Week 14

Information Fusion

What is Multimedia?

Exploiting Multiple sources of information

Representation   Compression   Storage   Integration

S1

S2

1 2 4 2 6
5 4 2 6 7
8 4 2 7 9
2 6 2 5 9

S3

HTML

DB

Output

Analysis
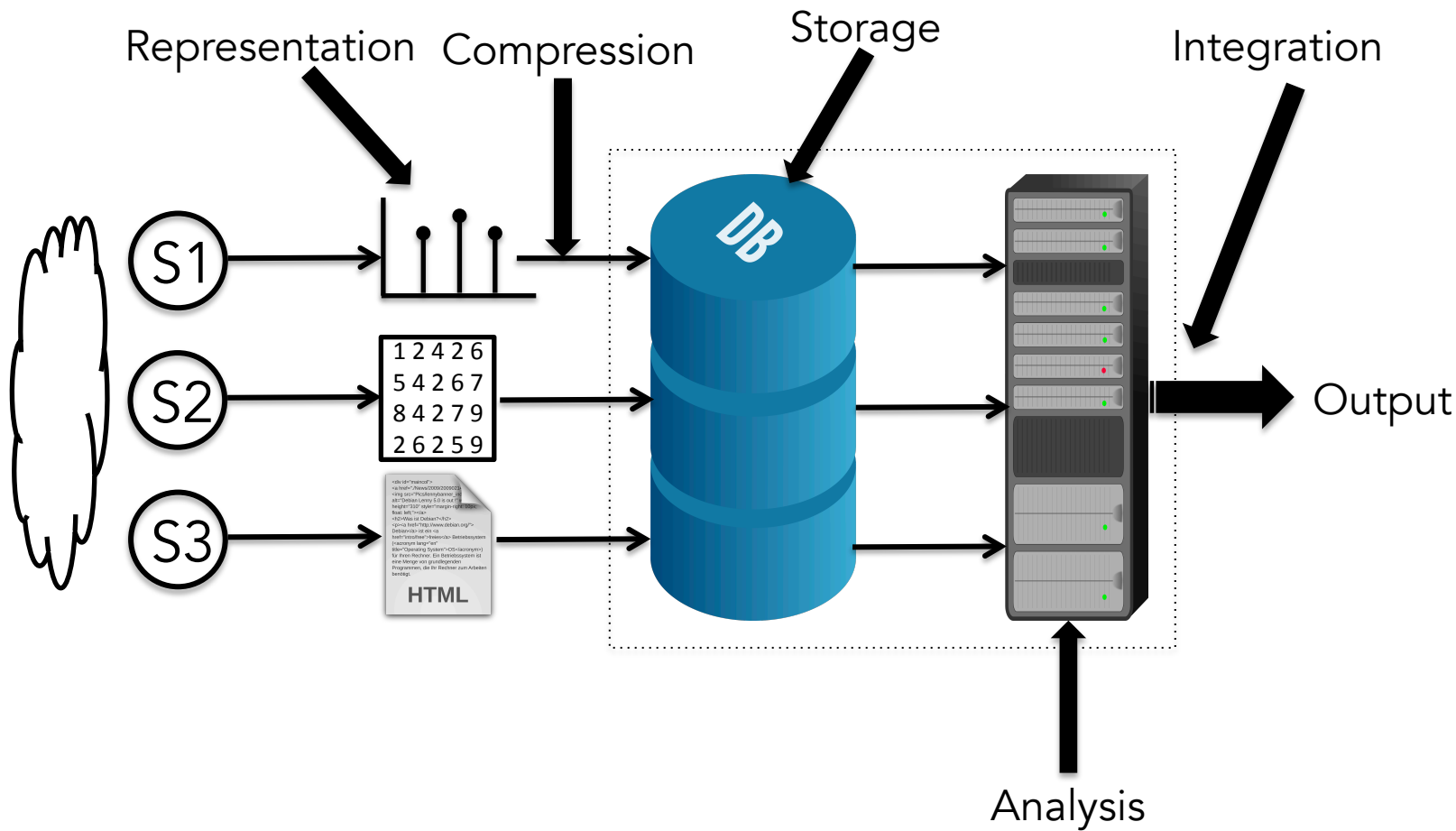
Why multiple sources?

Physical Perspective

- Have multiple instances of the same type of source
    - Multiple cameras
- Have different types of sources
    - Audio, video, and text
- Both

Information Perspective

- Redundant source: providing correlated information

  – Two cameras with overlapping views

- Complementary source: capturing different information

  – Two cameras with orthogonal views

Multimedia to improve robustness

- Multiple sources providing redundant information can also increase the robustness of the system.
- If one sensor is weak in some scenarios, other sensor may be employed which is strong in those scenarios.
- Audio-visual surveillance.
  - During day it is noisy, so video is better than audio
  - In night it is dark but relatively silent, so we can rely more on audio

Multimedia to improve reliability

- Keep duplicate sources/sensors in different physical settings

- If one source stops giving information, other source is still working

- E.g. Two camera surveillance in ATM to detect vandalism

Multimedia to improve efficiency

- Multiple sources can be processed in parallel
- Quick sources (processing efficient) can be processed first to generate early warning
- A slow source can be replaced by an efficient source completely
- Motion/heat sensor and video for person detection in a battle-field

Multimedia to make it cost effective

- Use cheaper sources (in terms of money) to implement intermediate steps to reduce number of expensive sources needed.

- Use heat/motion sensor with one PTZ camera to monitor a large area

Multimedia to make it more accurate

- The fusion of redundant information from different sources can reduce overall uncertainty.
- Complementary information provided by different sources results in an information gain in comparison to a single source.
- E.g. audio and gesture-based anger detection

Multimedia novel applications

CSL607 Projects

Why multiple sources?
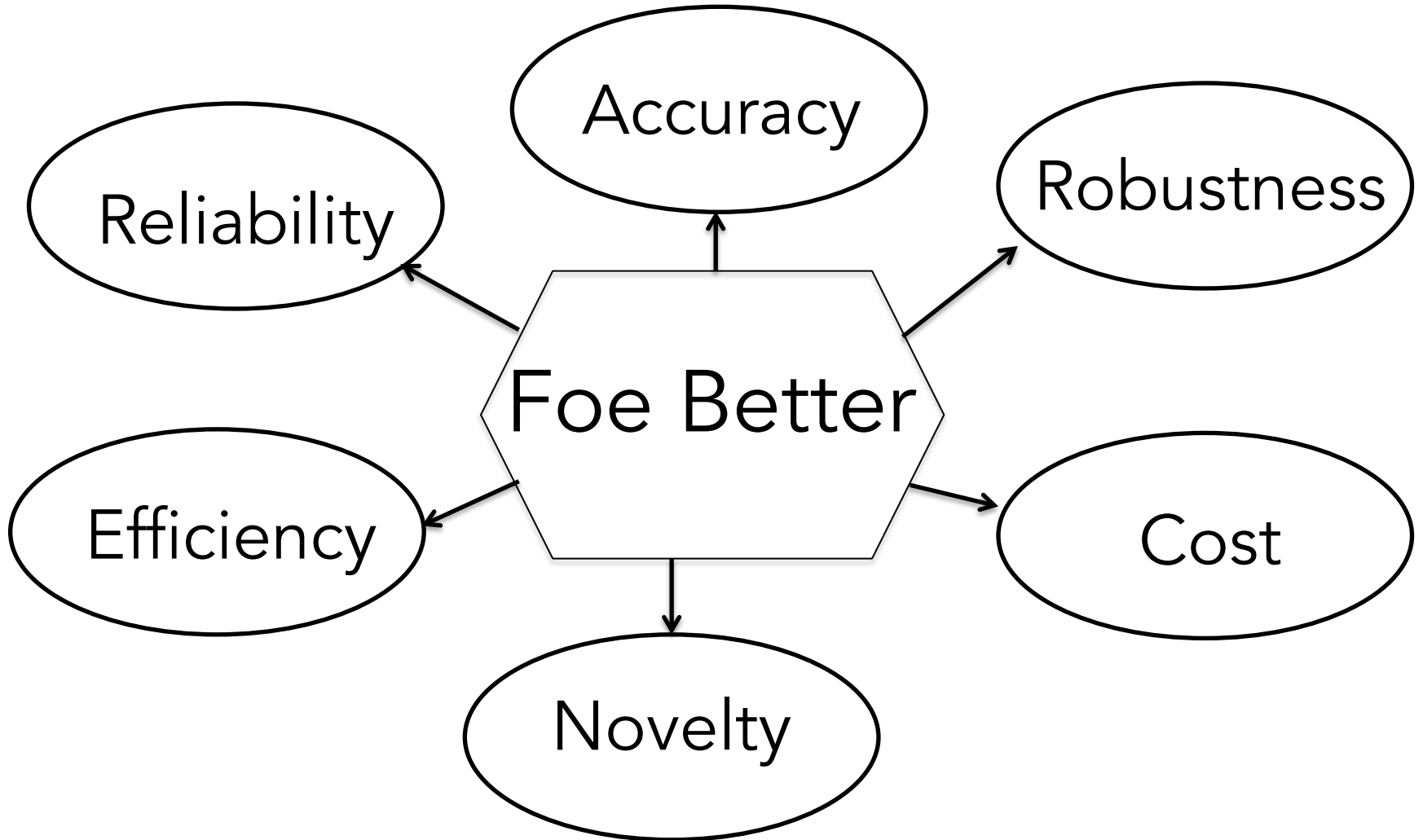
Information fusion is the process of combining information from multiple sources!

Challenges in Information Fusion

- Different media are captured at different rates and in different formats
- Processing times for different media streams is different which affects the real-time systems
- Varying confidence in the analysis tasks

How to deal with synchronization?

- Increase sampling rate with linear interpolation
- Reduce the sampling rate
- Take similar approach for feature synchronization

Questions

- When to fuse?
- What to fuse?
- How to fuse?

When to fuse?

- At equal time intervals?
- At certain events that are detected using subset of sources
- The subset may contain cheaper sensors that are processed more frequently
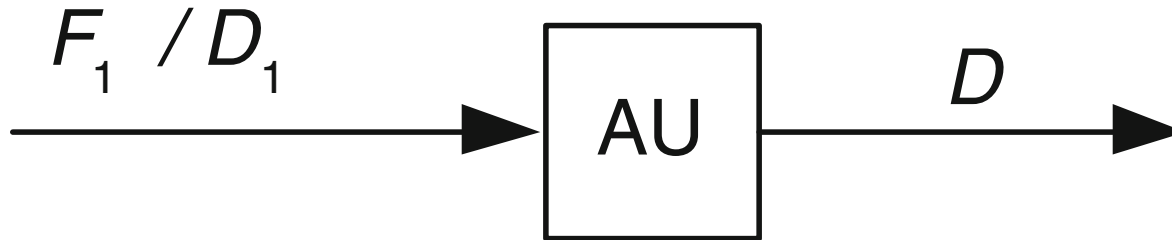
What to fuse?

The key factors are:

- Cost

- Processing time

- Accuracy/effectiveness
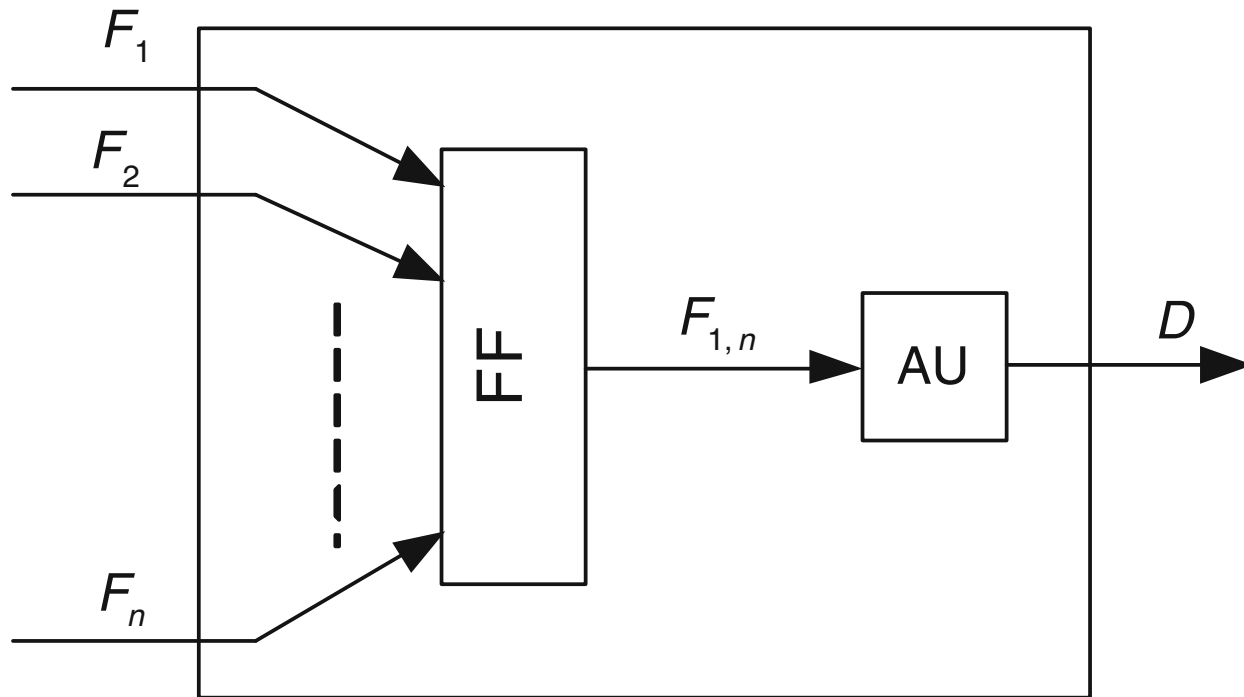
- End goal

How to fuse?

Fusion Levels

- Feature level (early fusion)
- Decision level (late fusion)
- Features as well as decisions (Hybrid)

Analysis Unit

$$F_1 \;/\; D_1 \longrightarrow \boxed{\text{AU}} \xrightarrow{\;\;D\;\;}$$

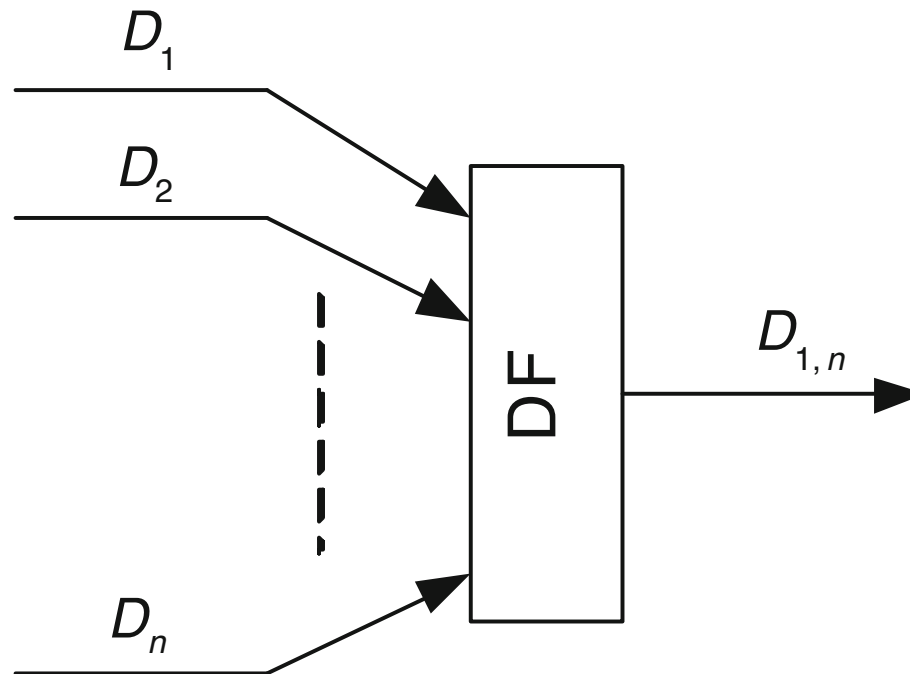# Early Fusion (Feature Level Fusion)

Early fusion requires single learning phase on combined vector!

Challenges

1. Synchronization
2. Feature representation
3. Single model may not fit all sources

Late Fusion (Decision Level Fusion)

Advantage

- Unlike features, the decisions have same representation

- Analysis units can be optimized to work for the given modality

Decision level fusion is unable to exploit feature level correlation!

Fusion Models

- Rule-based fusion methods
- Machine learning based methods

Rule-based fusion method

- MAX/MIN
- AND/OR
- Majority voting
- Linear weighted sum
- Weight powered product

MIN/MAX/AND/OR

- No additional training needed
- AND/OR need binary representation
- Which method is more prone to a noisy classifier?

What fusion strategy will lead to high recall?

- E.g. fire alarm
- Even if one of the classifiers is positive (such as fire alarm), use that classifier

Majority Voting

- Does not need training
- Special case of weighted sum with all weights equal
- Generally used for decision level fusion
- The final decision is the one for which majority of the classifiers agree

Linear Weighted Sum

$$I = \sum_{i=1}^{n} w_i \times I_i$$

Needs additional training to get weights!

Examples

- Track people using multiple sources (e.g. IR sensor and video), and average the tracks.

- Recognize humans based on face detection as well as speech detection and take average score.

- Take average of audio, video, and text matching scores in information retrieval!

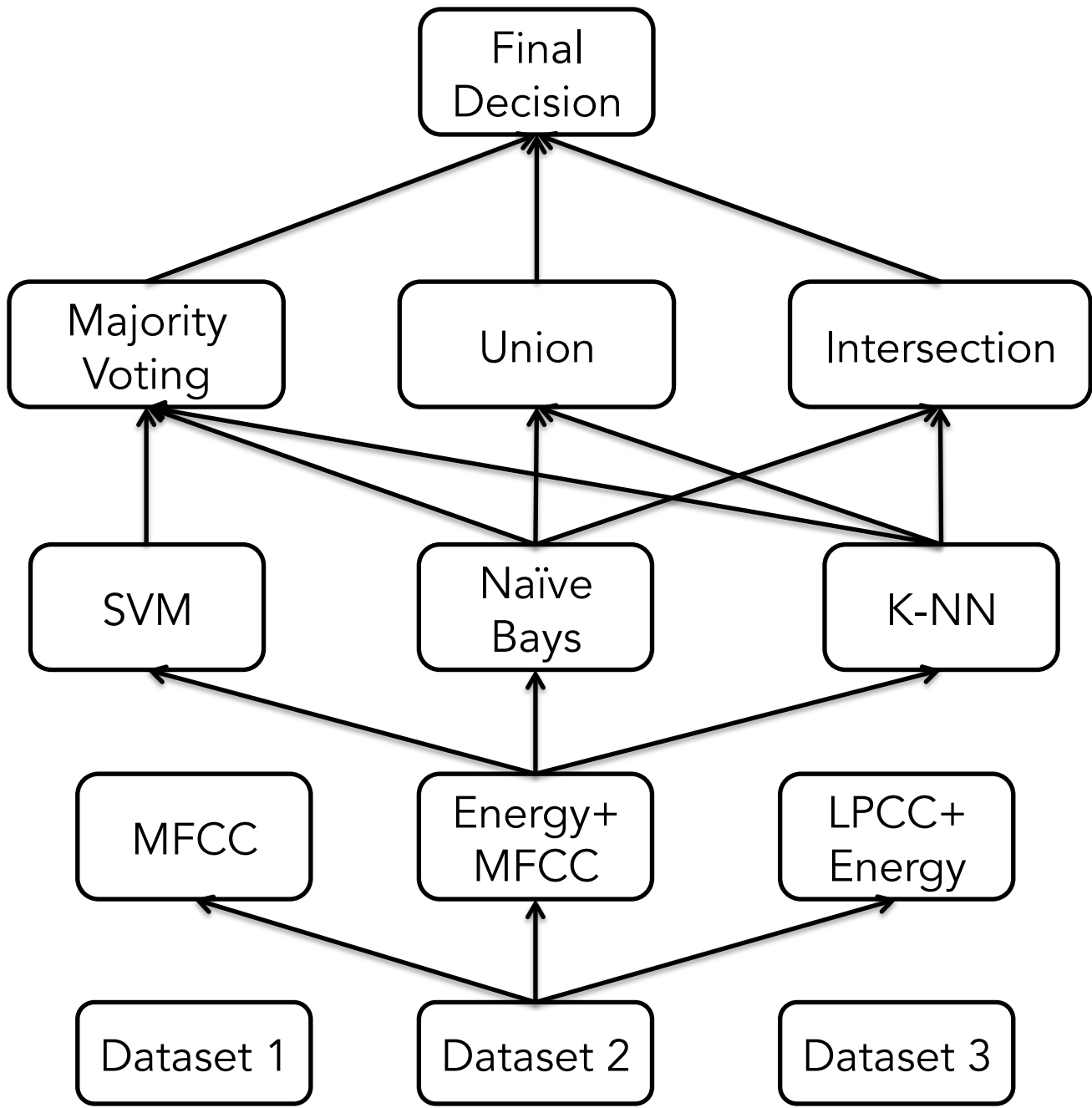# Weight Powered Product

$$I = \prod_{i=1}^{n} I_i{}^{w_i}$$

Needs additional training to get weights!

Machine Learning Based Methods

Multiple Classifier System

- The aim is to get more accurate classification at the expense of increased complexity.

- If you are using more than one classifier, it is called ensemble of classifiers or ensemble method.

How to obtain multiple classifiers or classifier ensemble?

Final Decision

Majority Voting   Union   Intersection   Use different combiners

SVM   Naïve Bays   K-NN   Train different base classifiers

MFCC   Energy+ MFCC   LPCC+ Energy   Use different feature sets

Dataset 1   Dataset 2   Dataset 3   Use different datasets