DAY 5

Intelligent Audio Systems:
A review of the foundations and applications of semantic audio analysis and music information retrieval

Imagine Research

Center for Computer Research in Music and Acoustics

Jay LeBoeuf
Imagine Research
jay{at}imagine-research.com

Kyogu Lee
Gracenote
Kglee{at}ccrma.stanford.edu

June 2009

These lecture notes contain hyperlinks to the CCRMA Wiki.

On these pages, you can find supplemental material for lectures - providing extra tutorials, support, references for further reading, or demonstration code snippets for those interested in a given topic .

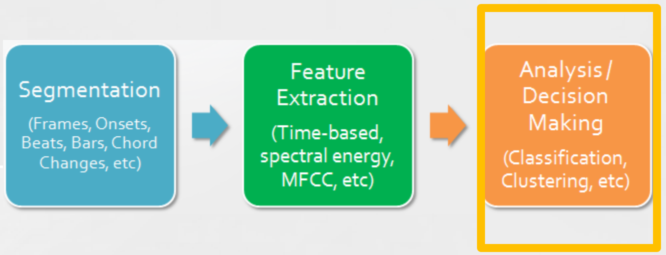
Click on the symbol on the lower-left corner of a slide to access additional resources.

WIKI REFERENCES…

Review from Day 4

- Your lab questions
- Explain “Novelty Score”
- What does the “# of components” in a gmm adjust?
- k-means is an algorithm that does “hard clustering.” What does that mean?

- Gautham
- Steve
- After lunch: listening room (Craig)



ANALYSIS AND DECISION MAKING: BUILDING AND EVALUATING A CLASSIFIER

Building a classifier

- Define classes (through training examples)
- Define features
- Define decision algorithm (parameters tuned through training data)
- Evaluate performance (error rate)

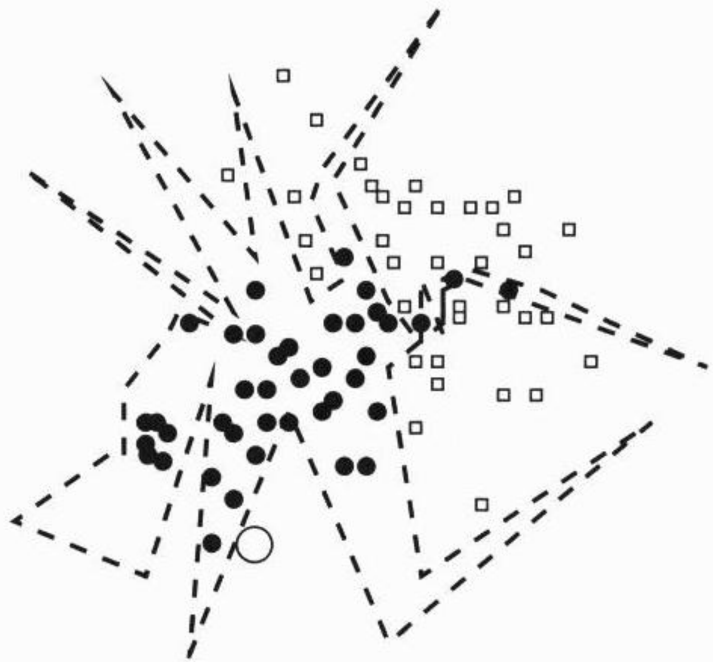


Fig. 2.13. Supervised classification into two classes with 2-dimensional data. In the training set (X, Y), data with label $\mathsf{y} = -1$ are represented with dots, whereas data with label $\mathsf{y} = 1$ are represented with squares. The dotted line is a classification function F such that $\mathsf{R}^{\text{emp}}_{(\mathsf{X},\mathsf{Y})}[\mathsf{F}] = 0$. Though it achieves zero empirical risk, F is not a good classification function, as it makes an error for a new datum which is not in the training set (circle at the bottom, with the true label $\mathsf{y} = -1$).

DATA PREPARATION

Gathering training data

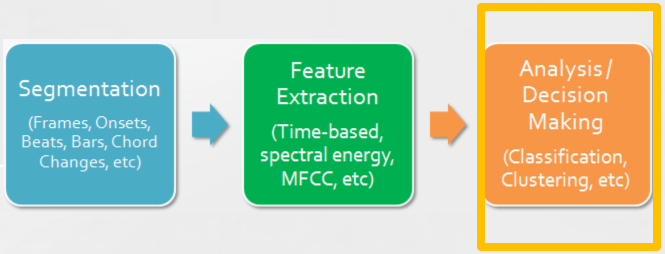
- Hand-annotate
- One technique is to use existing ground-truth MIDI files to synthesize audio.
- Contact researchers regarding their annotations (e.g., Beatles collection)
- Creative Commons, research collections (e.g., RWC database, OLPC data)

Data preparation

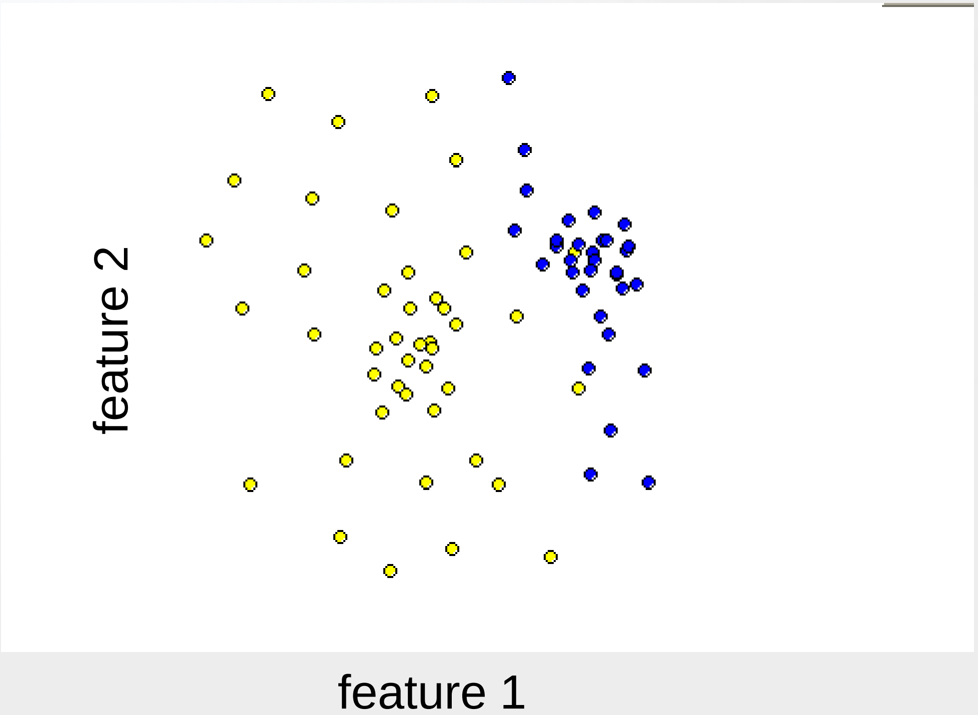
- Examine your data at every chance. (means, max, min, std, “NaN”, “Infs”)
- Try to visualize data when possible to see patterns and see if it makes. Incredible sanity check.
- Eliminate noisy data
- Data preparation
  - Cleaning
    - Open up and examine
    - Handle missing values
  - Relevance / Feature analysis
    - Remove irrelevant or redundant attributes
  - Data Transformation
    - Generalize or normalize data

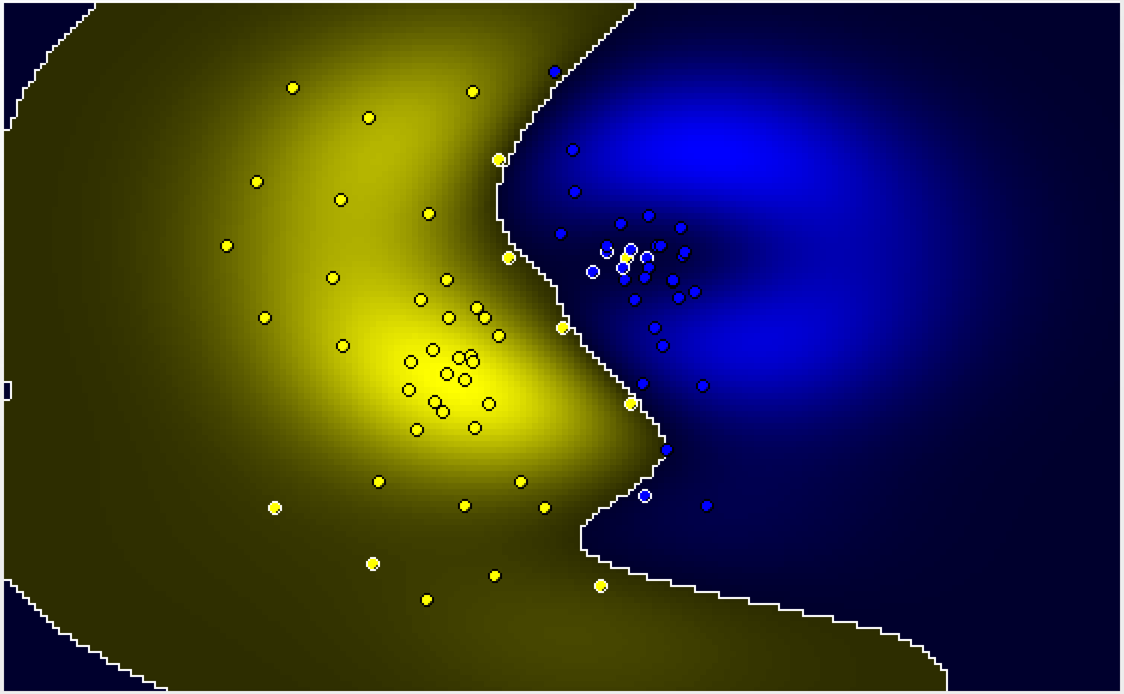
Classification

- Classification – class labels (discrete or nominal)
- Regression – models continuous-valued function



ANALYSIS AND DECISION MAKING: SUPPORT VECTOR MACHINES (SVM)

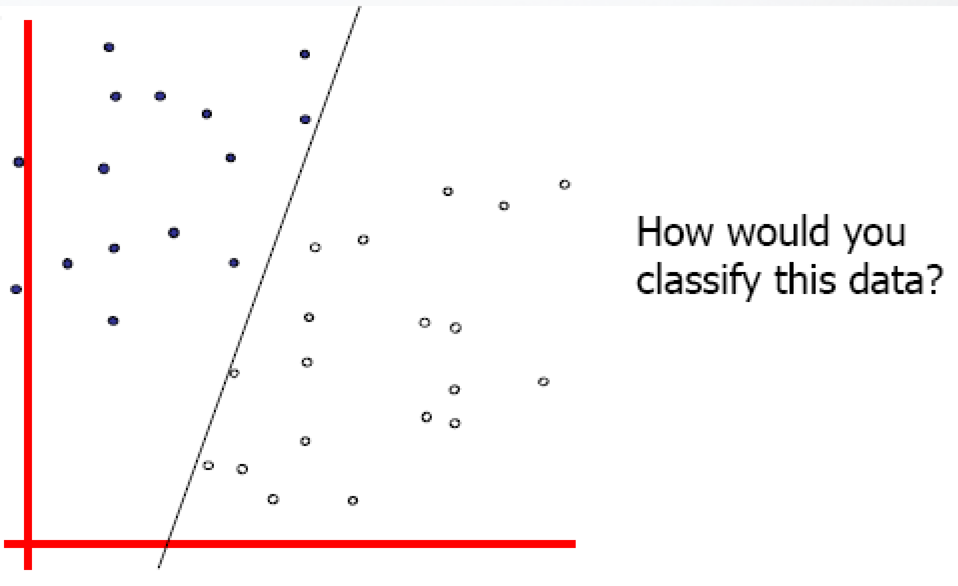

feature 2
feature 1

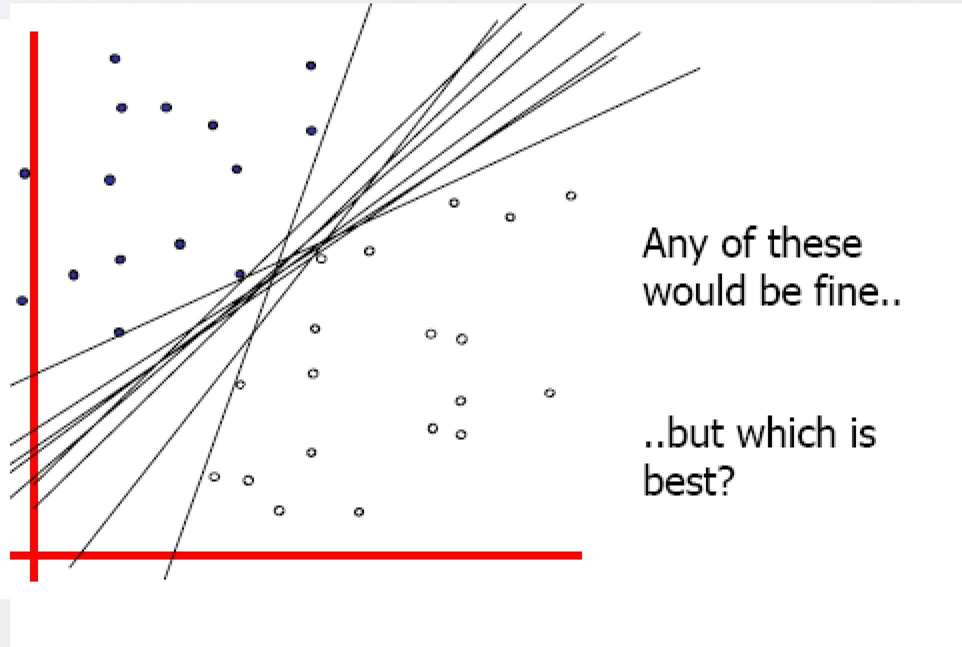


SVM

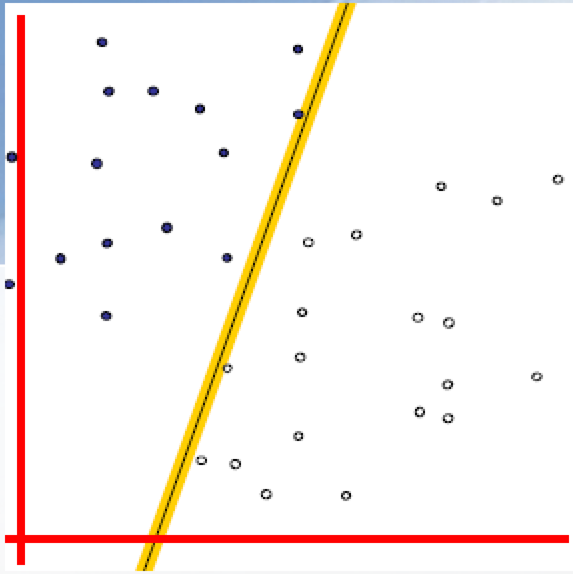
- Hyperplane separates the data from the two classes with a “maximum margin”.
- Support Vectors - are those data points that the margin pushes up against
- SVM training is guaranteed to find the global minimum of the cost function.
- Less experience needed - fewer parameters to tune

- >> svmdemo

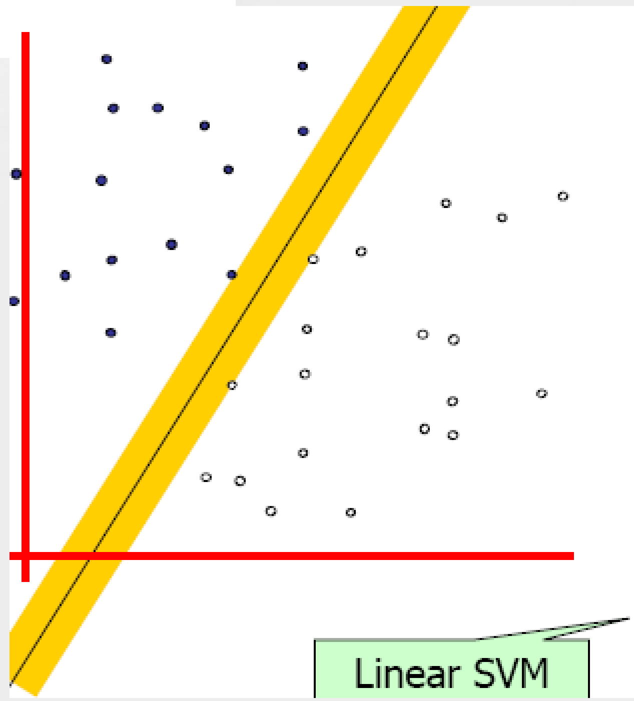




From : http://www.autonlab.org/tutorials/svm15.pdf



Define the margin of a linear classifier as the width that the boundary could be increased by before hitting a datapoint.

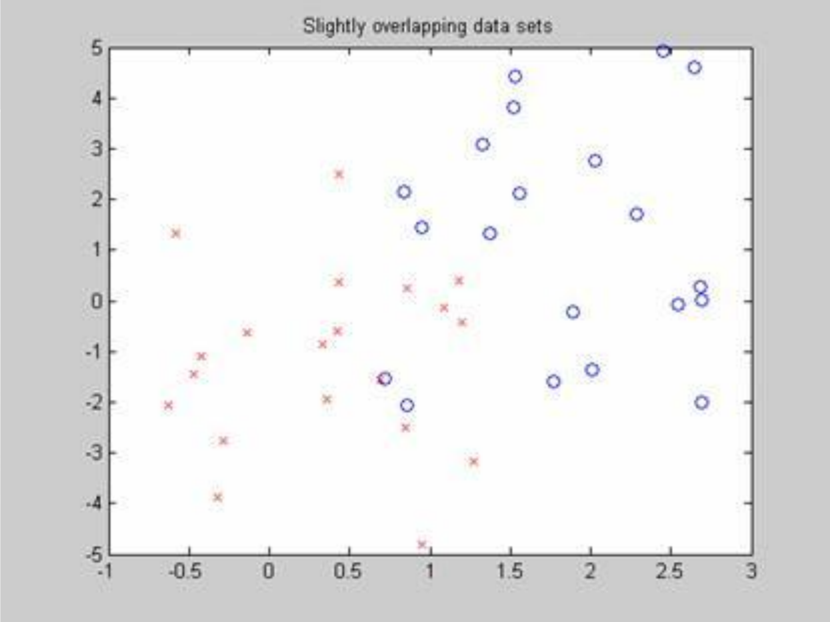


The maximum margin linear classifier is the linear classifier with the, um, maximum margin.

This is the simplest kind of SVM (Called an LSVM)

From : http://www.autonlab.org/tutorials/svm15.pdf

SVM Parameters



What effect do the parameters of an radial-basis-function SVM have on the separating the two data sets?

Using the RBF kernel, we have to choose values of :

gamma = degree of curviness of the hyperplane / complexity of the contour

C = allowance for points to overlap into each other's class

Video 1

Video 2

RBF Parameters: C and gamma

- Grid search using cross-validation to find the best one. Coarse then fine grid search.
- e.g., 2-5, 2-3, … 2+15, gamma = 2-15, 2-13, 2+3
- Why grid search
  - Psychological (If you *have time* for brute force… why chance it on approximations or heuristics)
  - Since there are only 2 params, grid search isn't all the different from advanced estimation techniques
  - Easily parallelized (C and gamma are independent)
- Large datasets
  - You can take random sample as approximation

Practical Guide to SVM: The Lab

- Feature selection?
- Scale feature data
  - Save scaling stats so we can scale the test data to be in the same range
- Feature format
- Class labels {1,-1} or {0,1}
- Kernels (linear, polynomial, RBF, sigmoid)
- Find best C and gamma (cross-validation)
- Train with entire training set
- Test with validation or test set

- easy.py or grid.py

Evaluation Measures

True+	correct	Classifier correctly predicted something in it's list of known positives
False-	absent	Classifier did not hit, for a known positive result.
False+	incorrect	Classifier said that something was positive when it's actually negative

Evaluation Measures

"Accuracy"

↑ is good

Precision - "Positive Predictive Value"

↓ = high F+ rate, the classifier is hitting all the time

↑ = low F+ rate, no extraneous hits

Recall – "Missed Hits"

↓ = high F- rate, the classifier is missing good hits

↑ = low F- rate, great at negative discrimination – always returns a negative when it should

F-Measure – a blend of precision and recall (harmonic-weighted mean)

↑

Evaluate Measures

P = T+ / (T+ + F+)	[0…1]
R = T+ / (T+ + F -)	[0…1]
F = 2*P*R/(P+R)	[0…1]

Training and test data

- An overfit model matches every training example (Now it's "overtrained.")
- Training Error AKA "Class Loss"
- Generalization
  - The goal is to classify new, unseen data.
  - The goal is NOT to fit the training data perfectly.
- An overfit model will not be well-generalized, and ***will*** make errors.
- Rule of thumb: favor simple solutions and more "general" solutions.

Training and test data

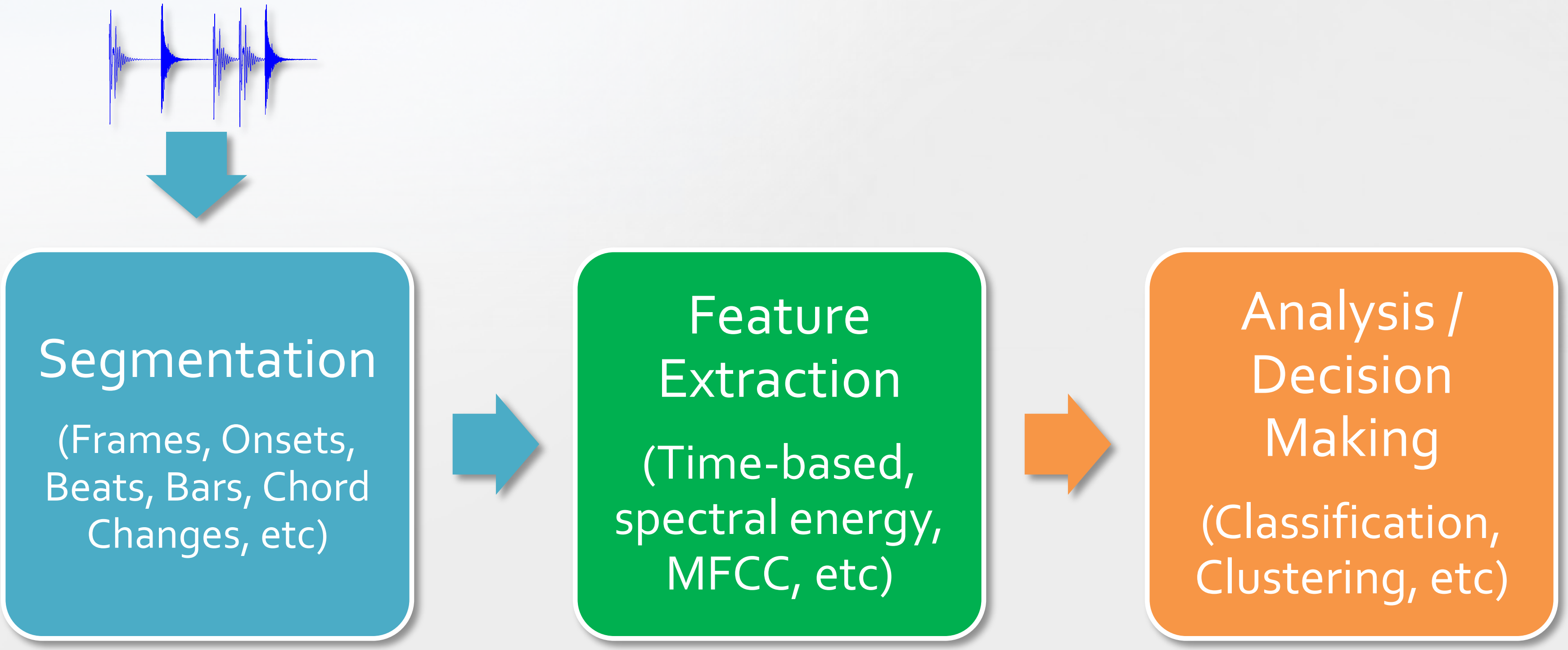
- Cross-validation
- Training, Validation, and Test set
  - Partition randomly to ensure that relative proportion of files in each category was preserved for each set
    - Weka or Netlab has sampling code
- Warnings:
  - Don't test (or optimize, at least) with training data
  - Don't train on test data (no!)

WORKSHOP SUMMARY

SUMMARY

- Building blocks can be rearranged to create various MIR applications.

Basic system overview



SUMMARY

- Building blocks can be rearranged to create various MIR applications.
- Research can examine:
  - "Inventing" new features
  - Examining techniques, parameters, and tuning in certain building blocks to enhance performance
  - Defining similarity measures
  - Selecting "labels" for data; building test / train datasets

Additional information…

ISMIR

October, 2009
Kobe Japan

AES 2009

October 2009
New York City

MUSIC-IR list

Websites

- https://cm-wiki.stanford.edu/wiki/MIR_workshop_2009
- www.iua.upf.es/mtg
- Queen Mary C4DM
- LabROSA
- http://iua.upf.edu/smc/
- www.bmat.com
- www.imagine-research.com
- www.Gracenote.com

Seminal Papers

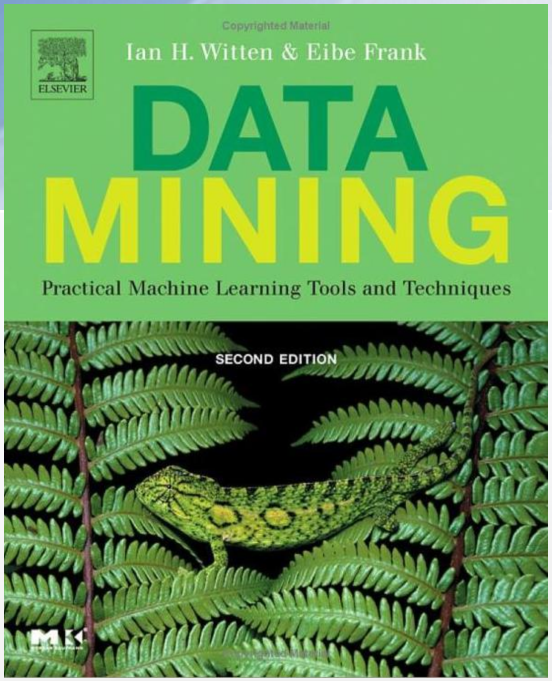
- References listed in Klapuri & Davy book
- Proceedings of ISMIR (International Symposium on Music Information Retrieval)
- *Papers listed on Pg 136-7 of MIR Toolbox: http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/userguide1.1*
- Audio Engineering Society
- IEEE ICASSP (International Conference on Acoustics, Speech, and Signal Processing)

Please add your ***favorites*** to the resources page at:

- http://cm-wiki.stanford.edu/wiki/MIR_workshop_2008_notes

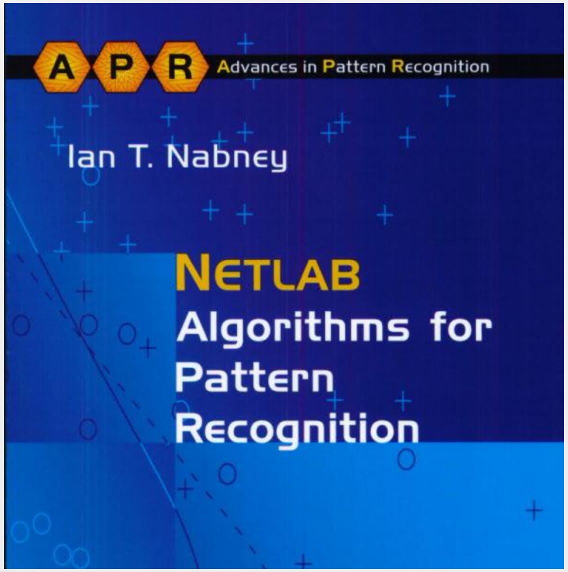
Acknowledgments

Dan Ellis
Juan Bello
Ge Wang
Perry Cook
Olivier Lartillot (MIR Toolbox)
Perfe Herrera
Mark Sandler
Òscar Celma
Carr Wilkerson
Sasha, Tricia, Chris Chafe and the whole CCRMA crew!



Data Mining: Practical Machine Learning Tools and Techniques, Second Edition
by Ian H. Witten , Eibe Frank

Weka Toolbox (Java / Standalone)
http://www.cs.waikato.ac.nz/ml/weka/



Netlab
By Ian T. Nabney (Author)

Netlab (Matlab) Software

Thank you!

Jay LeBoeuf
jay   imagine-research.com

Kyogu Lee
kglee@ccrma.stanford.edu