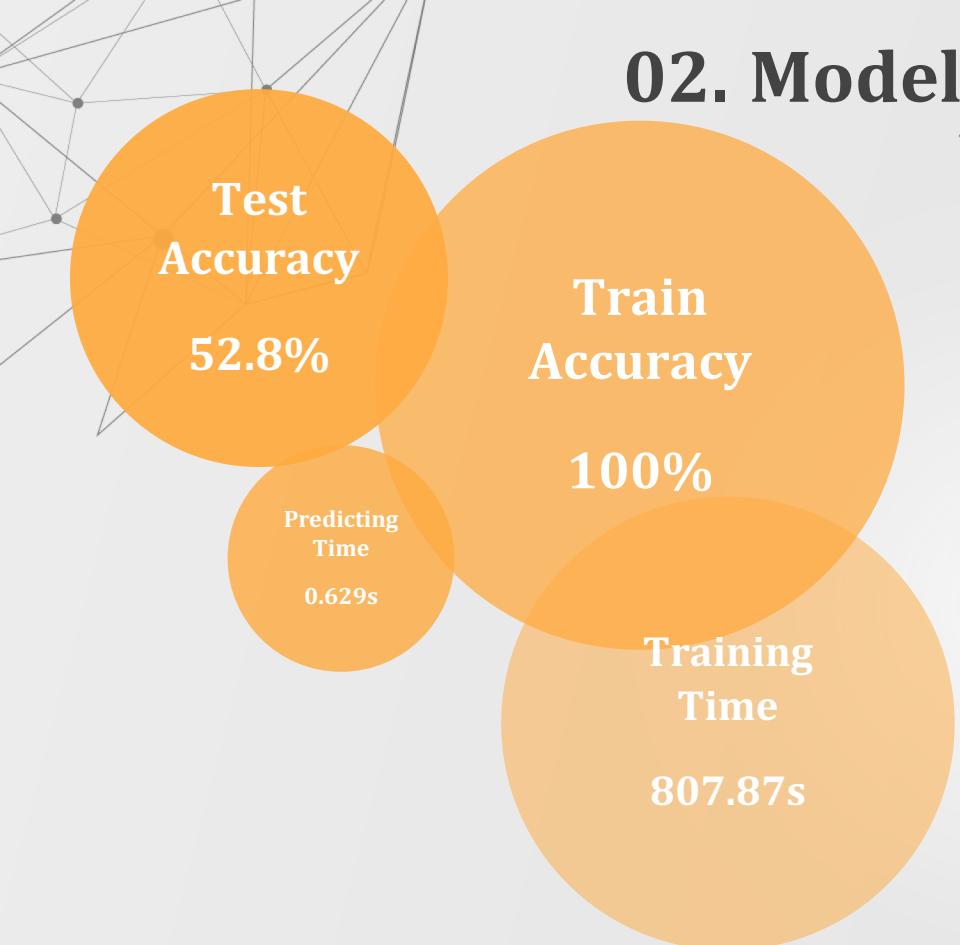
Training Time

1165.23 s

Gradient Boosting Machine (GBM)

- Baseline model
- Disadvantages
 - Can overemphasize outliers and cause overfitting
 - Often requires many trees (>1000) which can be time and memory exhaustive
 - Requires a large grid search during tuning

02. Model Selection

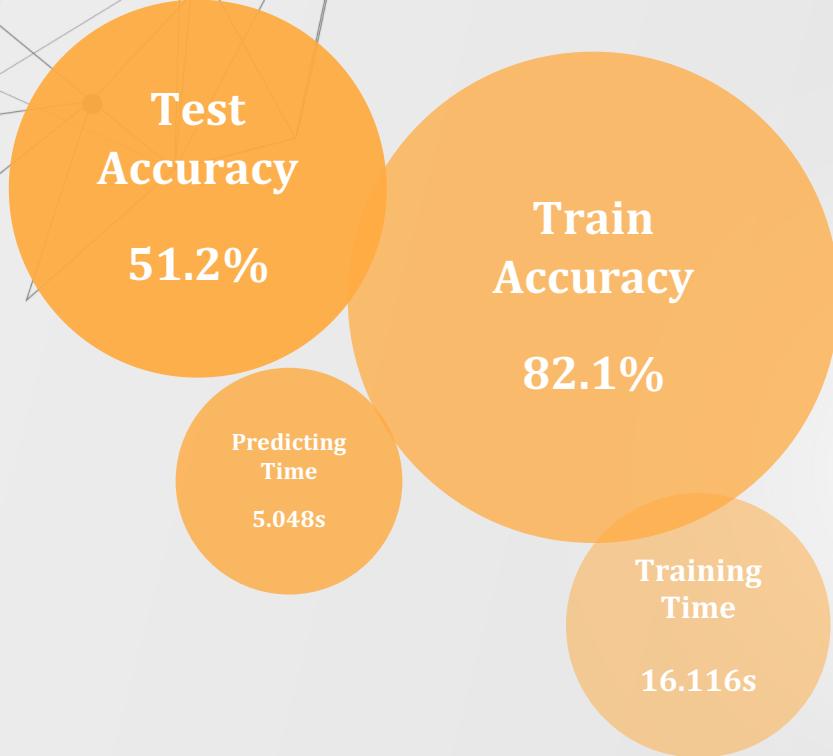


Model Choices

eXtreme Gradient Boosting (XGBOOST)

- XGBOOST vs GBM
 - Both follow Gradient Boosting
 - XGBOOST uses more regularized model formalization to control overfitting (better performance)
 - XGBOOST is more memory efficient
- Chosen Parameters
 - `n_estimators = 1000` (Number of Trees to fit)
 - `min_child_weight = 1`

02. Model Selection



Model Choices

Support Vector Machine (SVM)

- Why SVM?
 - Effective in high dimensional spaces
 - Uses a subset of training points (support vectors) in the decision function, so it is memory efficient.
- Chosen Parameters
 - Kernel : RBF
 - $c = 10$ (Penalty parameter)
 - $\text{gamma} = 0.0001$ (Kernel coefficient)

03. Conclusion

	Training Time	Test Accuracy
GBM : Baseline	1165.23s	46.8%
XGBOOST	807.87s	52.8%
SVM	15.9s	50.6%

Summary