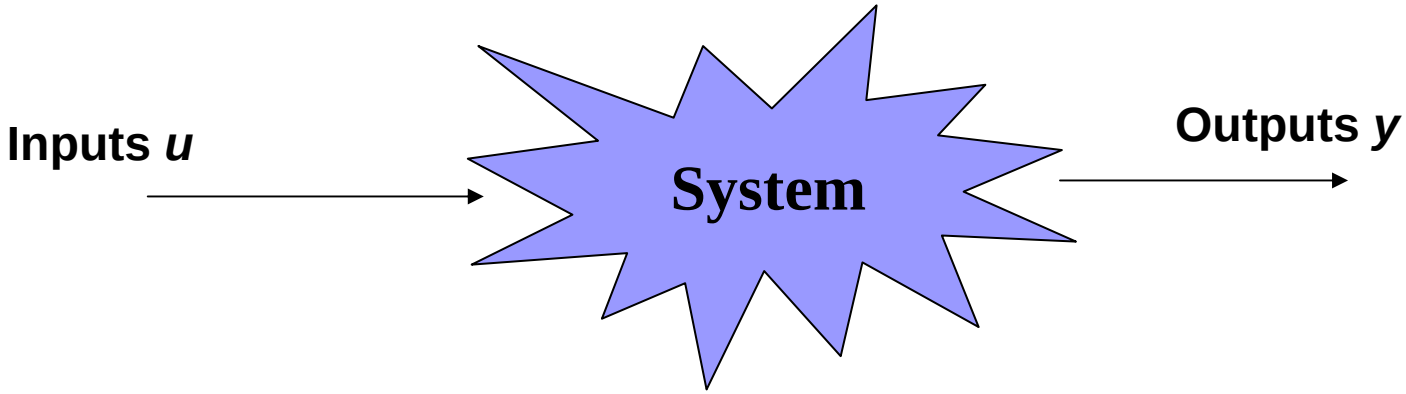
Knowledge Helper for Medical and Other Information users

Moamar SAYED MOUCHAWEH
Centre de Recherche en STIC (CReSTIC)
Université de Reims

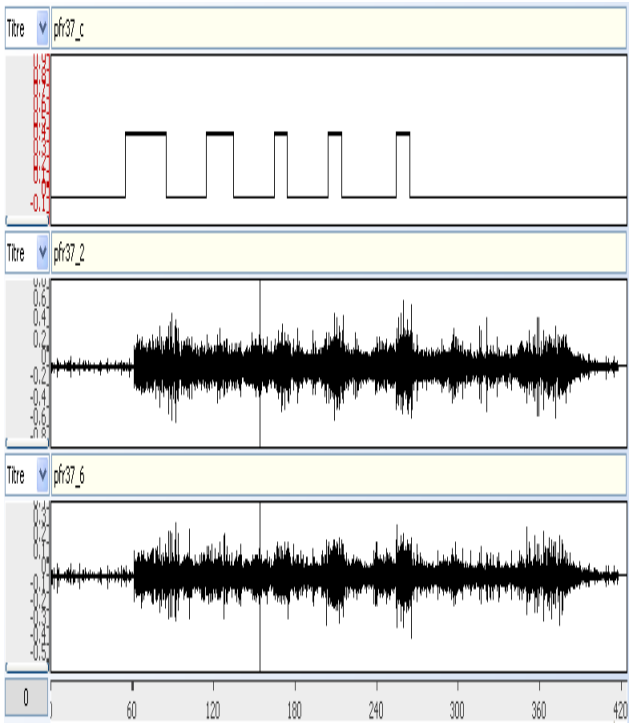
Introduction (What)

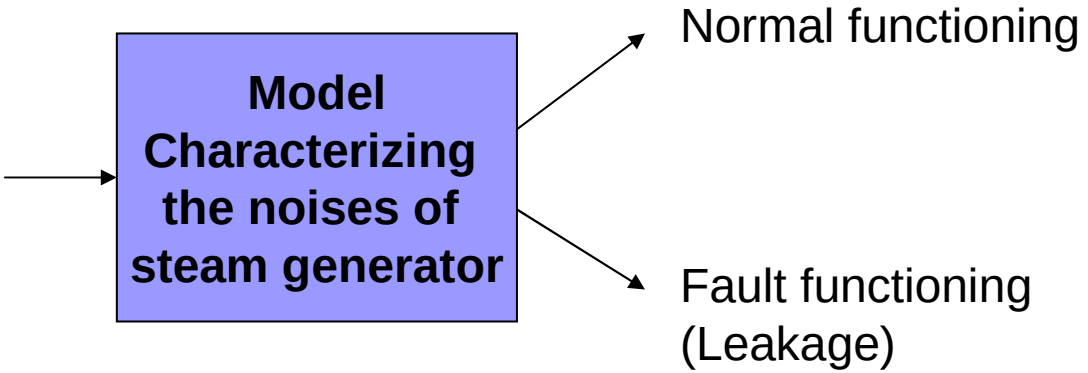


- Prediction of outputs in response to inputs
- Construction of a model based on the available information about system behavior

Introduction (Examples)

- Enhancing the monitoring online of the functioning of Sodium Fast Reactors using acoustic signals

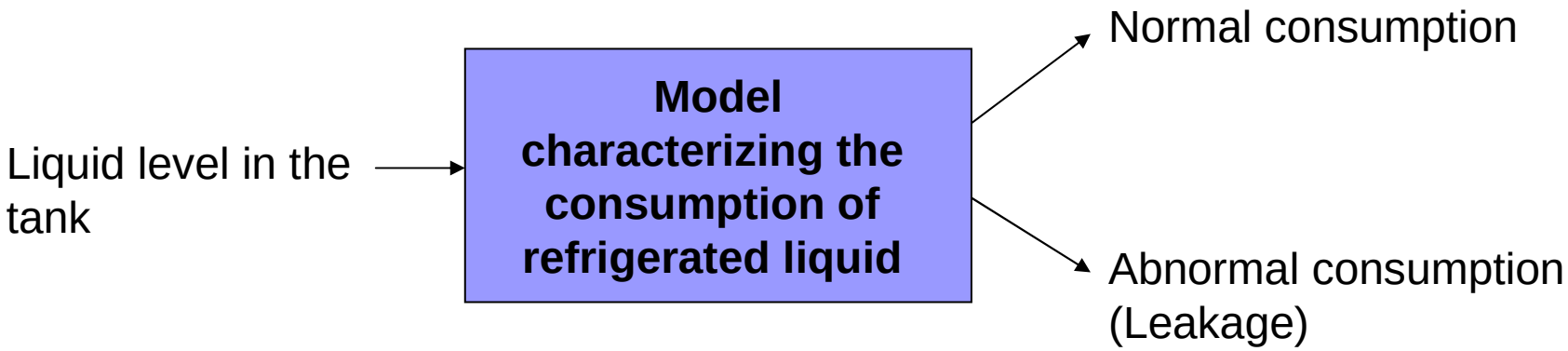




Available information: Measurements of the noises of the steam genretaor

Introduction (Examples)

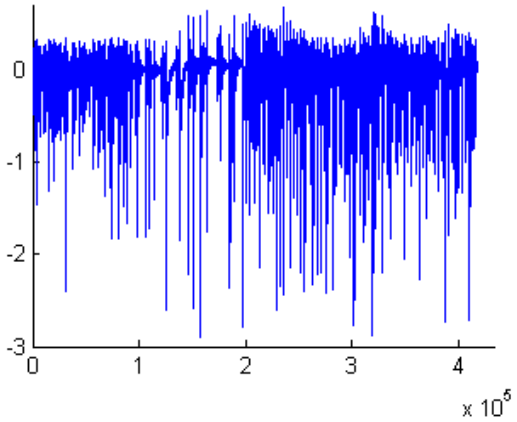
- Premature detection of micro leaks in refrigerated liquid tanks



Available information: measurements about the liquid level in the tank during the past consumption

Introduction (Examples)

- Monitoring of the quality of folded pieces



Model characterizing the quality of folded pieces

Good quality

Bad quality

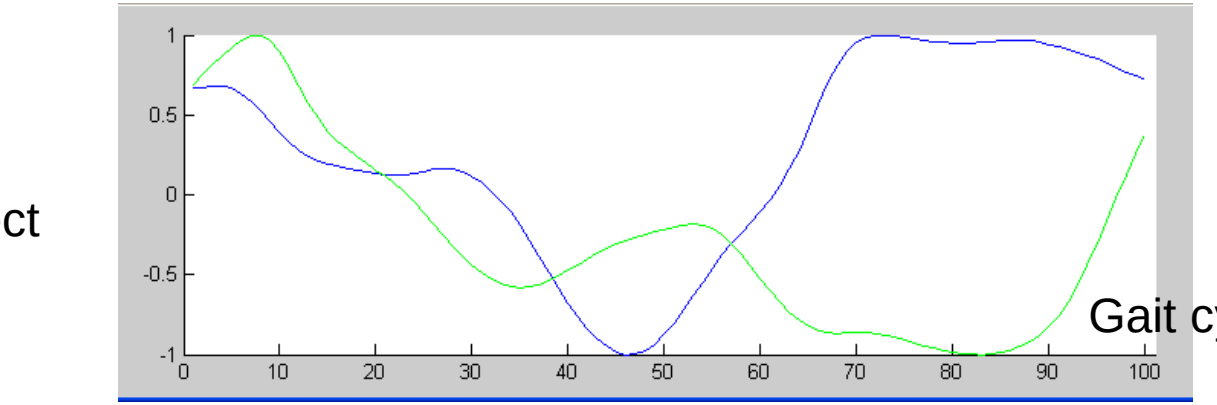
Available information: Measurements about the noises of folding process + human experience

Introduction (Examples)

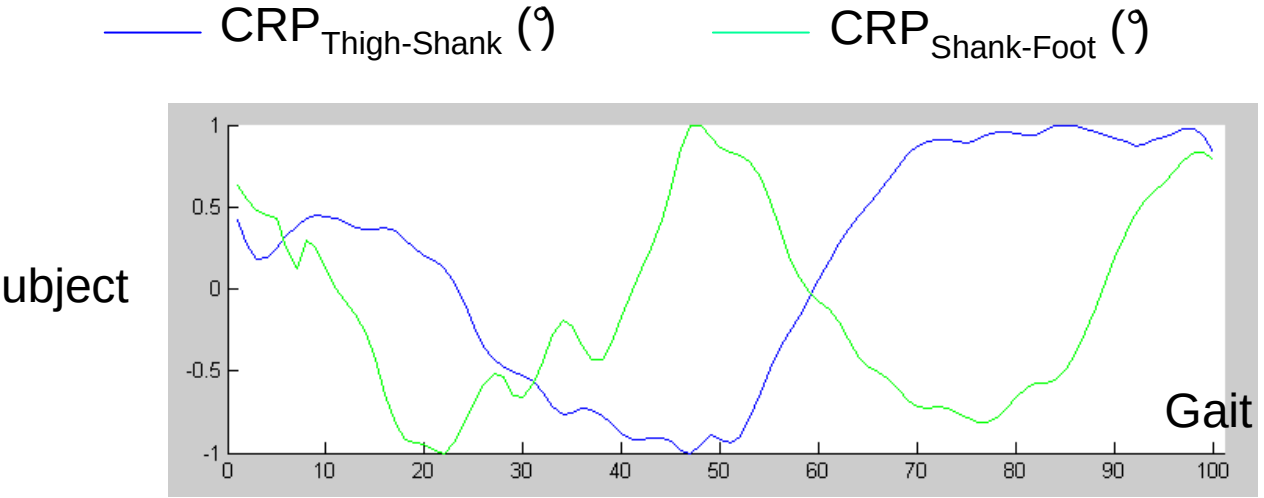
- Descrimination of hemiparetic patients and follow up of their response to a specific medical treatment
  - Hemiparetic patients present gait disturbances resulting of lesions in their central nervous system
  - Quantifying of the inter-segmental coordination between two body segments using the Continuous Relative Phase (CRP) measure

Introduction (Examples)

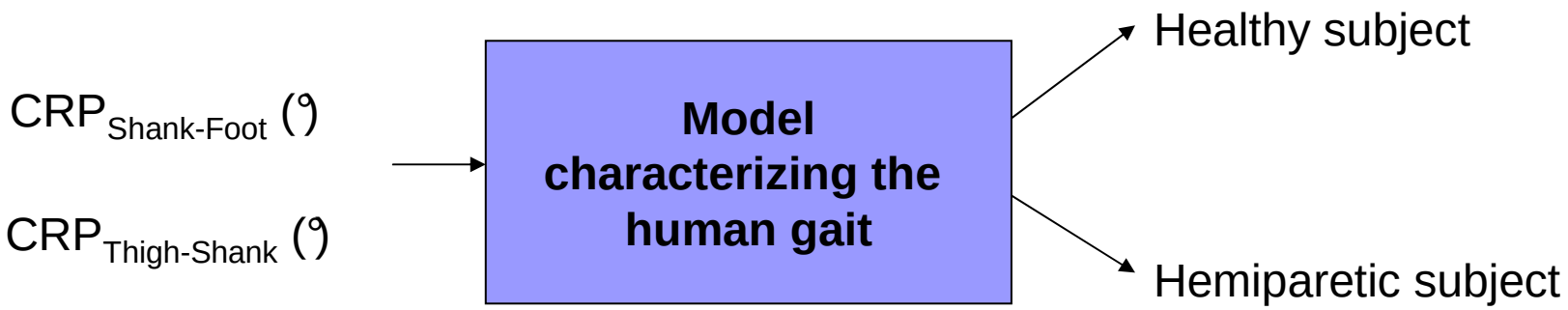
Healthy subject



Hemiparetic subject

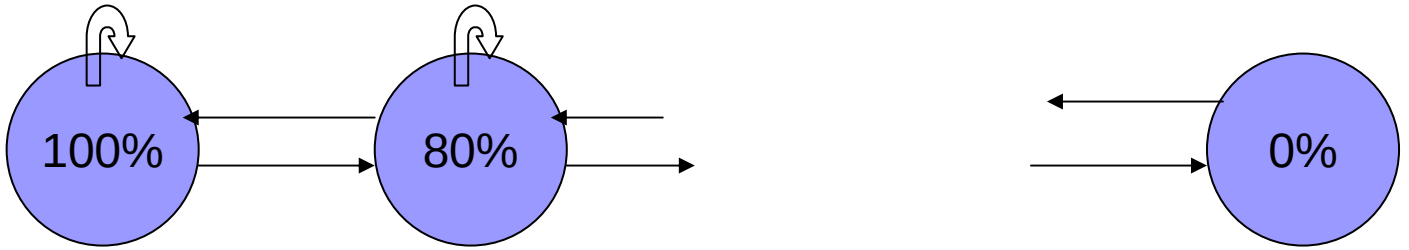


Introduction (Examples)



Available information: Two signals + doctors evaluations

Quantification of the positive response to treatment



Introduction (Aims)

- Effective automated information extraction
  - Automated analysis and indexing for medical information (images, data, texts, ..)
  - Trustable results at a level of understability adapted to users (personalization to the class and expertise of user
  - Helping clinicians in reasoning over similar cases (Diagnosis aid)
  - Providing updated, related and timely results

Methods (How)

- Two categories of methods:
  - Model-based methods (Quantitative [GEN 07]/ Qualitative [CAS 99])
  - Reasoning-based methods [DUB 90], [DUD 01]
- The available information, objectives and system size determine the method(s) to be used

Methods (Constraints)

- Available information
  - Many sources (Images 2D, 3D and 4D, Texts, measurements, reports on patients, diagnosis, ..)
  - Different representation
  - Different level of trust and target user expertise
  - Huge size (125 TB)
  - Exponential increasing rate (100 GB of images per day from 400 patients)
  - Evolving environement

Methods (Constraints)

- Available information
  - Conflict data
  - Incomplete, imprecise and uncertain information
  - High dimensional data
- Objectives
  - Making decision in real time
  - Taking into account the collaboration between different actors of different expertise levels and domains
  - Following a situation evolution over time

Methods (Constraints)

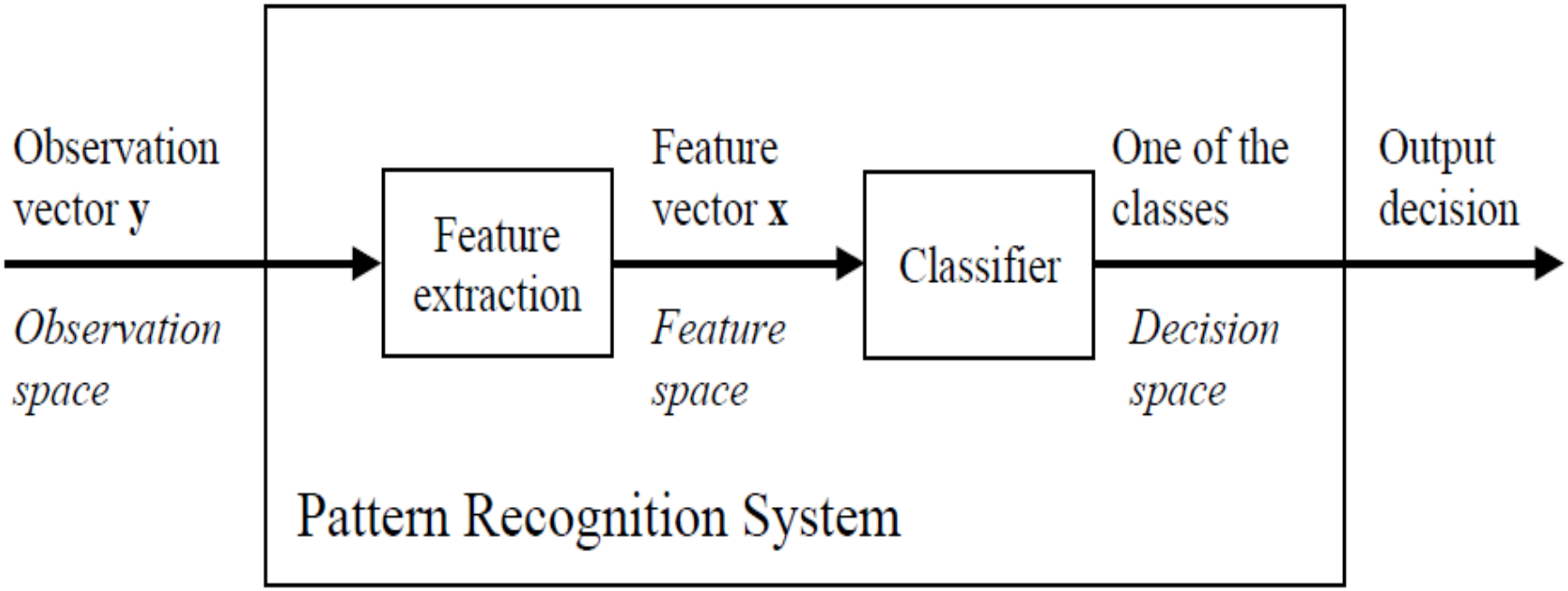
- System size
  - Centralized
  - Dcentralized

Methods (How)

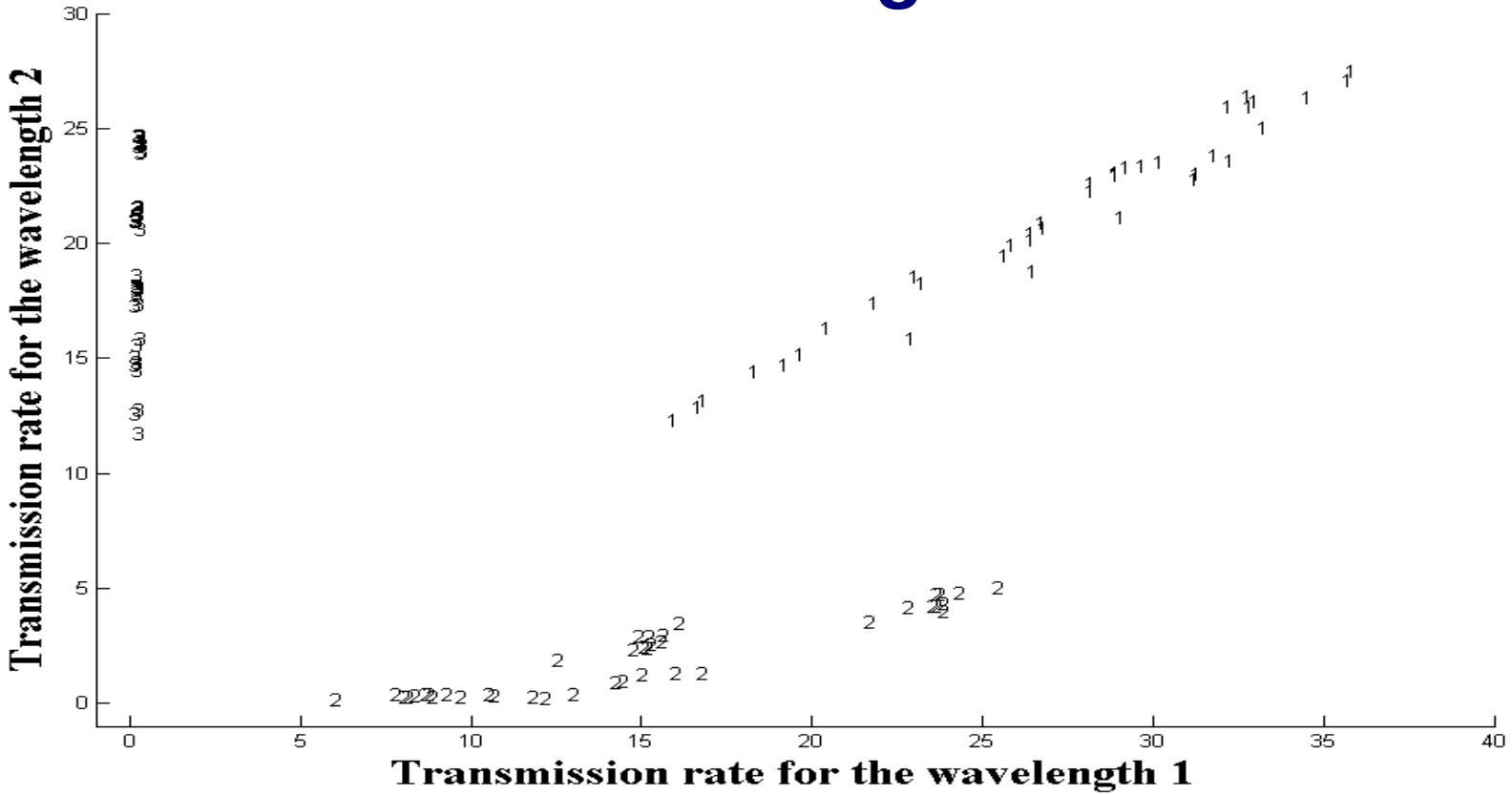
- Reasoning-based methods:
  - Knowledge-based methods [DUB 01]
  - Machine learning methods
- Machine learning methods
  - Statistical Pattern Recognition [DUD 01]
    - ✓ Parametric [DUB 90]
    - ✓ Non-parametric [PAR 62] [ANI 00]
  - Structural/syntaxic Pattern Recognition [CHE 90] [FU 82]
  - Neural networks [ZWI 95], [DUB 01]
  - Support Vector Machines (SVMs) [VAP 99], [CHA 08]

Pattern Recognition

- Pattern recognition principal



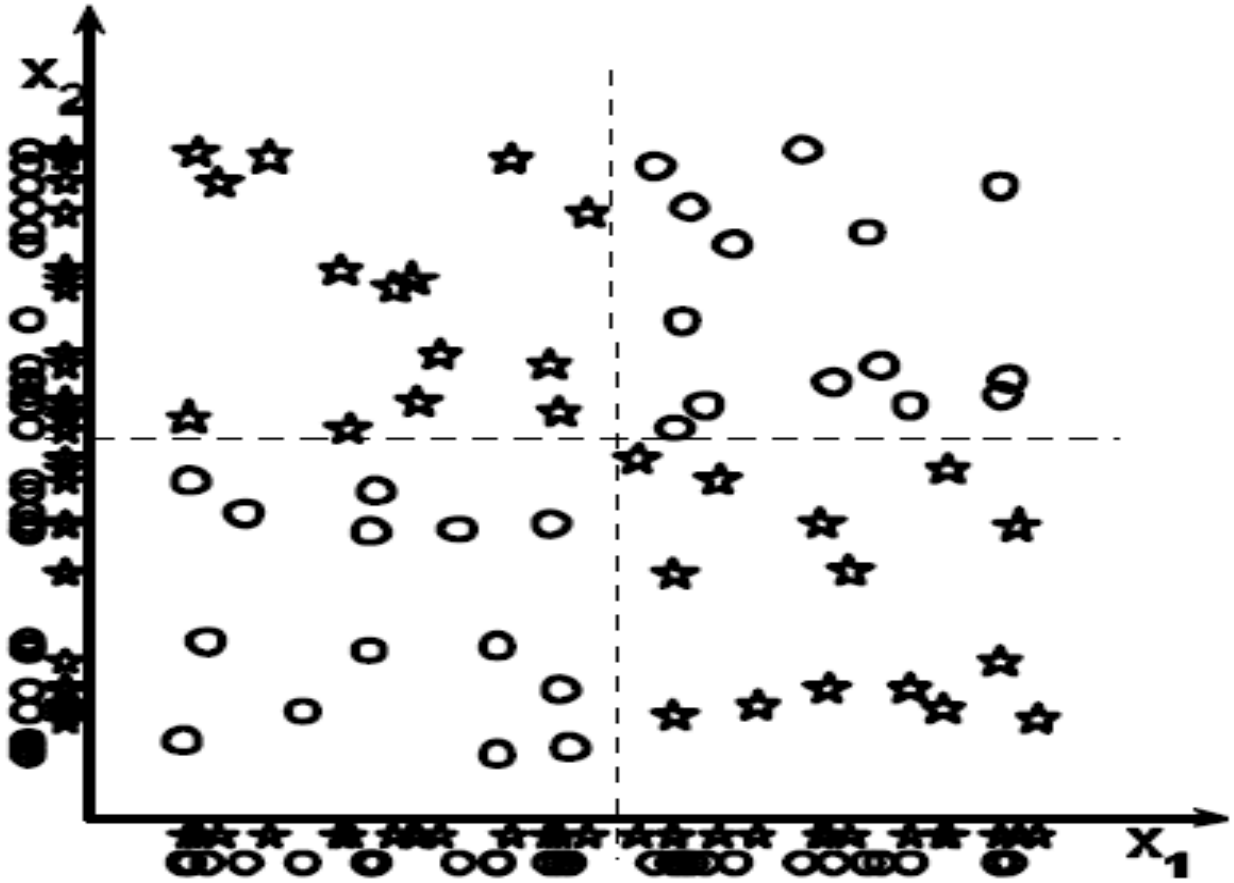
Pattern Recognition



Transmission rates according to the two wavelenghs → **Model characterizing the type of plastic material** → PVC (Class 1)
→ PEHD (Class 2)
→ PET (Class 3)

Pattern Recognition

- XOR example

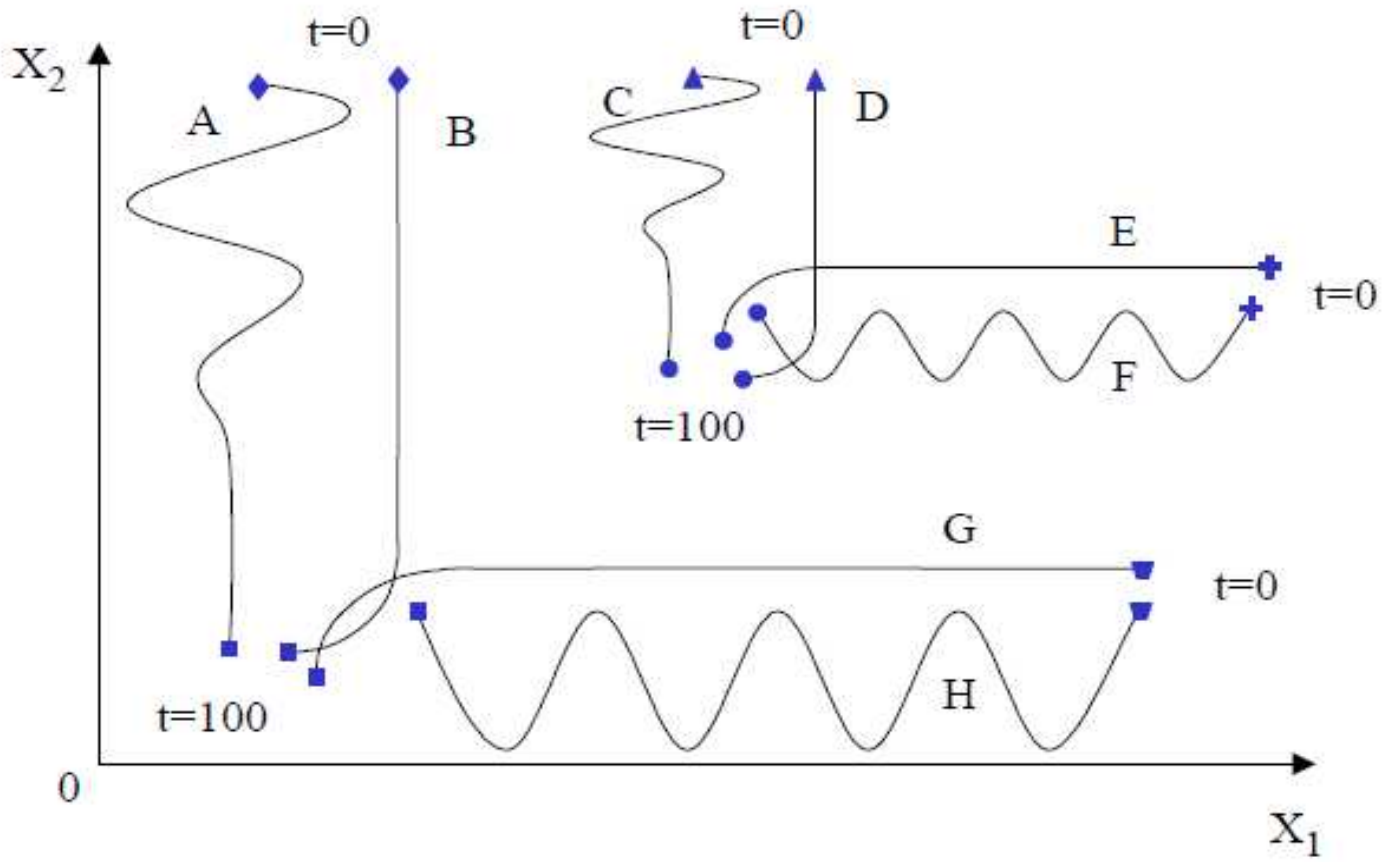


Pattern Recognition (How)

- Filter method ranks features or feature subsets independently of the predictor (classifier) [SAE 07]
  - Univariate method: considers one variable at a time [BEN 82]
  - Multivariate method: considers subsets of variables together [HAL 99]
- Wrapper method**:** uses a classifier to assess features or feature subsets [KIT 78]

Pattern Recognition (How)

- Structural Pattern recognition principal



Pattern Recognition (How)

- Supervised classification methods [DUD 01]
  - Fuzzy methods [ZAD 65]
  - Possibilistic methods [DUB 88], [ZAD 78]
  - Evidence methods [SHA 76]
- Unsupervised classification methods
  - Hierarchical methods [LEB 95]
  - Partitioning methods [BEZ 81]
- Semi-supervised classification methods [CHA 06]
- Semi-supervised clustering methods [BAS 02]
- Semi-supervised learning [SAY 10]

Pattern Recognition (How)

- Active learning [OLS 09]
- Incremental learning [SAY 02]
- Multi-classifiers [CHE 97]
  - Classifier fusion (Serial/Parallel)
  - Classifier selection (static/dynamic)
- Dynamic Pattern Recognition
  - Substitution of patterns [NAK 97], [LEC 06]
  - Selection of useful patterns [ANG 00], [MAR 98], [HAR 10]

Summary

- Active, incremental, dynamic semi-supervised learning
- Dynamic multi-classifiers selection system
- Dynamic feature space
- Hybrid (structural/statistical) Pattern Recognition
- Adaptive Human-Machine Interface
- Decentralized structure of processing

References

[ANG 00] Angstenberger, L., Dynamic Fuzzy Pattern Recognition, Dissertation, Fakultät für Wirtschaftswissenschaften der Rheinisch-Westfälischen Technischen Hochschule, Aachen, Germany, 2000.

[ANI 00] Anil, K.J., Robert, P.W., Mao, J., Statistical Pattern Recognition : A review, IEEE Transactions on pattern analysis and machine intelligence, Vol. 1, N°22, 4-36, 2000.

[BAS 02] Basu, S., Banerjee, A., Mooney, R. J., Semi-supervised clustering by seeding, 19th International Conference on Machine Learning (ICML-2002), 2002.

[BEN 82] Ben-Bassat, M., Pattern recognition and reduction of dimensionality, Handbook of Statistics II, Vol. 1, Holland, Amsterdam, 773–791, 1982.

References

[BEZ 81] Bezdek, J.C., Pattern recognition with fuzzy objective function algorithms, New-York, Plenum Press, 1981.

[CAS 99] Cassandra, C.G., Lafortune, S., Introduction to Discrete Event Systems", Kluwer Academic Publisher, 1999.

[CHA 06] Chapelle, O., Schölkpf, B., Zien, A., Semi-Supervised Learning, The MIT Press, 2006.

[CHA 08] Chapelle, O., Sindhwani, V., Keerthi, S.S., Semi-Supervised Support Vector Machines, Journal of Machine Learning Research 9, 203-233, 2008.

[CHE 90] Cheung, J.T., Stephanopoulos, G., Representation of process trends Part I. A formal representation framework., Computer & Chemical Engineering 14, N°4-5, 495-510, 1990.

References

[CHE 97] Chen, K., Wang, L., Chi, H., Methods of combining multiple classifiers with different features and their application to text-independent speaker identification, International Journal of Pattern Recognition and Artificial Intelligence, 417-445, 1997.

[DUB 88] Dubois, D., Prade, H., Possibility Theory, Plenum Press, New-York, 1988.

[DUB 90] Dubuisson, B., Diagnostic et reconnaissance des formes, Traité des Nouvelles Technologies, série Diagnostic et Maintenance, Hermès Sciences, 1990.

[DUB 01] Dubuisson, B., Automatique et statistiques pour le diagnostic, Traité IC2 Information, commande, communication, Hermès Sciences, 2001.

[DUD 01] Duda, R.O., Hart, P.E., Stork, D.E., Pattern Classification, Wiley, New York, Second edition, 2001.

References

[FU 82] Fu, K.S., Syntactic Pattern Recognition and Applications, Prentice Hall, Englewood Cliffs, New Jersey, 1982.

[FUN 01] Fung, G., Mangasarian, O., Optimization Techniques for Semi-supervised support vector machines for unlabeled data classification, Optimization Methods and Software 15, 29–44, 2001.

[GEN 07] Gentil, S., Supervision des procédés complexes, Traité IC2, Série systèmes automatisés, 2007.

[HAL 99] Hall, M., Correlation-based feature selection for machine learning. PhD Thesis, Department of Computer Science, Waikato University, New Zealand, 1999.

[HAR 10] Hartert, L., Sayed Mouchaweh, M., Billaudel, P., Monitoring of non stationary systems using dynamic Pattern Recognition, Chapitre N° 15 of IGI book on "Intelligent Industrial Systems: Modelling, Automation and Adaptive Behaviour", IGI Global, 2009.

References

[KIT 78] Kittler, J., Pattern Recognition and Signal Processing, Chapter Feature Set Search Algorithms Sijthoff and Noordhoff, Alphen aan den Rijn, Netherlands, 41–60, 1978.

[KIT 98] Kitttler, J., Hatef, M., Duin, R.P.W., Matas, J., On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell. 20, N°3, 226-239, 1998.

[LEB 95] Lebart, L., Morineau, A., Piron, M., Statistique exploratoire multidimensionnelle, Dunod, Paris, 1995.

[LEC 06] Lecoeuche, S., Mercère, G., Amadou-Boubacar, H., Modelling of Non-stationary Systems based on a Dynamical Decision Space, 14th IFAC Symposium on System Identification (SYSID06), 2006.

[MAR 98] Marsili-Libelli, S., Adaptive Fuzzy Monitoring and Fault Detection, International Journal of COMADEM 1, N°3, 31-37, 1998.

References

[NAK 97] Nakhaeizadeh, G., Taylor, C., Kunisch, G., Dynamic Supervised Learning. Some Basic Issues and Application Aspects, Classification and Knowledge Organization, Springer Verlag, Berlin, Heidelberg, 123-135, 1997.

[OLS 09] Olsson, F., A literature survey of active machine learning in the context of natural language processing, SICS Technical Report T2009:06. Swedish Institute of Computer Science, Kista, Sweden, 2009.

[PAR 62] Parzen, E., On the estimation of a probability density function and mode, Annals of Mathematical. Statistics 33, 1065-1076, 1962.

[SAE 07] Saeys, Y., Inza, I., Larranaga, P., A review of feature selection techniques in bioinformatics, bioinformatics 23, N° 19, 2507–2517, 2007.

[SHA 76] Shafer, G., A mathematical theory of evidence, Princeton, University press, 1976.

References

[SAY 02] Sayed Mouchaweh, M., Devillez, A., Lecolier, V. G., Billaudel, P., Incremental learning in Fuzzy Pattern Matching, Fuzzy Sets and Systems 132, N°1, 49-62, 2002.

[SAY 10] SAYED MOUCHAWEH, M., Semi Supervised Classification Method for Dynamic Applications, Fuzzy Sets and Systems 161, 544-563, 2002.

[VAP 99] Vapnik, V., The Nature of Statistical Learning Theory, Springer, 1999.

[ZAD 65] Zadeh, L.A., Fuzzy Sets, Information and Control 8, 338-353, 1965.

[ZAD 78] Zadeh, L.A., Fuzzy sets as a basis for a theory of possibility, Fuzzy sets and systems 1, 3-28, 1978.

[ZWI 95] Zwingelstein, G., Diagnostic des défaillances, Traité des Nouvelles Technologies, série Diagnostic et Maintenance, Hermès, 1995.