Introduction to the AA2 Course

Davide Bacciu

Dipartimento di Informatica
Università di Pisa
bacciu@di.unipi.it

Machine Learning: Neural Networks and Advanced Models (AA2)

Objectives

Train machine learning (ML) specialists capable of

- designing novel learning models
- developing advanced applications using ML solutions

Address complex data domains

- Noisy, hard-to-interpret, semantically rich information (natural language, images, videos)
- Non-vectorial relational information (sequences, trees, graphs)

Expected Outcome

Students completing the course are expected to

- Gain in-depth knowledge of advanced machine learning topics
- Understand their theory and applications
- Be able to individually read, understand and discuss research works in the field

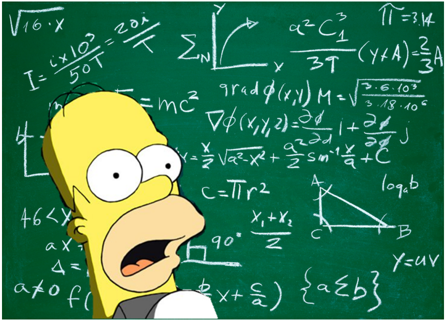
The course is targeted at

- Students specializing in
  - Machine learning and computational intelligence
  - Data mining and information retrieval
  - Robotics
  - Bioinformatics, ...
- Students seeking machine learning theses

Prerequisites

- Knowledge of machine learning fundamentals
  - Having taken the AA1 course..
  - ...or discuss your ML skills with me
- It will be of great help if you have knowledge of
  - Algebra and calculus
  - Probability and statistics

...and, above all, a disposition not to get easily scared by math!



Organization

The course is articulated in four parts

- Recurrent/recursive neural networks
- Probabilistic Learning and Graphical Models
- Kernel methods
- Advanced Applications

Introduce learning models with an incremental approach: from sequential data processing to complex structured domains

Guest seminars by researchers and (possibly) companies

- Alessio Micheli (@di.unipi)
- Claudio Gallicchio (@di.unipi)
- Alexander Schulz (@uni-bielefeld)

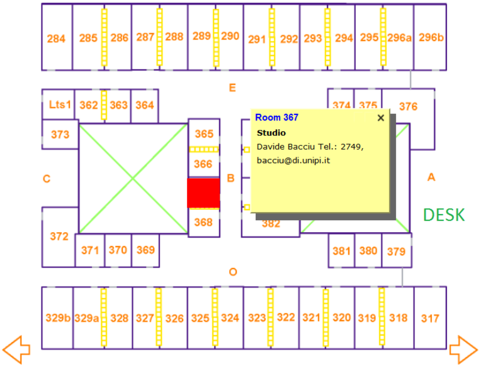
Topics

- Recurrent neural networks
  - Reservoir computing
- Probabilistic Learning and Graphical Models
  - Hidden Markov models
  - Markov random fields
  - Latent variable models
- Non-parametric and kernel-based methods
  - Unsupervised learning for complex data
- Learning in structured domains (sequences, trees and graphs)
- Emerging topics and applications in machine learning
  - Deep learning, machine vision, ChemInformatics, BioInformatics, AAL

Course Instructor

Davide Bacciu

Email - bacciu@di.unipi.it
Tel - 050 2212749
Office - Room 367, Dipartimento di Informatica
Office hours - Tuesday 17-19 (email me!)



Course Schedule

Weekly Timetable:

Day	Time	Room
Monday	16-18	C1
Thursday	16-18	C1

Talk now if you need to change course weekly schedule!

Course comprises 24 lectures

- Not enough dates in the academic calendar
- Will need to accommodate some (2,3) extra dates

Course Homepage

Reference Webpage on Didawiki:

http://didawiki.cli.di.unipi.it/doku.php/magistraleinformatica/aa2/start

Here you can find

- Course information
- Lecture slides
- Articles and course materials

You can subscribe to get RSS feeds on page updates

Reference Books

No official textbook

A standard neural networks reference book:

Simon O. Haykin, *Neural Networks and Learning Machines*, Pearson (2008)

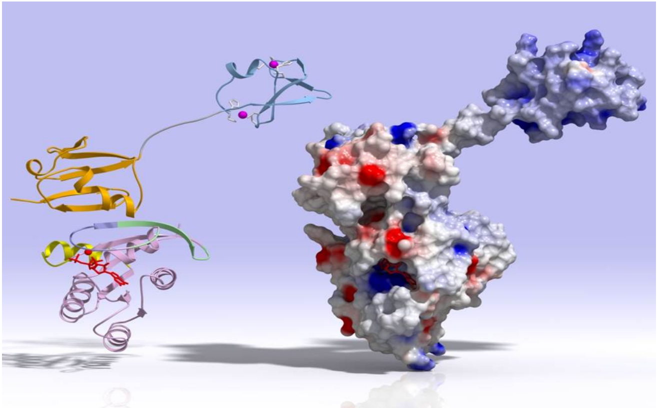
Probabilistic learning (free pdf, with code):

David Barber, *Bayesian Reasoning and Machine Learning*, Cambridge University Press (2012)

Inference and learning (free pdf):

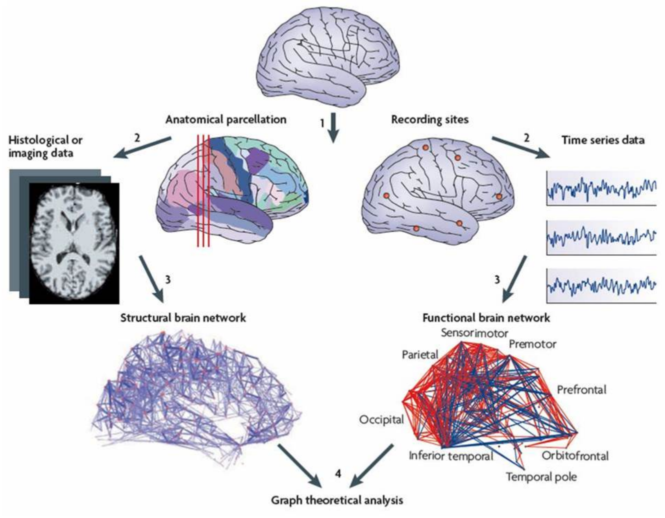
David J.C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press (2003)

Complex Data

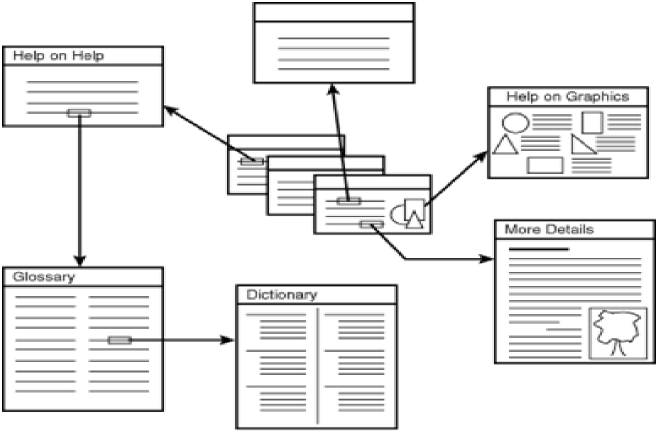


BioChemistry: protein sequences and molecular graphs

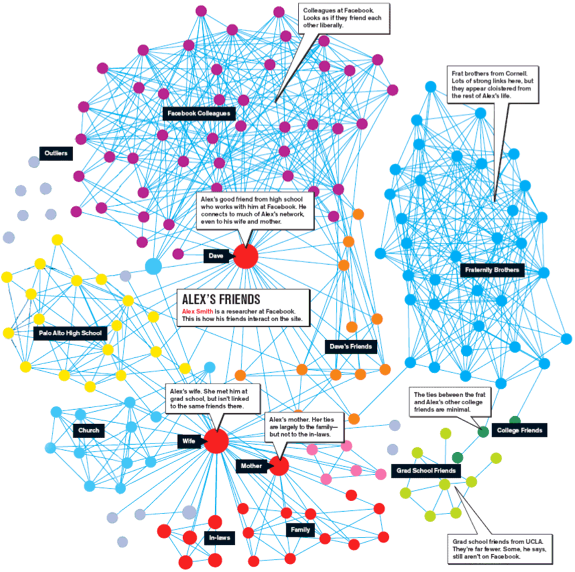
High-throughput Bio-Medical Data



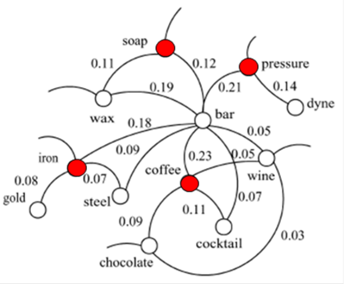
Structured Documents and Relational Data

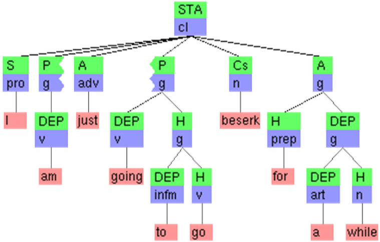


Social Data

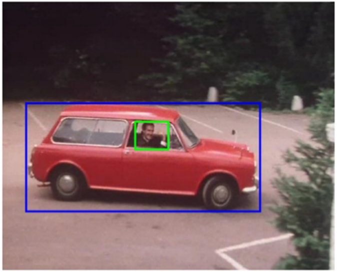

Colleagues at Facebook. Looks as if they friend each other liberally.
Frat brothers from Cornell. Lots of strong links here, but they appear cloistered from the rest of Alex's life.
Facebook Colleagues
Outliers
Alex's good friend from high school who works with him at Facebook. He connects to much of Alex's network, even to his wife and mother.
Dave
ALEX'S FRIENDS
Alex Smith is a researcher at Facebook. This is how his friends interact on the site.
Fraternity Brothers
Palo Alto High School
Dave's Friends
Alex's wife. She met him at grad school, but isn't linked to the same friends there.
The ties between the frat and Alex's other college friends are minimal.
Church
Alex's mother. Her ties are largely to the family—but not to the in-laws.
College Friends
Wife
Mother
Grad School Friends
Family
In-laws
Grad school friends from UCLA. They're far fewer. Some, he says, still aren't on Facebook.

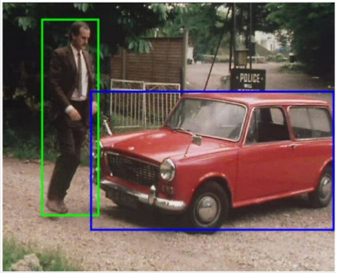
Natural Language





Machine Vision





Understanding if a man is inside/outside the car impacts scene interpretation

Complex data

- Addressing complex tasks in challenging applications
  - Need to design novel machine learning models
  - Noisy, high-dimensional, semantically-rich data
- Information is no longer a single atomic piece
  - Need to take context into account
  - Relational data

The Course in One Word

The secret word is structures

Structured Data

- Compound information representing the relationships between its composing elements
- Different degrees of complexity and expressivity (classes of structures)

Structure in Learning Models

- Represent the model components and the information they share (context)
- Change model expressivity and the complexity of inference and learning
- May follow the structure of data and can be dynamical

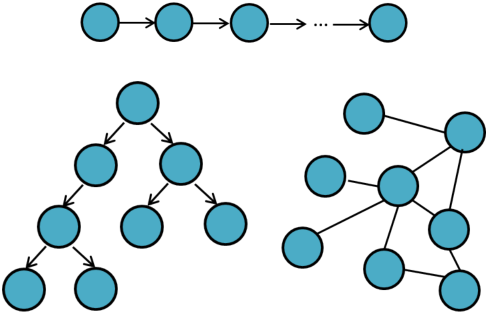
Data (Flat)

- Data is the set of available observations (facts) of the process we want to model
  - Need an appropriate representation to capture relevant information
  - Data structure defines its representation capabilities
- Flat data
  - Population: Fixed-size vectors of features
  - Represent a set of relevant object properties
  - Patient records, DNA micro-arrays, images, ...

Age	Sex	Weight	Smoke

Data (Structured)

- Structured data
  - Population: a set of structures of variable dimension
  - Represent features as well as relationships between them
  - Sequences, trees, graphs



Examples: DNA sequences, language sentences, image segmentations, molecules, ...

Learning Tasks (Flat)

Approximate general functions f from observations $(x, f(x))$ where x and $f(x)$ are vectors

- Face recognition
  - $x \rightarrow$ face image (descriptor)
  - $f(x) \rightarrow$ person identifier
- Handwriting recognition
  - $x \rightarrow$ character image (descriptor)
  - $f(x) \rightarrow$ character/digit
- Patient diagnosis
  - $x \rightarrow$ patient record
  - $f(x) \rightarrow$ disease classification
- Spam detection
  - $x \rightarrow$ email bag of words
  - $f(x) \rightarrow$ spam/no-spam classification

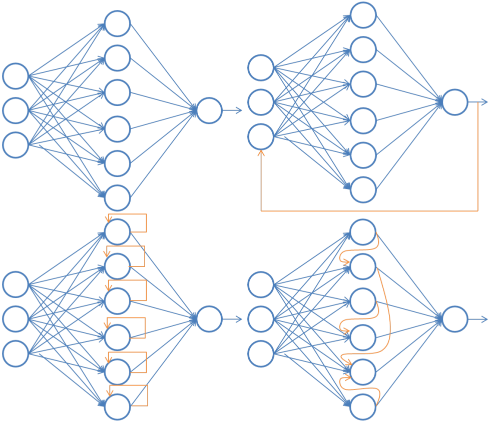
Learning Tasks (Structured)

Approximate general trasductions f from observations $(x, f(x))$ where at least x is structured data

- Protein folding
  - $x \rightarrow$ amino-acids sequence
  - $f(x) \rightarrow$ sequence of atoms' 3D coordinates
- Machine translation
  - $x \rightarrow$ English sentence parse tree
  - $f(x) \rightarrow$ Italian sentence parse tree
- Molecule characterization
  - $x \rightarrow$ molecular graph
  - $f(x) \rightarrow$ toxicity prediction

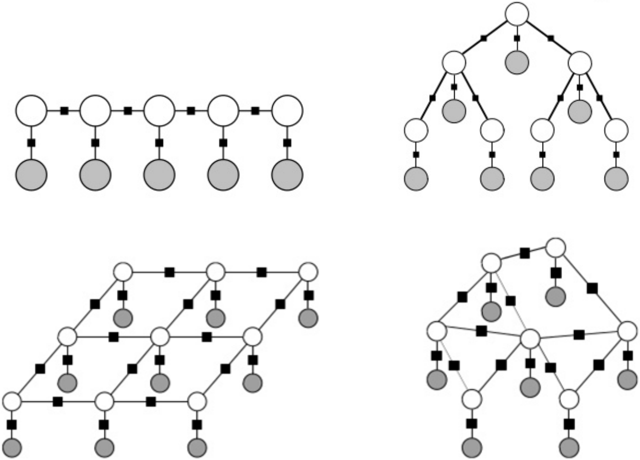
Learning Model Structure (I)

Study how varying the structure of a learning model influences its computational capabilities and complexity



Learning Model Structure (II)

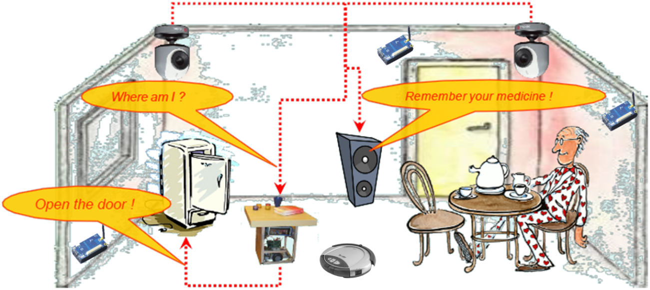
In some cases the learning model structure conforms with the data structure



Ambient Assisted Living

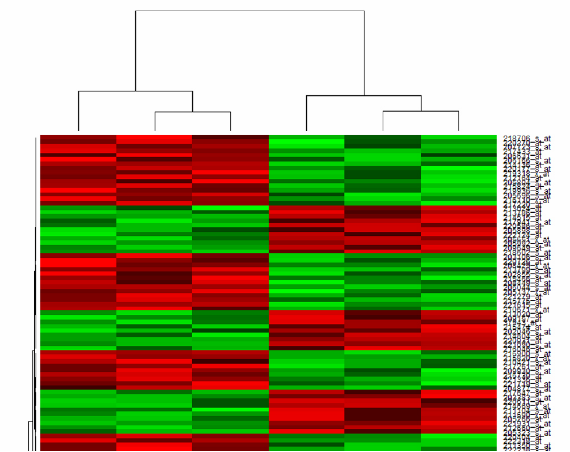
Learning models for the analysis of sensor timeseries

- EU FP7 RUBICON - Cognitive Robotic Ecologies
- EU FP7 DOREMI - Smart environment for empowering elderly lifestyle



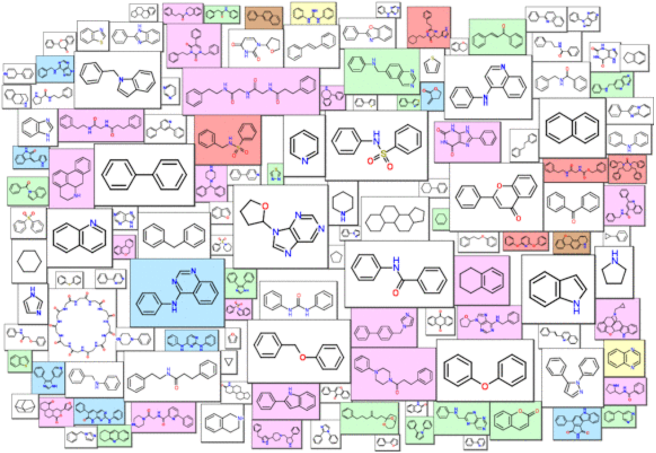
Machine Learning for BioInformatics

- Exploratory analysis of bio-medical data
- Predictive models for personalized medicine



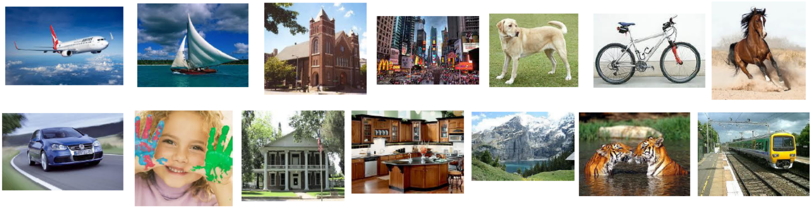
Machine Learning for ChemInformatics

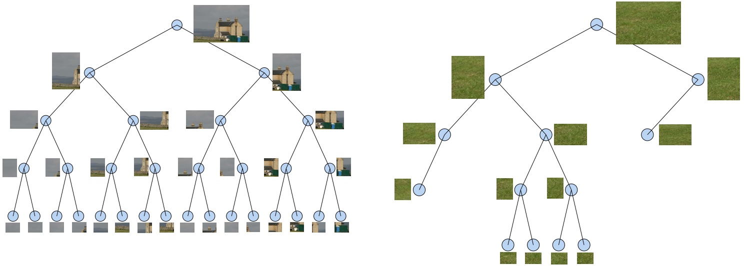
Learning to predict functional properties from chemical structures



Machine Vision

Learning structured image representations





Machine Learning for Ethology

Pattern discovery from sensor data for endangered species preservation

Exams

The (quick) way (there is no such thing as an easy way)

Midterm Assign. - A short presentation on a work from a reading list to be performed at midterm

Final Project - A written report on a topic of interest for the course (and for you)

Oral Presentation - A presentation of the final project

The alternative way (for working students, those not attending classes and those who fail the other way)

Final Project - A written report on a topic of interest for the course, as above

Oral Exam - A presentation of the final project plus examination on the course program

Midterm Assignment

- Pick one article from a reading list and prepare a short presentation (20 minutes) to be given in front of the class
- The presenter should
  - Answer reading-list's questions on the article
  - Include a mathematical derivation of a theoretical result or learning algorithm in the paper
  - Be able to answer my (and your collegues') questions on the presentation
- Timeline
  - Reading list published mid-march
  - Article bidding by end of march
  - Presentation during midterm exams (tentative date: 13/04/2015)

Final Project (I)

- Choose from a set of suggested topics or propose your own topic of interest and prepare a report (6-10 pages)
- Timeline (quick way)
  - Suggested topics list: early-may
  - Choose project: strictly before the last lecture (late may)
  - Report delivery: 7 days before the oral (strict)
- Timeline (alternative way)
  - Choose project: email me to arrange a topic
  - Report delivery: By the standard exam date (appello) (strict)

Final Project (II)

- Possible project types

  Survey Read at least three relevant and distinct papers on a topic and write a report: not a simple summary, rather try to find connections between the works and highlight interesting open problems

  Original Propose a research project and develop it (with my help) as much as possible: must have a substantial innovative component

  Software Develop a well-written, tested and commented software (with doc) implementing a non-trivial learning model and/or an application relevant for the course

Oral Presentation (Quick Way)

- Prepare a presentation (40 minutes) on the final project and discuss it in front of me (and anybody interested)
- The presenter should
  - Summarize the ideas in the report
  - Be able to answer my questions on the presentation
  - Be confident on the course topics: no mathematical derivations asked, but questions on the concepts seen in the course may pop-up
- Timeline
  - Arrange a presentation date with (10 days in advance)
  - Handle your report at least 7 days before presentation
  - Send me your slides the day before the presentation

Oral Exam (Alternative Way)

- Prepare a presentation (40 minutes) on the final project following the same rules and timeline of the quick way
- In addition, candidates will be subject to an oral exam with questions covering the course contents
  - Mathematical derivations and details of the models discussed during the lectures will be asked
  - Examination will be performed jointly with the presentation

How to get past this course?

Grading (quick way) $\frac{(G_P + G_O)}{2} + G_M$

- Midterm: $G_M \in [0, 3]$ (0 means failure)
- Project: $G_P \in [0, 30]$ ($<$ 18 means failure)
- Oral: $G_O \in [0, 30]$ ($<$ 18 means failure)

Grading (alternative way) $\frac{(G_P + G_O)}{2}$

- Project: $G_P \in [0, 30]$ ($<$ 18 means failure)
- Oral: $G_O \in [0, 32]$ ($<$ 18 means failure)

Upcoming..

Recurrent Neural Networks Module

Extend connectionist paradigm with a new class of neural models capable of representing the history of the input signals in its internal state

Topics

- Introduction to Recurrent Neural Networks
- Processing history/context dependant input signals
- Learning in Recurrent Neural Networks
- Recurrent/recursive neural networks for structures
- Reservoir Computing for sequences

Next Lecture

Introduction to Recurrent Neural Networks

- Neural networks for history/context dependant task
- Representing time in neural networks: explicit and implicit approach
- Elman networks
- Application examples

Guest lecture by Alessio Micheli

Contacts and Information

Course Didawiki

`http://didawiki.cli.di.unipi.it/doku.php/magistraleinformatica/aa2/start`

Before leaving please write down your email address for the course mailing list

Questions?