VIBE PROJECT

Virtual Biomedical and STEM/STEAM Education

2021-1-HU01-KA220-HED-000032251

Funded by the European Union

Erasmus+

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

PÉCSI TUDOMÁNYEGYETEM
UNIVERSITY OF PÉCS

U.PORTO

Politechnika Śląska

DEX innovation centre

VIBE PROJECT

PATTERN RECOGNITION IN BIOMEDICAL ENGINEERING

MATHEMATICAL FOUNDATIONS

Introduction

Books


WILLIAM GIBSON
Pattern Recognition
PATTERN RECOGNITION THIRD EDITION
SERGIOS THEODORIDIS KONSTANTINOS KOUTROUMBAS
WILLIAM GIBSON PATTERN RECOGNITION
PATTERN RECOGNITION AND MACHINE LEARNING
CHRISTOPHER M. BISHOP
PATTERN RECOGNITION SECOND EDITION
Pattern Classification
Mathematical Methodologies in Pattern Recognition and Machine Learning
Pattern Recognition Mechanisms
模式识别
Pattern Recognition for Medical Imaging
Anke Meyer-Bäse
Pattern Recognition
PATTERN RECOGNITION IN SPEECH AND LANGUAGE PROCESSING
PATTERN RECOGNITION IN BIOLOGY
Marsha S. Corrigan
Pattern Recognition and Machine Learning
T. Pavlidis
Structural Pattern Recognition
Pattern Recognition and Machine Learning
PATTERN RECOGNITION ON ORIENTED MATROIDS
Pattern Recognition with Neural Networks in C++
Abhijit S. Pandya
Robert B. Macy
Vein Pattern Recognition
A Privacy-Enhancing Biometric
Chuck Wilson
Train Your Chess Pattern Recognition
More Key Moves and Motifs in the Middlegame
ARTHUR VAN DE OUDEWEETERING
Pattern Recognition and Computer Vision
Syntactic Pattern Recognition
PATTERN RECOGNITION Human and Mechanical
Satosi Watanabe
PATTERN RECOGNITION BY SELF-ORGANIZING NEURAL NETWORKS
PATTERN RECOGNITION APPROACH TO DATA INTERPRETATION
Pattern Recognition and Image Preprocessing
Chess Pattern Recognition for Beginners
The Fundamental Guide to Spotting Key Moves in the Middlegame
Arthur van de Oudeweetering
The New York Times Bestseller
WILLIAM GIBSON

The Term "Pattern Recognition"

Pattern Recognition

is a field whose objective is to assign an object or event to one of a few categories, based on features derived to emphasize commonalities. In practice, features are often extracted from sensory signals, such as images or audio.

is the act of taking in raw data and taking an action based on the category of the pattern.

Terminology

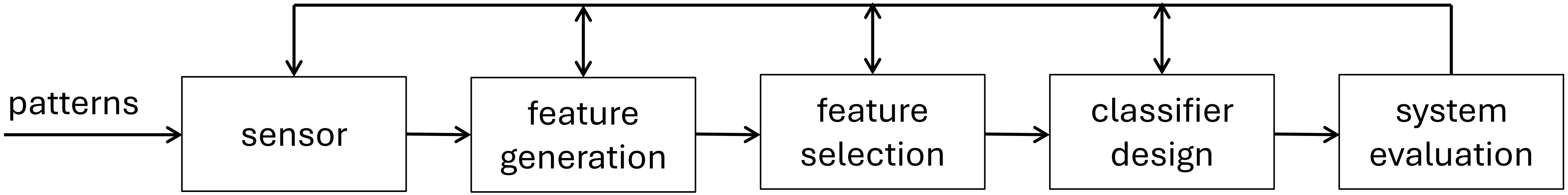
What is the difference between

Image Processing,
Image Recognition, and
Pattern Recognition?

Pattern recognition applications

- Computer Vision
- Character Recognition
- Computer-aided Diagnosis
- Speech Recognition
- Data Mining and Knowledge Discovery
- …

Basic Stages of Pattern Analysis



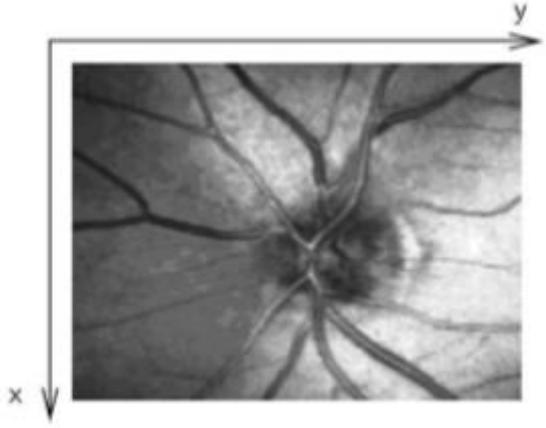
The stages are highly dependent on each other.

In order to design an optimal pattern recognition system, they all have to be optimised at once.

Patterns are analysed at different levels of abstraction.

Integration of background knowledge into the process may be very useful.

Low-Level Interpretation of Patterns

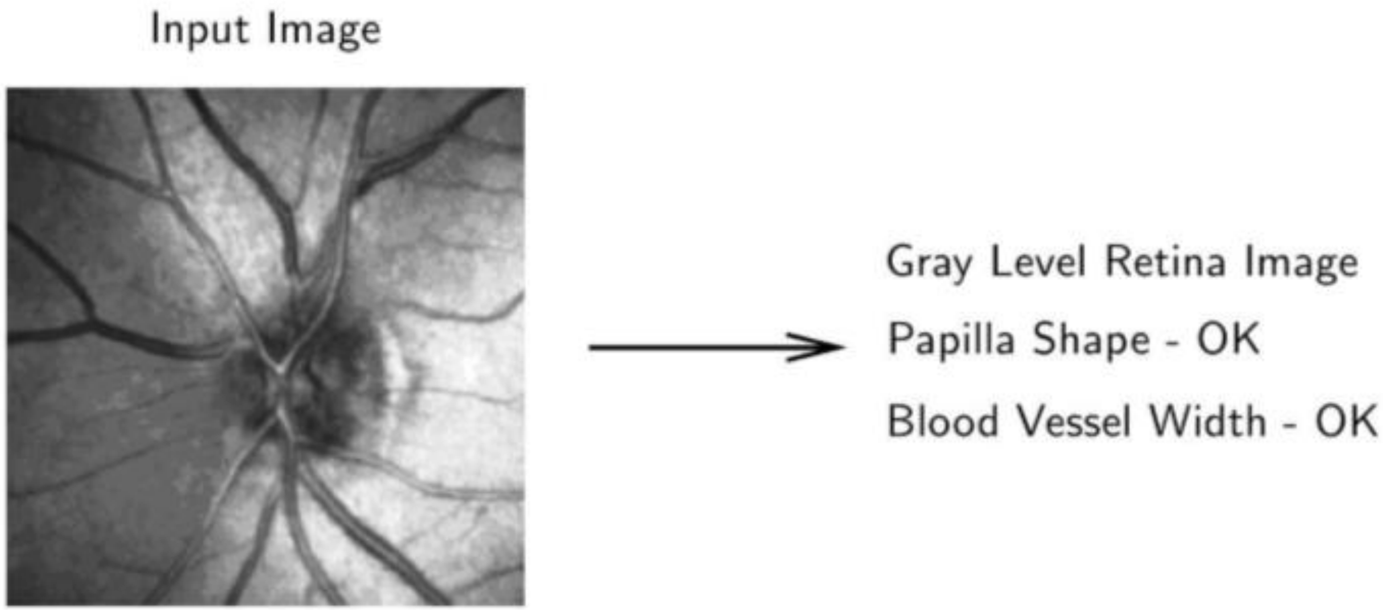


$$\longrightarrow \begin{pmatrix} 039 & 041 & 041 & 043 & \dots & 057 \\ 041 & 043 & 043 & 043 & \dots & 066 \\ 041 & 051 & 053 & 055 & \dots & 066 \\ 041 & 051 & 053 & 056 & \dots & 066 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 054 & 052 & 042 & 042 & \dots & 095 \end{pmatrix}_{150\times 200}$$

$$f(x,y) = \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0,199) \\ f(1,0) & f(1,1) & \dots & f(1,199) \\ \vdots & \vdots & & \vdots \\ f(149,0) & f(149,1) & \dots & f(149,199) \end{pmatrix} \quad ; \quad f(x,y) \in \{0,1,2,\dots,255\}$$

Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

High-Level Interpretation of Patterns



Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

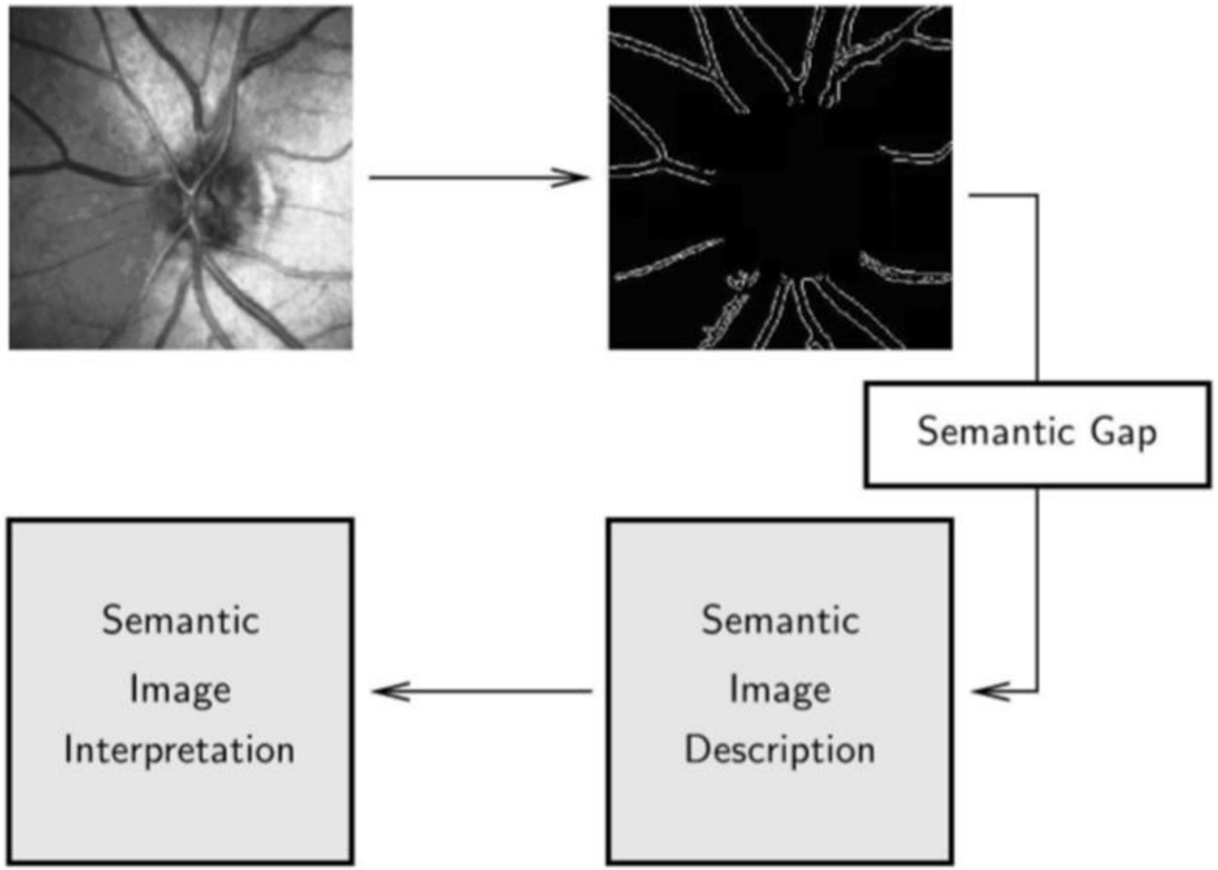
Basic Stages of Pattern Analysis

Optimization of the Entire Processing Chain at Once

Combination of the Different Levels of Abstraction

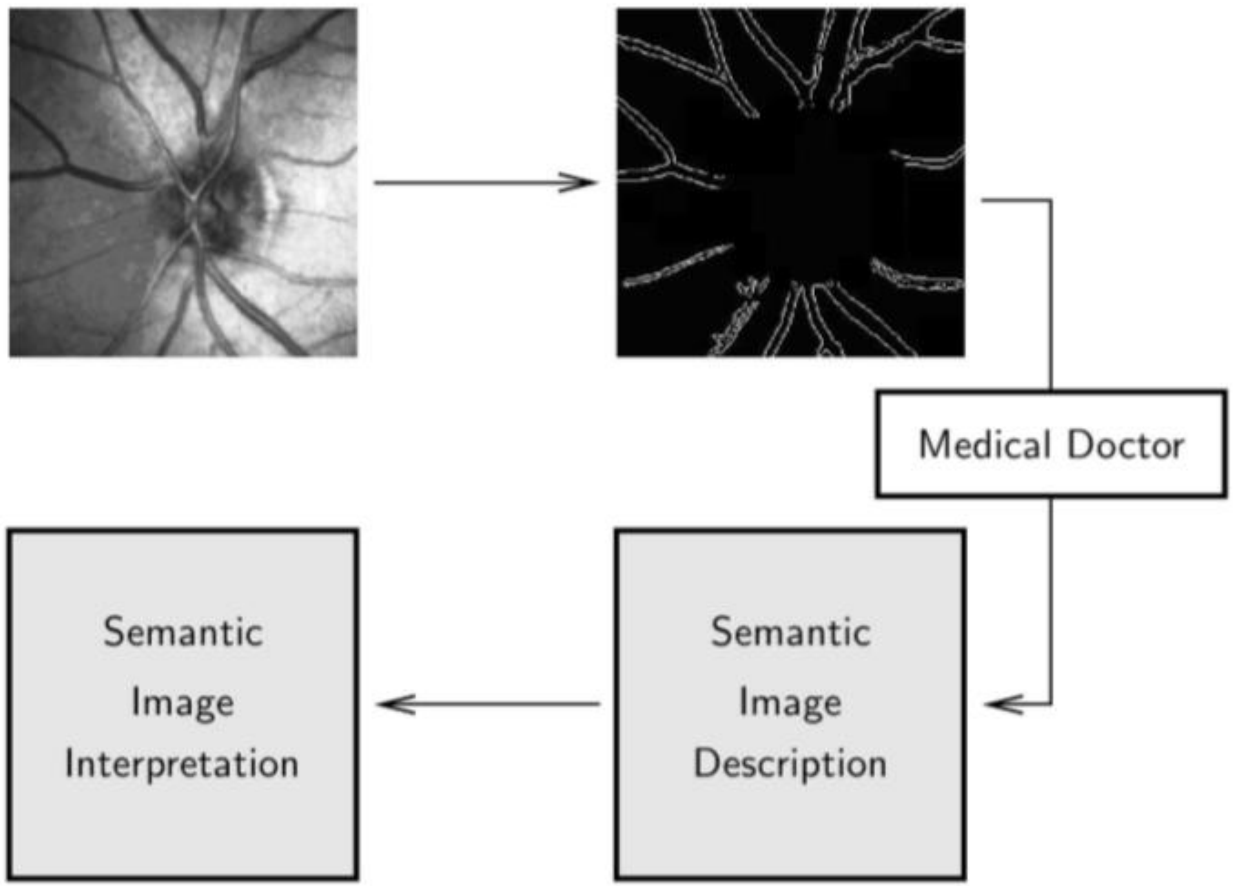
Integration of Background Knowledge into the Process

Semantic Gap in Image Understanding



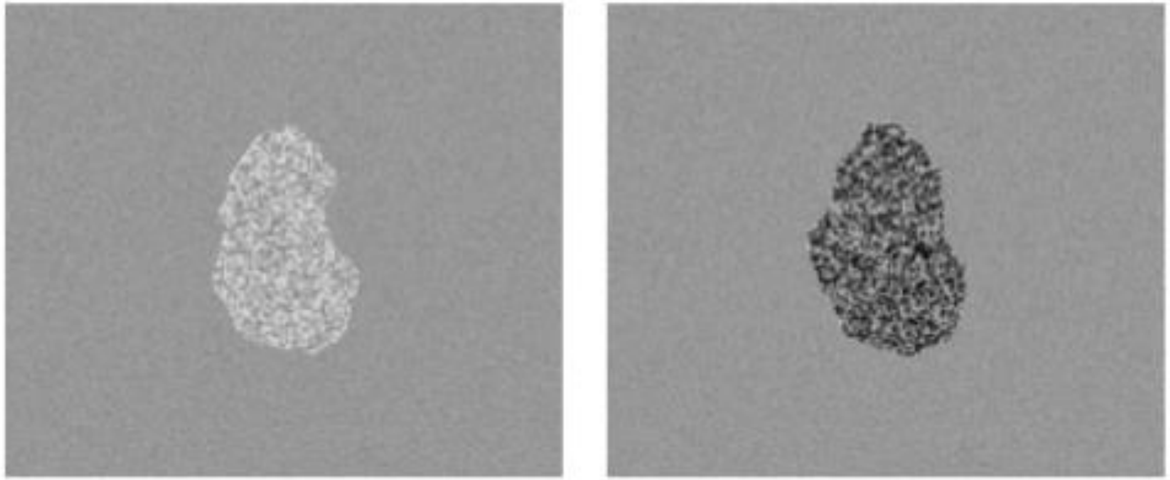
Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

Semantic Gap in Image Understanding



Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

Example for Medical Image Classification

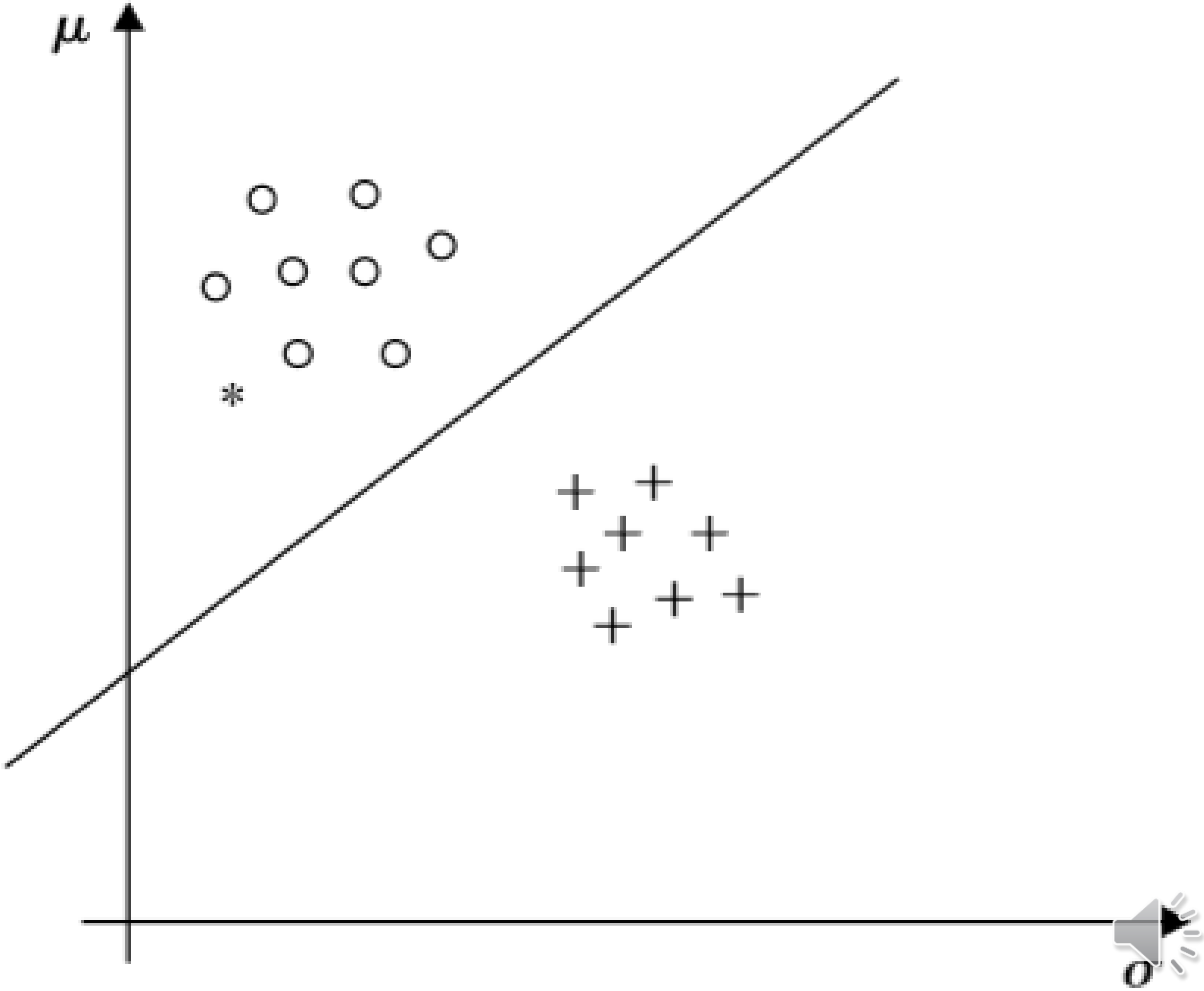


(a) (b)

Examples of image regions corresponding to (a) class A and (b) class B.

Example Descriptors for the Image Regions

- Plot of the mean value μ and standard deviation σ for a numer of different images originating from class A (◦) and class B (+).



Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

Feature Vectors → Random Vectors

Each feature vector identifies a single pattern (object)

Descriptors are called feature vectors

$$x = [x_1, x_2, \ldots, x_l]^T$$

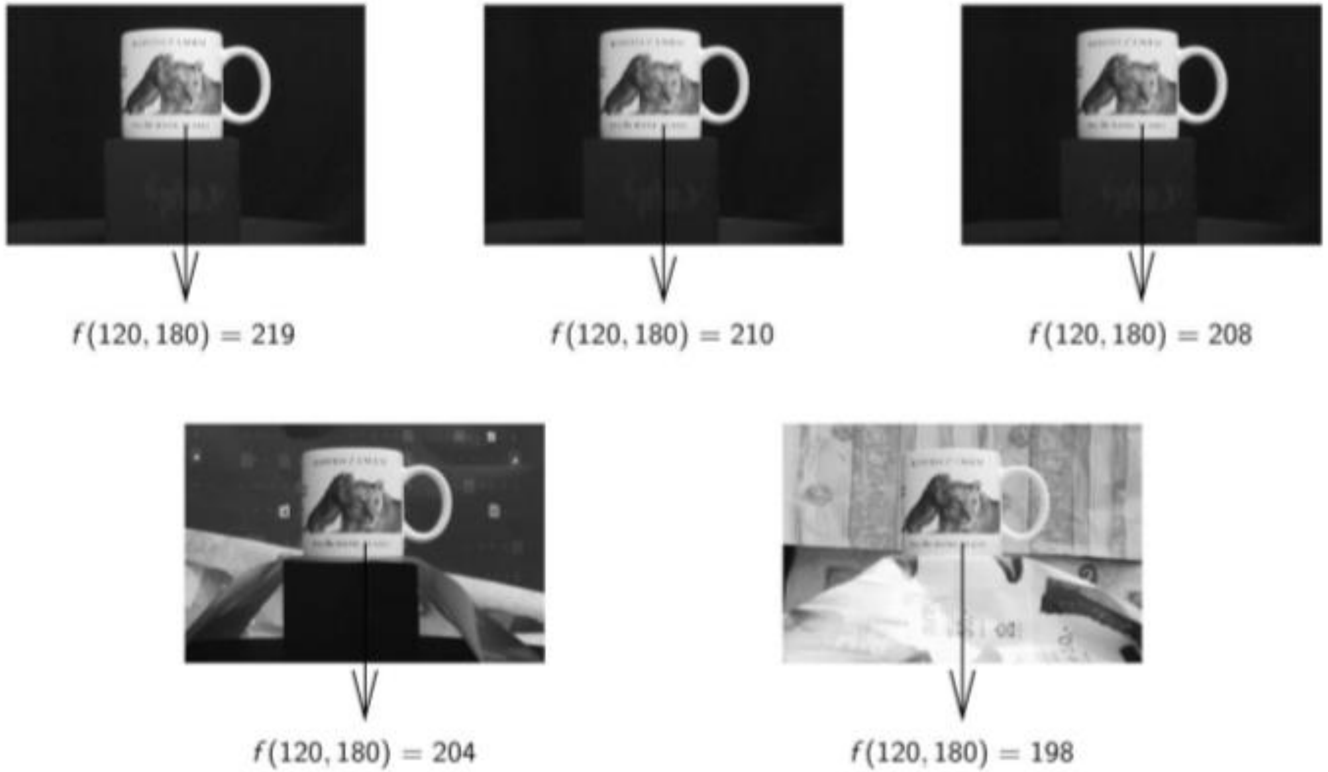
Feature vectors are treated as random vectors

Signal Acquisition - Stochastic Process

- Stochastic processes are processes that proceed randomly in time.
- Rather than consider fixed random variables X, Y, etc. or even sequences of i.i.d random variables, we consider sequences X_0, X_1, X_2, ... Where X_t represent some random quantity at time t.
- In general, the value X_t might depend on the quantity X_{t-1} at time t-1, or even the value X_s for other times $s < t$.

Signal Acquisition - Stochastic Process

Example

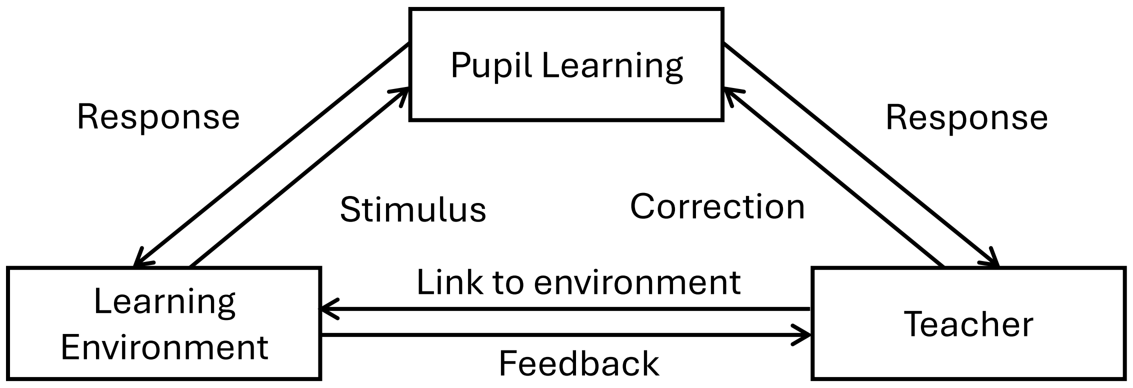


Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

Learning Strategies

Supervised Learning

assumes that a set of labelled training data is available and the classifier is designed by exploiting this a-priori known information.

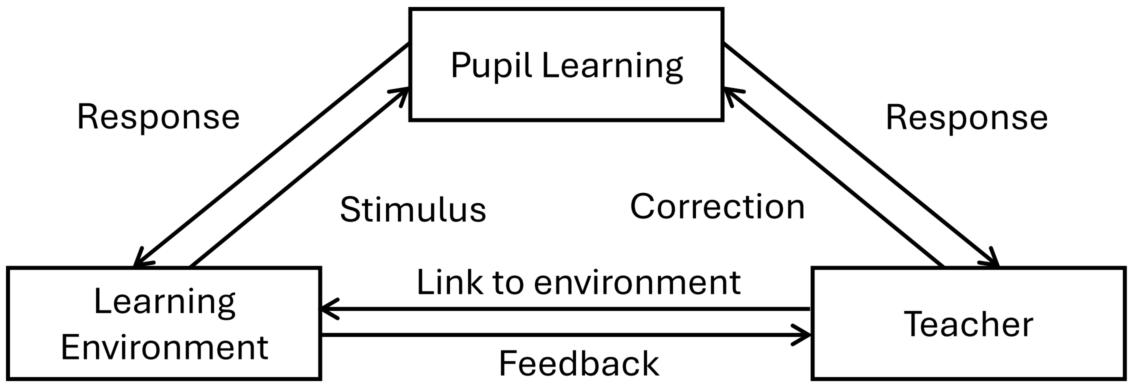


Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

Learning Strategies

Supervised Learning

assumes that a set of labelled training data is available and the classifier is designed by exploiting this a-priori known information.



Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

Learning Strategies

Semi-supervised Learning

applies both the labelled and unlabelled training for designing a classification system.

Statistical Classification - Problem Statement

Classification of an unknown pattern in the most probable of the classes!

- Set of classes: $\{\omega_1, \omega_2, \ldots, \omega_M\}$
- Unknown pattern represented by its feature vector x
- Conditional probabilities: $P(\omega_i|x)$, $i = 1,2,\ldots,M$
- Classification result: the class with the maximum conditional probability

But how to compute the conditional probability for a particular class?

Probability P vs. Density p

Probability P
is a real number describing an event belonging to the range <0,1>.

Density p
is a value of a function[1] p(x) describing the distribtion of the random variable x.

If the random variable takes only discrite values, the densities become probabilities!

[1]This function is often reffered as pdf – probability density function.

Bayes Decision Theory

A Priori Probability vs. A Posteriori Probability

A priori probability – probability before classification

- How probable is a particular class ω_i for a pattern **x** before applying any classificaion algorithm?
- Answer: $P(\omega_i)$

A posteriori probability – probability after classification

- How probable is a particular class ω_i for a pattern **x** after applying any classificaion algorithm?
- Answer: $P(\omega_i|x)$

Likelihood Density Function

Likelihood Density Function	How feature vector **x** are distributed in a class ω_i?
	Answer: $p(x\|\omega_i)$
	$p(x\|\omega_i)$ is the likelihood function of ω_i with respect to x
	$p(x\|\omega_i)$ can be trained from examples

Two-class Problem

Known

- Classes: $\{\omega_1, \omega_2\}$
- A priori probabilities: $P(\omega_1)$ and $P(\omega_2)$
- Likelihood density functions: $p(x|\omega_1)$ and $p(x|\omega_2)$
- Pattern to be classified: $x=[x_1, x_2, ..., x_l]^T$

Assumption

- The feature vectors can take any value in the l-dimensional feature space: $x=[x_1, x_2, ..., x_l]^T \in \mathbb{R}$

Unknown

- A posteriori probabilities: $P(\omega_1 |x)$ and $P(\omega_2 |x)$

Computation of the A Posteriori Probability

Using the Bayes Rule

$$P(\omega_i|x) = \frac{p(x|\omega_i)P(\omega_i)}{p(x)}, \qquad \mathrm{i} = 1, 2 \qquad (1)$$

p(x) – density function for x

Bayes Classification Rule (1)

Higher a posteriori probability wins

If $P(\omega_1|x) > P(\omega_2|x)$, x is classified to ω_1

If $P(\omega_1|x) < P(\omega_2|x)$, x is classified to ω_2

Bayes Classification Rule (2)

Considering the Bayes Rule (Eq. 1)

If $\frac{p(x|\omega_1)P(\omega_1)}{p(x)} > \frac{p(x|\omega_2)P(\omega_2)}{p(x)}$, x is classified to ω_1

If $\frac{p(x|\omega_1)P(\omega_1)}{p(x)} < \frac{p(x|\omega_2)P(\omega_2)}{p(x)}$, x is classified to ω_2

Bayes Classification Rule (3)

p(x) can be discarded, because it is the same for all classses

If $p(x|\omega_1)P(\omega_1) > p(x|\omega_2)P(\omega_2)$, x is classified to ω_1

If $p(x|\omega_1)P(\omega_1) < p(x|\omega_2)P(\omega_2)$, x is classified to ω_2

Bayes Classification Rule (4)

If the a priori probabilities are equal: $P(\omega_1) = P(\omega_2)$

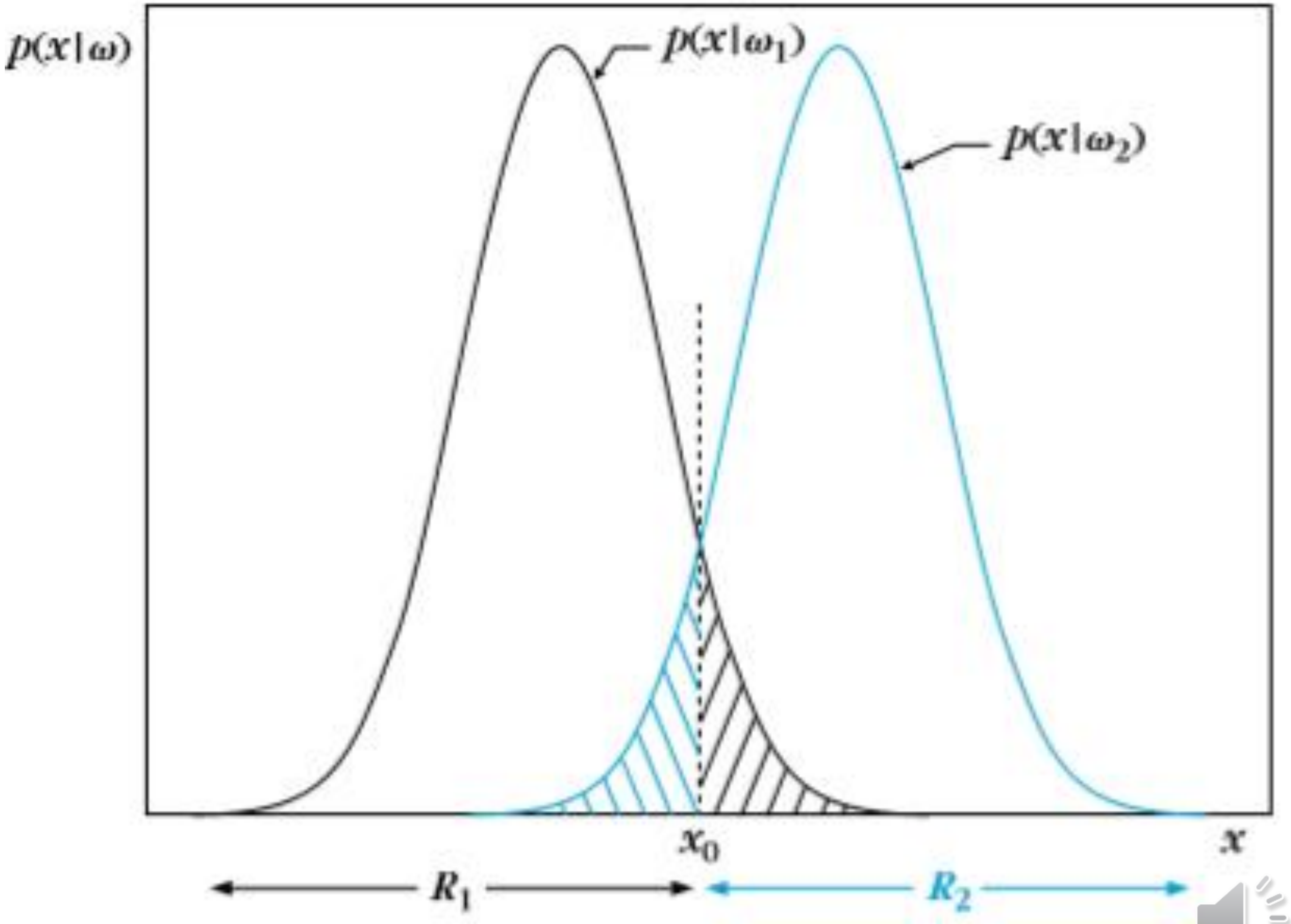
If $p(x|\omega_1) > p(x|\omega_2)$, x is classified to ω_1

If $p(x|\omega_1) < p(x|\omega_2)$, x is classified to ω_2

We are done, since the likelihood density functions $p(x|\omega_1)$ and $p(x|\omega_2)$ are assumed to have been trained from examples!

Classification Error Probability

Error Probability: $P_e = \frac{1}{2}\int_{-\infty}^{x_0} p(x|\omega_2)\mathrm{dx} + \frac{1}{2}\int_{x_0}^{\infty} p(x|\omega_1)\mathrm{dx}$



Source: https://medium.com/@thommaskevin/tinyml-gaussian-naive-bayes-classifier-31f8d241c67c

Classification Error Probability in General

- A priori probabilities are not equal: $P(\omega_1) \neq P(\omega_2)$
- Feature vectors have more than one dimension: $l > 1$

$$\mathbf{x} = [x_1, x_2, \ldots, x_l]^{\mathrm{T}}$$

- General form:

$$P_e = P(\omega_1) \int_{R_2} p(\mathbf{x}|\omega_1) d\mathbf{x} + P(\omega_2) \int_{R_1} p(\mathbf{x}|\omega_2) d\mathbf{x}$$

Grzegorzek Marcin & Doniec Rafał ,. (2024). *Pattern Recognition*. University: Universität zu Lübeck

Pattern Recognition Quiz

1. What is Pattern Recognition?

•A. Pattern recognition is identifying patterns without using any pre-learned information.

•B. Pattern recognition is the classification of data based on previously gained knowledge or statistical information extracted from patterns.

2. Is Speech Recognition an example of Pattern Recognition?

•A. Yes, as it involves processing raw data and classifying patterns for machine use.

•B. No, speech recognition doesn't involve identifying or classifying patterns.

3. What is the Difference Between Classification and Clustering?

•A. Classification assigns labels based on training patterns, while clustering groups data without predefined labels.

•B. Classification and clustering are the same process, as both involve labeled data.

4. Can Binary Quantities be Used as Features?

•A. No, features can only be represented as continuous variables.

•B. Yes, features can be continuous, discrete, or binary variables.

5. How are Features Obtained?

•A. Features are randomly generated without any measurement criteria.

•B. A feature is a function of measurements that quantify significant characteristics of an object.

Pattern Recognition Quiz with Correct Answers

1. What is Pattern Recognition?

•A. Pattern recognition is identifying patterns without using any pre-learned information.

•B. Pattern recognition is the classification of data based on previously gained knowledge or statistical information extracted from patterns.

Answer: B

2. Is Speech Recognition an example of Pattern Recognition?

•A. Yes, as it involves processing raw data and classifying patterns for machine use.

•B. No, speech recognition doesn't involve identifying or classifying patterns.

Answer: A

3. What is the Difference Between Classification and Clustering?

•A. Classification assigns labels based on training patterns, while clustering groups data without predefined labels.

•B. Classification and clustering are the same process, as both involve labeled data.

Answer: A

4. Can Binary Quantities be Used as Features?

•A. No, features can only be represented as continuous variables.

•B. Yes, features can be continuous, discrete, or binary variables.

Answer: B

5. How are Features Obtained?

•A. Features are randomly generated without any measurement criteria.

•B. A feature is a function of measurements that quantify significant characteristics of an object.

Answer: B

Thank you for your attention