MACHINE LEARNING

Liviu Ciortuz
Department of CS, University of Iaşi, România

What is Machine Learning?

- **ML studies algorithms that $\underbrace{\text{improve with}}_{\text{learn from}}$ experience.**

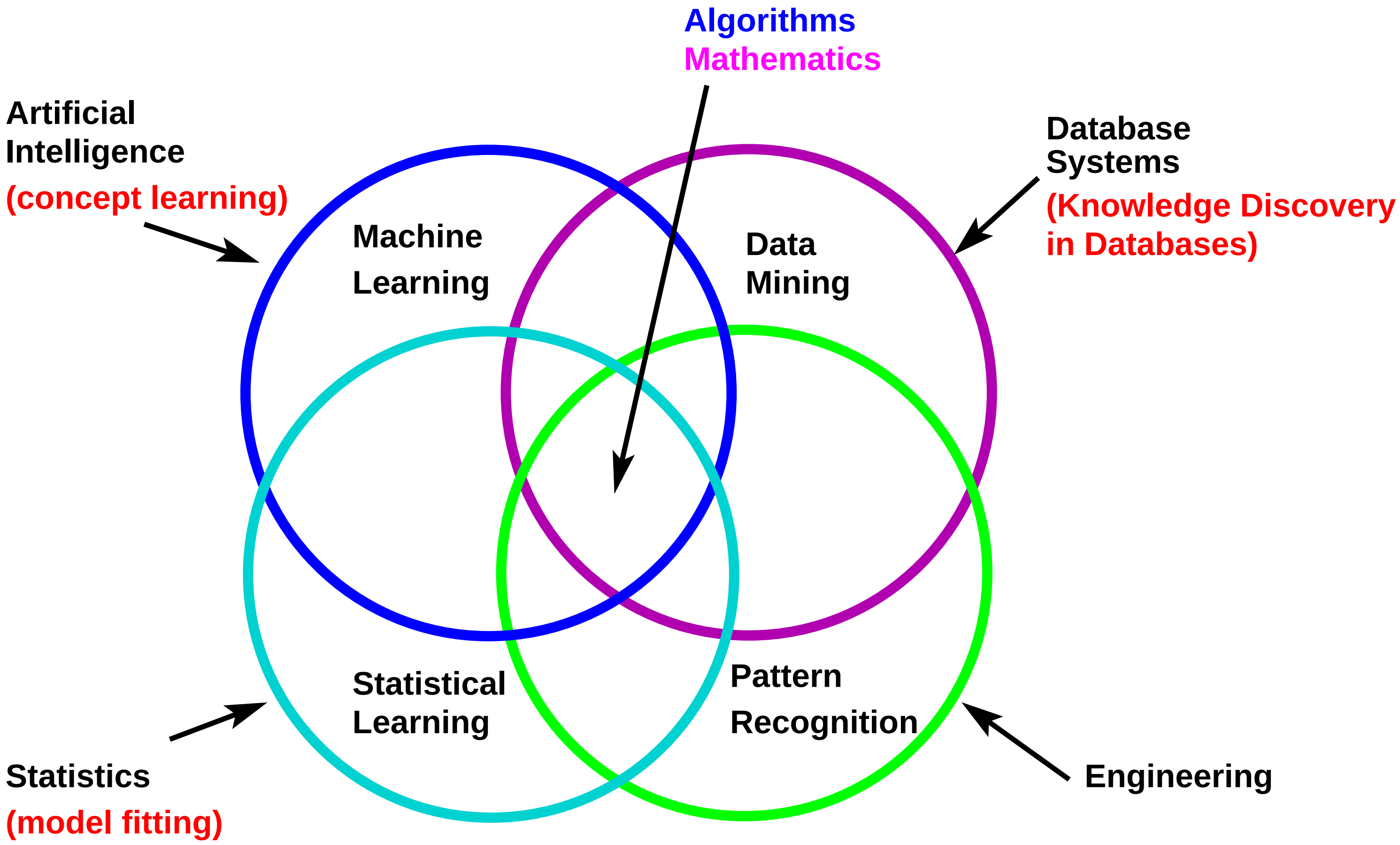
  Tom Mitchell's Definition of the [*general*] *learning problem*:
  "A computer program is said to *learn* from experience E with respect to some class of *tasks* T and *performance measure* P, if its performance on tasks in T, as measured by P, improves with experience E."

- **Examples of [specific] learning problems (see next slide)**
- **[Liviu Ciortuz:] ML is data-driven programming**
- **[Liviu Ciortuz:] ML gathers a number of well-defined subdomains/disciplines, each one of them aiming to solve in its own way the above-formulated [*general*] *learning problem*.**

What is Machine Learning good for?

- natural language (text & speech) processing
- genetic sequence analysis
- robotics
- customer (financial risc) evaluation
- terrorist threat detection
- compiler optimisation
- semantic web
- computer security
- software engineering
- computer vision (image processing)
- etc.

A multi-domain view



The Machine Learning Undergraduate Course: Plan

0. Introduction to Machine Learning (T. Mitchell, ch. 1)

---

1. Probabilities Revision (Ch. Manning & H. Schütze, ch. 2)

---

2. Decision Trees (T. Mitchell, ch. 3)
3. Parameter estimation for probablistic distributions
   (see *Estimating Probabilities*, additional chapter to T. Mitchell's book, 2016)
4. Bayesian Learning (T. Mitchell, ch. 6)
   and the relationship with Logistic Regression
5. Instance-based Learning (T. Mitchell, ch. 8)
6. Clustering Algorithms (Ch. Manning & H. Schütze, ch. 14)

The Machine Learning Master Course:

Tentative Plan

1. Probabilities Revision (Ch. Manning & H. Schütze, ch. 2)
2. Decision Trees: Boosting
3. Gaussian Bayesian Learning
4. The EM algorithmic schemata (T. Mitchell, ch. 6.12)
5. Support Vector Machines (N. Cristianini & J. Shawe-Taylor, 2000)

---

6. Hidden Markov Models (Ch. Manning & H. Schütze, ch. 9)
7. Computational Learning Theory (T. Mitchell, ch. 7)

Bibliography

0. **“Exerciţii de învăţare automată”**
   L. Ciortuz, A. Munteanu E. Bădărău.
   Iaşi, Romania, 2019
   www.info.uaic.ro/∼ciortuz/ML.ex-book/book.pdf

---

1. **“Machine Learning”**
   Tom Mitchell. McGraw-Hill, 1997

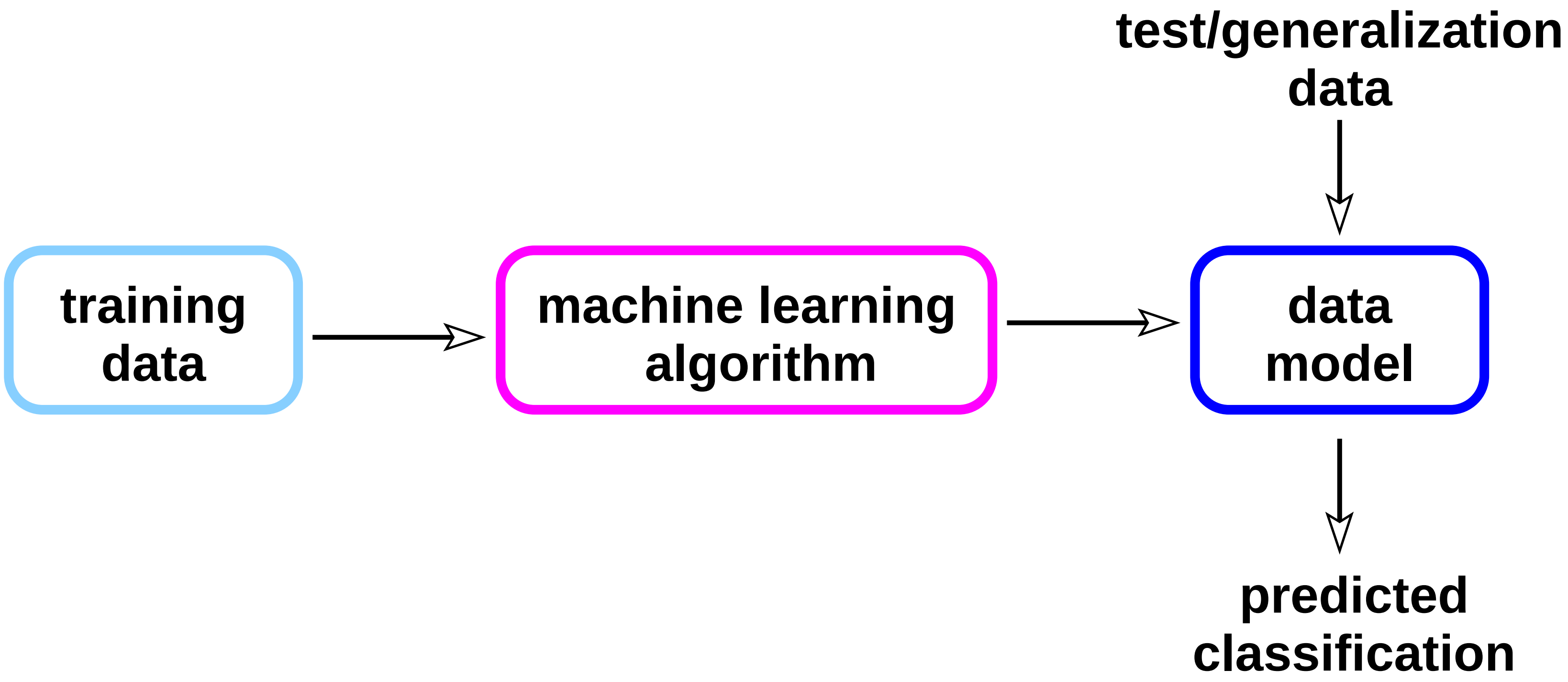
---

2. **“The Elements of Statistical Learning”**
   Trevor Hastie, Robert Tibshirani, Jerome Friedman. Springer, 2nd ed. 2009

---

3. **“Machine Learning – A Probabilistic Perspective”**
   Kevin Murphy, MIT Press, 2012
4. **“Pattern Recognition and Machine Learning”**
   Christopher Bishop. Springer, 2006
5. **“Foundations of Statistical Natural Language Processing”**
   Christopher Manning, Hinrich Schütze. MIT Press, 2002

A general schema for machine learning methods



"We are drawning in information but starved for knowledge."

John Naisbitt, "Megatrends" book, 1982

Basic ML Terminology

1. **instance x, instance set X**
   concept $c \subseteq X$, or $c : X \rightarrow \{0, 1\}$
   example (labeled instance): $\langle x, c(x) \rangle$; positive examples, neg. examples
2. **hypotheses $h : X \rightarrow \{0, 1\}$**
   hypotheses representation language
   hypotheses set H
   hypotheses consistent with the concept c: $h(x) = c(x), \forall$ example $\langle x, c(x) \rangle$
   version space
3. **learning = train + test**
   supervised learning (classification), unsupervised learning (clustering)
4. ***error$_h$* $= | \{x \in X, h(x) \neq c(x)\} |$**
   training error, test error
   accuracy, precision, recall
5. **validation set, development set**
   n-fold cross-validation, leave-one-out cross-validation
   overfitting

The Inductive Learning Assumption

Any hypothesis found to conveniently approximate the target function over a sufficiently large set of training examples

will also conveniently approximate the target function over other unobserved examples.

Inductive Bias

Consider

- **a concept learning algorithm** L
- **the instances** X**, and the target concept** c
- **the training examples** $D_c = \{\langle x, c(x) \rangle\}$**.**
- **Let** $L(x_i, D_c)$ **denote the classification assigned to the instance** x_i **by** L **after training on data** D_c**.**

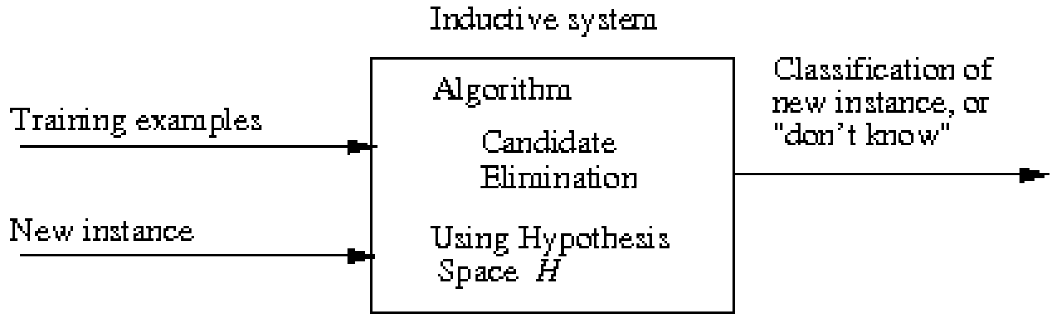
Definition:

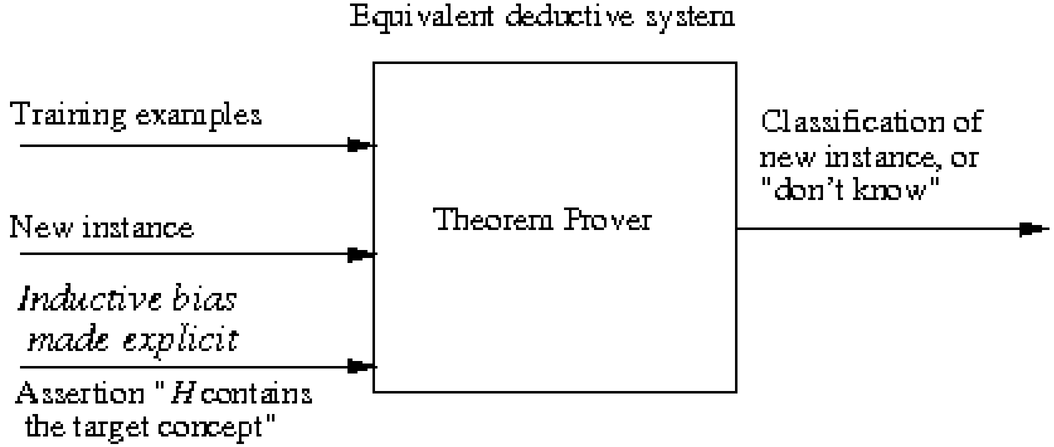
The inductive bias of L **is any minimal set of assertions** B **such that**

$$(\forall x_i \in X)[(B \vee D_c \vee x_i) \vdash L(x_i, D_c)]$$

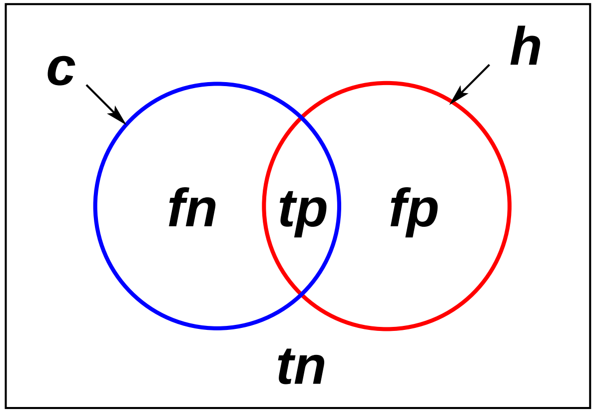
for any target concept c **and corresponding training examples** D_c**.**
($A \vdash B$ **means** A **logically entails** B**)**

Inductive systems can be modelled by equivalent deductive systems





Evaluation measures in Machine Learning



tp – true positives
fp – false positives
tn – true negatives
fn – false negatives

accuracy: $Acc = \frac{tp + tn}{tp + tn + fp + fn}$

precision: $P = \frac{tp}{tp + fp}$

recall (or: sensitivity): $R = \frac{tp}{tp + fn}$

F-measure: $F = \frac{2\ P \times R}{P+R}$

specificity: $Sp = \frac{tn}{tn + fp}$

follout: $= \frac{fp}{tn + fp}$

Mathew's Correlation Coefficient:

$$MCC = \frac{tp \times tn - fp \times fn}{\sqrt{(tp + fp)\times(tn + fn)\times(tp + fn)\times(tn + fp)}}$$

Lazy learning vs. eager learning algorithms

Eager: generalize before seeing query

- ◦ ID3, Backpropagation, Naive Bayes, Radial basis function networks, ...
- Must create global approximation

Lazy: wait for query before generalizing

- ◦ k-Nearest Neighbor, Locally weighted regression, Case based reasoning
- Can create many local approximations

Does it matter?

If they use the same hypothesis space H, lazy learners can represent **more complex functions.**

E.g., a lazy Backpropagation algorithm can learn a NN which is different for each query point, compared to the eager version of Backpropagation.

Who is Liviu Ciortuz?

- **Diploma (maths and CS) from UAIC, Iaşi, Romania, 1985**
  PhD in CS from Université de Lille, France, 1996
- **programmer:**
  Bacău, Romania (1985-1987)
- **full-time researcher:**
  Germany (DFKI, Saarbrücken, 1997-2001),
  UK (Univ. of York and Univ. of Aberystwyth, 2001-2003),
  France (INRIA, Rennes, 2012-2013)
- **assistant, lecturer and then associate professor:**
  Univ. of Iasi, Romania (1990-1997, 2003-2012, 2013-today)

ADDENDA

"...colleagues at the Computer Science department at Saarland University have a strong conviction, that nothing is as practical as a good theory."

Reinhard Wilhelm,
quoted by Cristian Calude,
in *The Human Face of Computing*,
Imperial College Press, 2016

“**Mathematics translates concepts into formalisms and applies those formalisms to derive insights that are usually NOT amenable to a LESS formal analysis.**”

Jürgen Jost,
***Mathematical Concepts*,**
Springer, 2015

"**Mathematics is a journey that must be shared, and by sharing our own journey with others, we, together, can change the world.**"

"**Through the power of mathematics, we can explore the uncertain, the counterintuitive, the invisible; we can reveal order and beauty, and at times transform theories into practical objects, things or solutions that you can feel, touch or use.**"

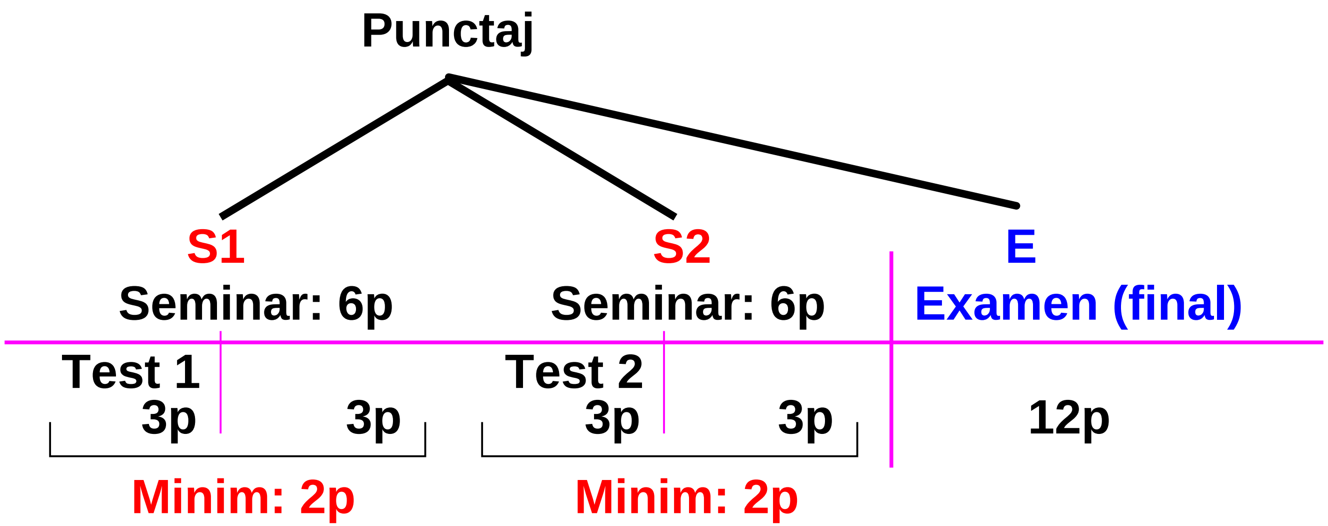
Cedric Villani,
winner of the Fields prize, 2010

cf. http://www.bbc.com/future/sponsored/story/20170216-inside-the-mind-of-a-mathematician, **15.03.2017**

ADMINISTRATIVIA

Grading standards for the ML undergraduate course 2019

Obiectiv: invatare pe tot parcursul semestrului!



Prezenta la seminar: obligatorie!
Penalizare: 0.1p pentru fiecare absenta de la a doua incolo

Nota = (6 + S1 + S2 + E) / 3
Pentru promovare: S1 + S2 + E >= 7.5

REGULI generale pentru cursul de Învăţare automată (cont.)

Sistemul de notare la licenţă

Nota = (6 + S1 + S2 + E) / 3,
unde

> **S1 = punctajul la seminar pe prima jumătate de semestru (0-6 puncte)**
> **S2 = punctajul la seminar pe a doua jumătate de semestru (0-6 puncte)**
> **E = punctajul la examenul din sesiune (0-12 puncte)**

Punctajele S1 si S2 se obţin (fiecare) ca suma a două punctaje, pentru
– test scris (anunţat în prealabil)
– răspunsuri "la tablă"

Condiţii de promovare:

S1 $\geq$ 2; S2 $\geq$ 2; nota $\geq$ 4.5
În consecinţă, punctajul minimal de îndeplinit din suma S1+S2+E este 7.5.

Atenţie:

S1 $<$ 2 (sau S2 $<$ 2) implică imediat nepromovarea acestui curs în anul universitar 2019-2020!

REGULI generale pentru cursul de Învăţare automată (cont.)

pentru cursul de la licenţă

- **Slide-uri de imprimat** (în această ordine şi, de preferat, COLOR):

http://profs.info.uaic.ro/~ciortuz/SLIDES/foundations.pdf

https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.ProbStat.pdf
[https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.EstimP.pdf]
[https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.Regression.pdf]
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.DT.pdf
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.Bayes.pdf
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.IBL.pdf
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.Cluster.pdf

(Atenţie: acest set de slide-uri poate fi actualizat pe parcursul semestrului!)

- De imprimat (ALB-NEGRU):

http://profs.info.uaic.ro/~ciortuz/SLIDES/ml0.pdf
http://profs.info.uaic.ro/~ciortuz/SLIDES/ml3.pdf
http://profs.info.uaic.ro/~ciortuz/SLIDES/ml6.pdf
http://profs.info.uaic.ro/~ciortuz/SLIDES/ml8.pdf
http://profs.info.uaic.ro/~ciortuz/SLIDES/cluster.pdf

REGULI generale pentru cursul de Învăţare automată (cont.)

pentru cursul de la master

- **Slide-uri de imprimat** (în această ordine şi, de preferat, COLOR):

http://profs.info.uaic.ro/~ciortuz/SLIDES/foundations.pdf

https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.ProbStat.pdf
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.EstimP.pdf
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.Regression.pdf
[https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.Cluster.pdf]
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.EM.pdf
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/SLIDES/ML.ex-book.SLIDES.SVM.pdf

(Atenţie: acest set de slide-uri poate fi actualizat pe parcursul semestrului!)

- De imprimat (ALB-NEGRU):

http://profs.info.uaic.ro/~ciortuz/SLIDES/svm.pdf

- De imprimat opţional (ALB-NEGRU):

Companion-ul practic pentru culegerea ,,Exerciţii de învăţare automată“:
https://profs.info.uaic.ro/~ciortuz/ML.ex-book/implementation-exercises/ML.ex-book.Companion.pdf

REGULI generale pentru cursul de Învăţare automată (cont.)

Observaţie (1)

Este recomandabil ca la fiecare curs şi seminar, studenţii să vină cu cartea de exerciţii şi probleme (de L. Ciortuz et al) şi cu o fasciculă conţinând slide-urile imprimate.

Observaţie (2)

Profesorul responsabil pentru acest curs, Liviu Ciortuz NU va răspunde la email-uri care pun întrebări pentru care raspunsul a fost deja dat

– fie în aceste slide-uri,
– fie la curs