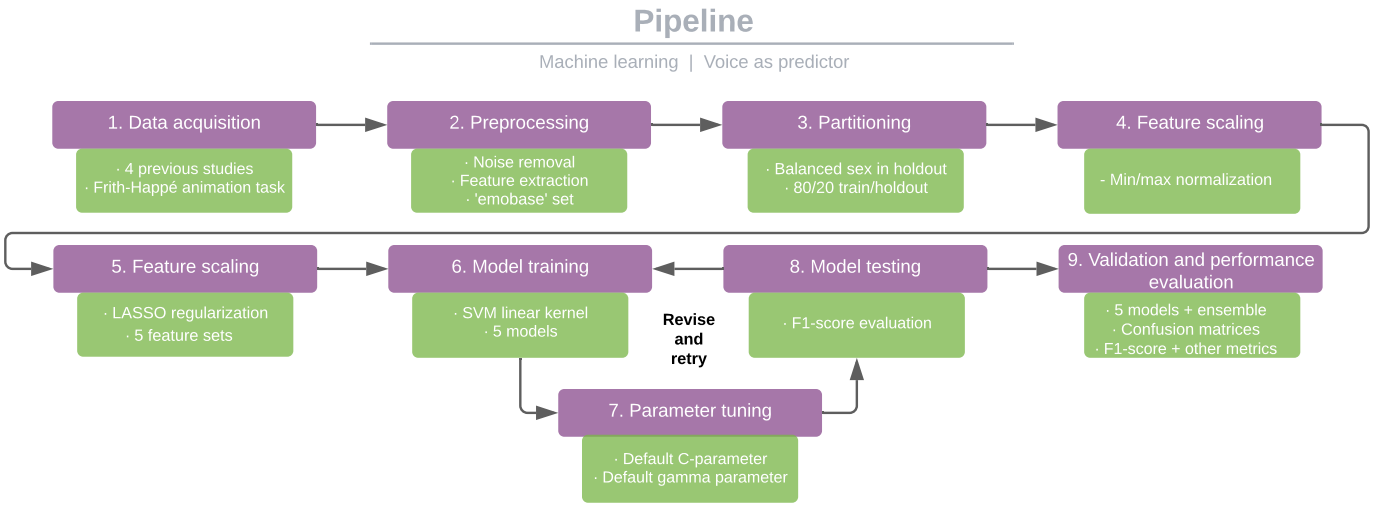
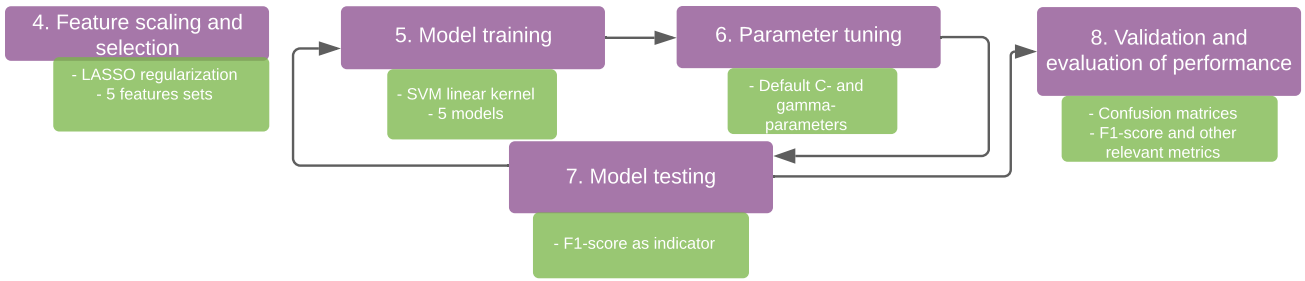
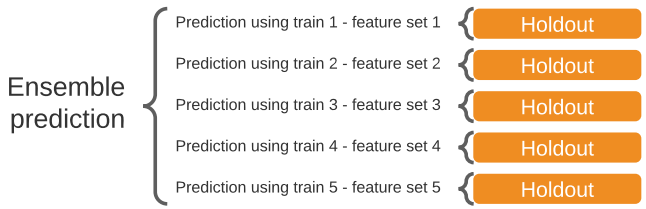

Pipeline
Machine learning | Voice as predictor
1. Data acquisition
· 4 previous studies
· Frith-Happé animation task
2. Preprocessing
· Noise removal
· Feature extraction
· 'emobase' set
3. Partitioning
· Balanced sex in holdout
· 80/20 train/holdout
4. Feature scaling
- Min/max normalization
5. Feature scaling
· LASSO regularization
· 5 feature sets
6. Model training
· SVM linear kernel
· 5 models
Revise and retry
8. Model testing
· F1-score evaluation
9. Validation and performance evaluation
· 5 models + ensemble
· Confusion matrices
· F1-score + other metrics
7. Parameter tuning
· Default C-parameter
· Default gamma parameter


4. Feature scaling and selection
- LASSO regularization
- 5 features sets
5. Model training
- SVM linear kernel
- 5 models
6. Parameter tuning
- Default C- and gamma-parameters
8. Validation and evaluation of performance
- Confusion matrices
- F1-score and other relevant metrics
7. Model testing
- F1-score as indicator


Ensemble prediction
Prediction using train 1 - feature set 1
Holdout
Prediction using train 2 - feature set 2
Holdout
Prediction using train 3 - feature set 3
Holdout
Prediction using train 4 - feature set 4
Holdout
Prediction using train 5 - feature set 5
Holdout