

**Project on “Machine Learning”
MSc in Artificial Intelligent**

Spilios Dellis

- Introduction
- Data presentation and problem determination
- Program workflow
 - Features extraction
 - Features preprocessing
 - Machine learning algorithms comparison
 - Features comparison
 - Best features and machine learning combination

Data description

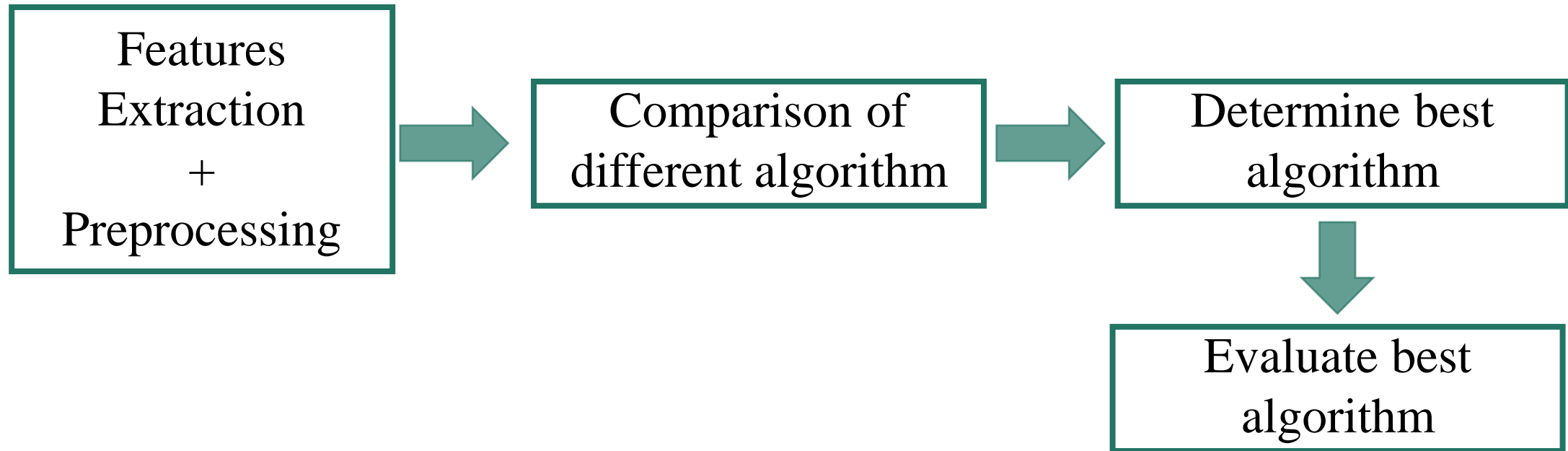
- Field audio recordings of 10 s duration from around the world (FreeSound project)
- Some contains bird sound
- Labeling by the Machine Listening Lab of the Queens Mary University of London
- More than 7000 audio recording (25 % of them contains bird sound)

Detect audio recording contains bird sound



Classification problem

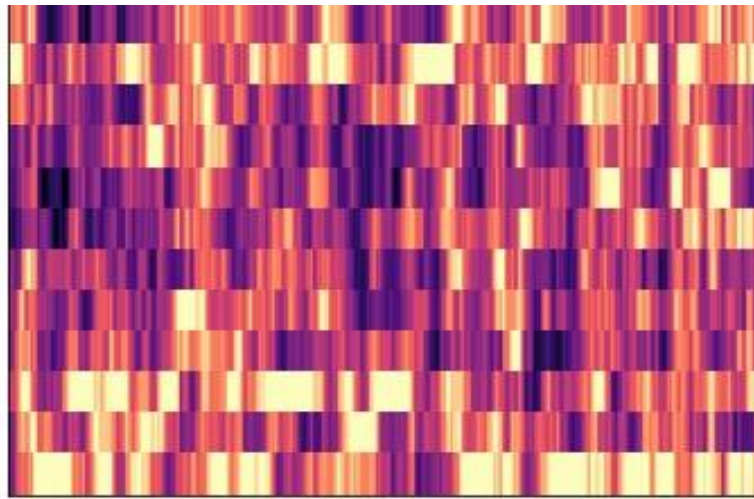
Workflow



Features. *Extraction*

Extraction

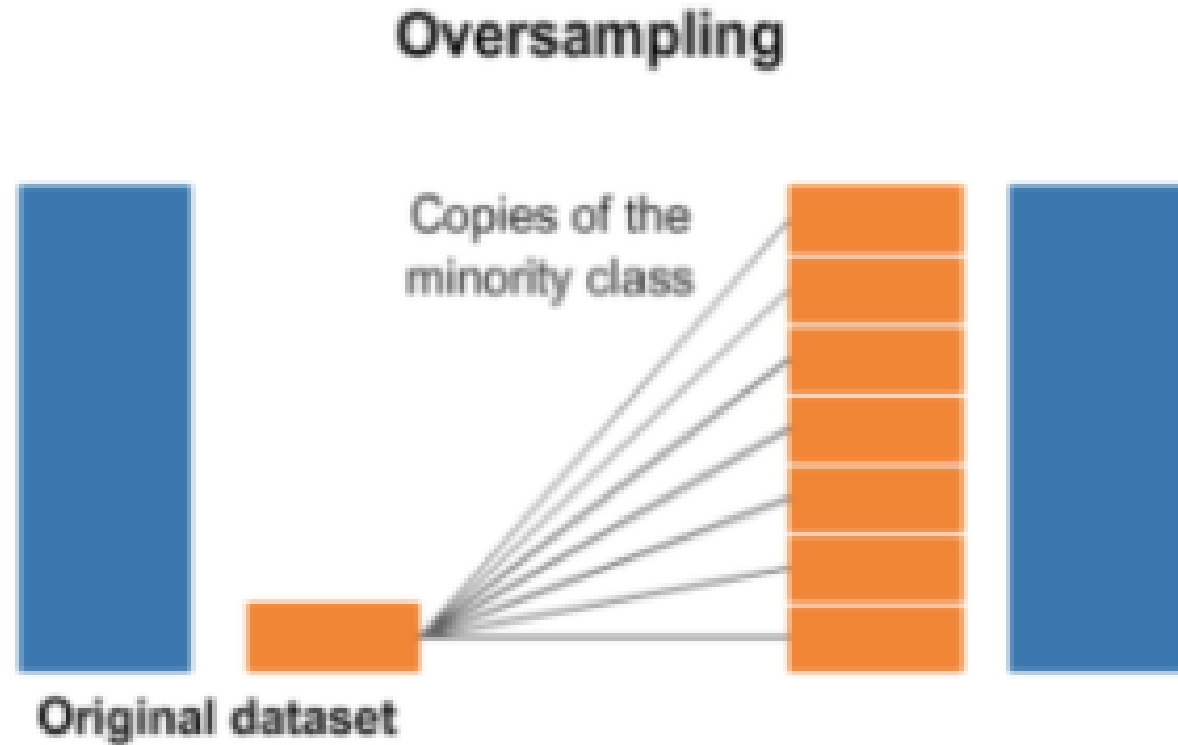
1. Prepare chromagrams for each audio recording using constant-Q method and waveform methods.
2. Use chromagrams as images and extract features using a CNN pretrained model (VGG16)



VGG16 → [1X4096]

Features. *Preprocessing*

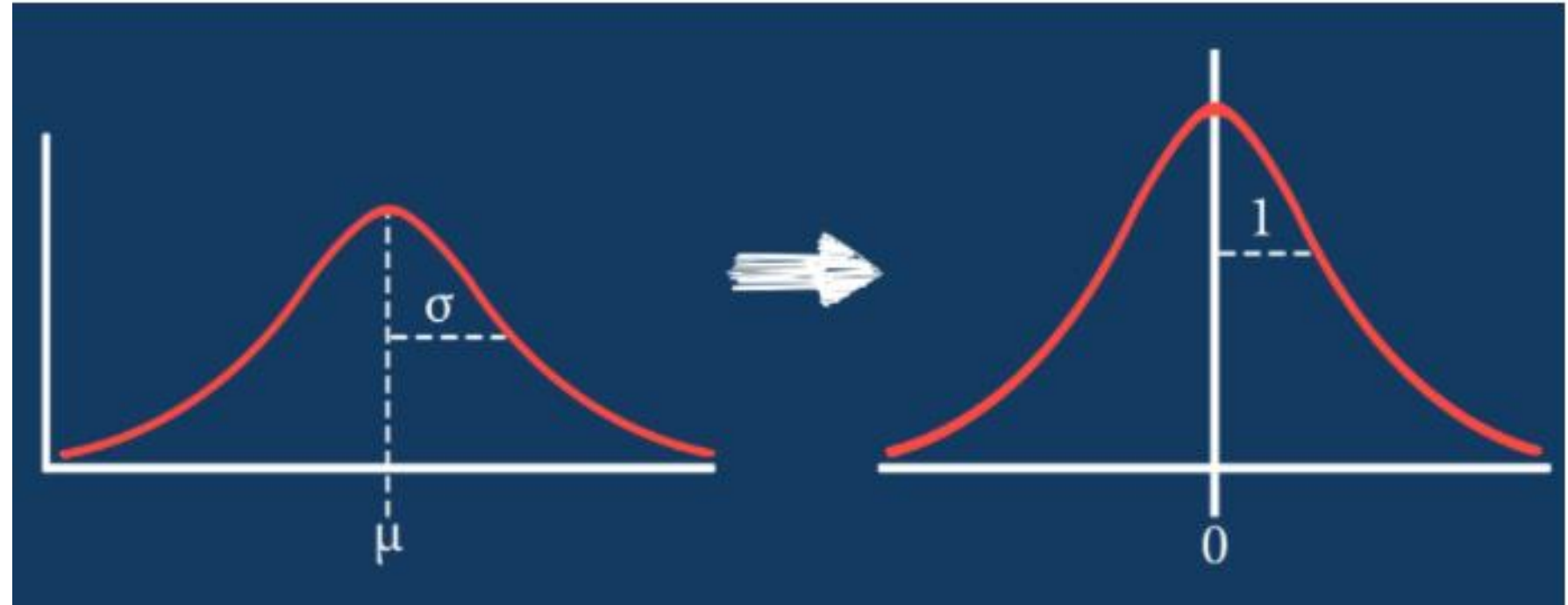
- **Resampling**



Features. *Preprocessing*

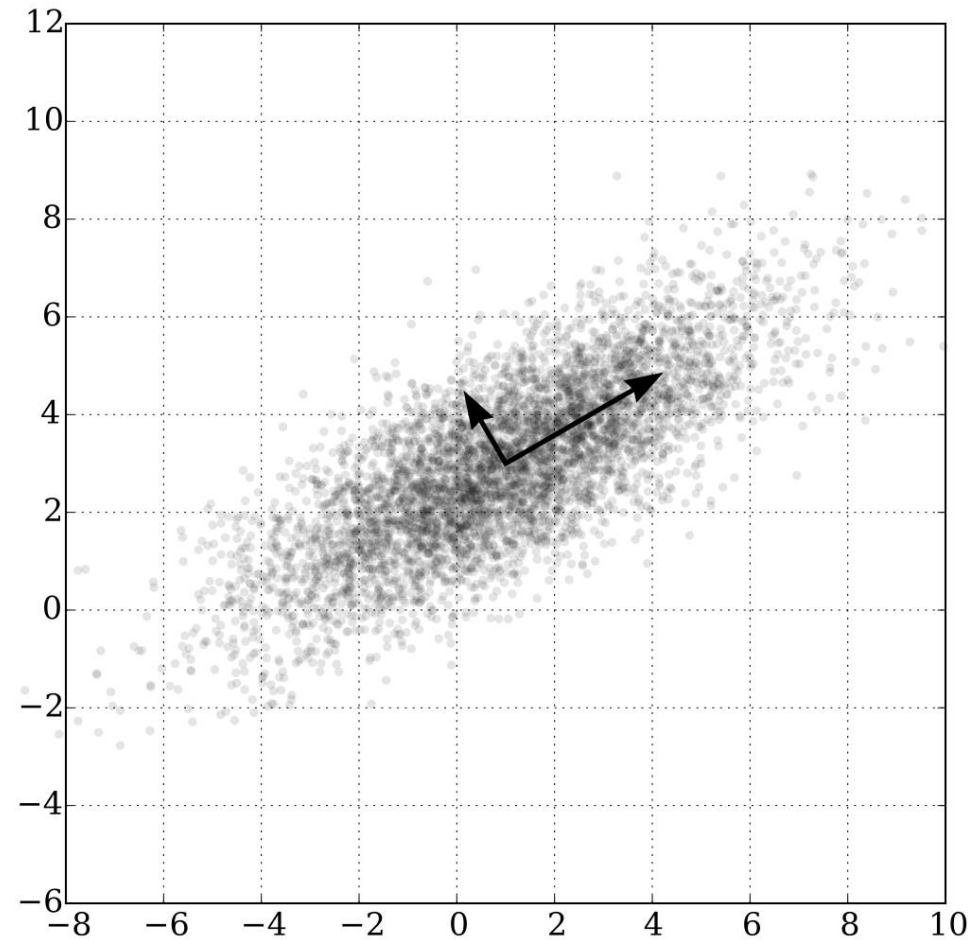
- **Resampling**

- **Standardization**



Features. *Preprocessing*

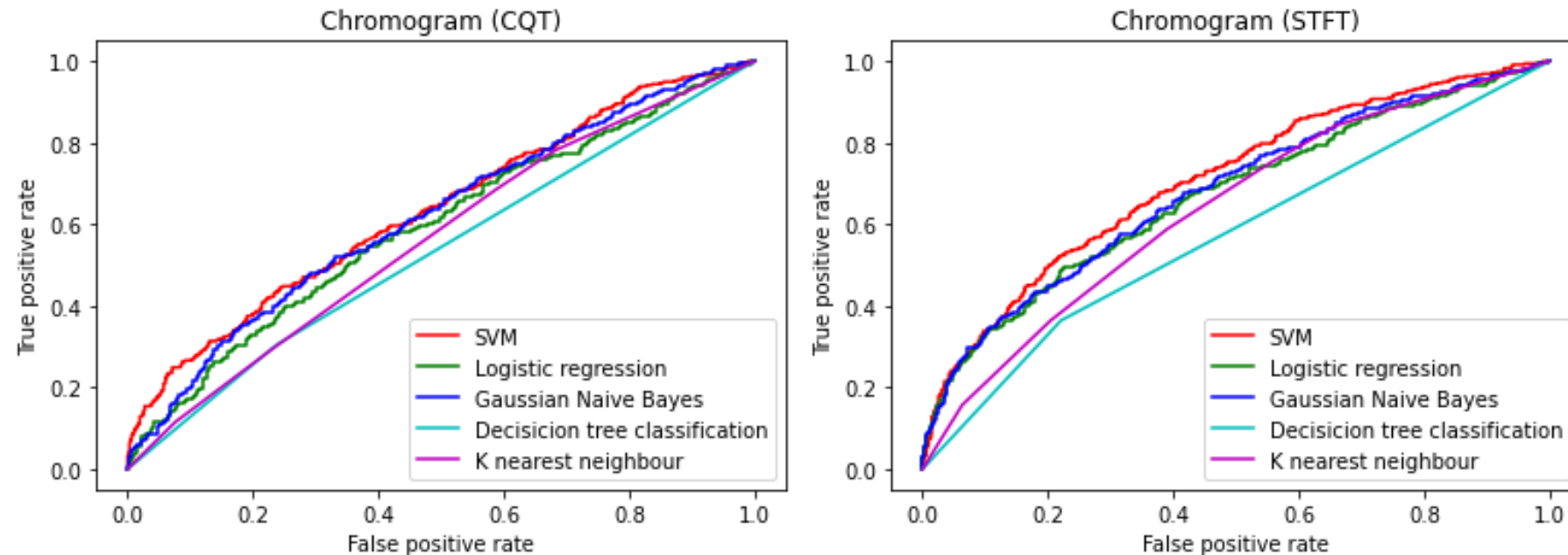
- **Resampling**
- **Standardization**
- **PCA**



ML algorithms comparison

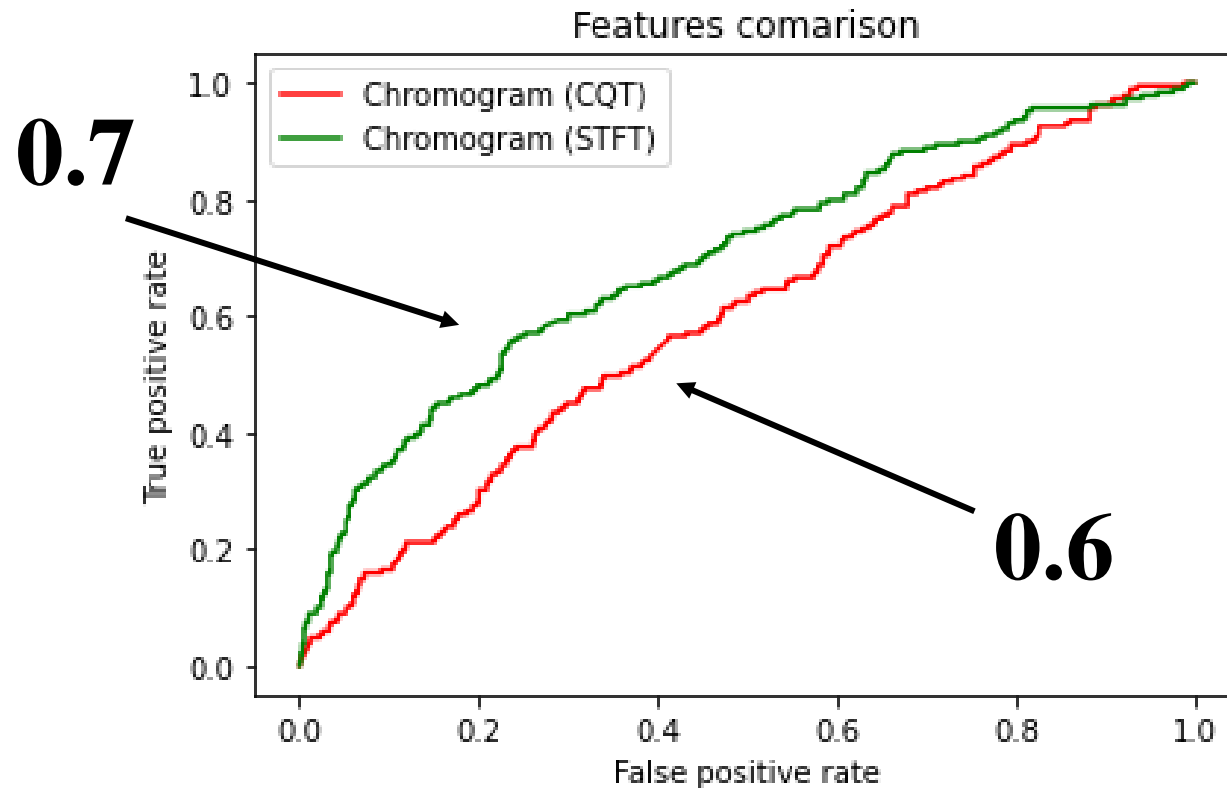
- 10 % of the original data kept as validation dataset.
- Rest used for model training with a 2/8 ratio between test and training datasets.
- The used algorithms are
 1. Support vector classification
 2. Logistic regression
 3. Gaussian naïve Bayes
 4. Decision tree classification
 5. K nearest neighbours
- ROC curves used for comparison
- Best algorithm derived for each type of features

Determine best algorithm in respect to features' type



	CQT	STFT
SVM	0.63	0.71
Logistic regression	0.59	0.67
Gaussian Naïve Bayes	0.61	0.68
Decision tree classification	0.53	0.57
K nearest neighbors	0.56	0.64

Comparison between features



Best combination

Features extracted from chromograms using waveform + SVM