CMP-6002B - Machine Learning

Gavin Cawley

Lecture 1 - Introduction

What is Artificial Intelligence?

- "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning..." [Bellman 1978].
- Symbolic reasoning:
  - GOFAI - *"Good Old Fashioned Artificial Intelligence"*.
- Non-symbolic reasoning:
  - Connectionist approaches (artificial neural networks).
  - Statistical approaches.
  - Swarm intelligence.
  - Robotics.
- But do we want systems that behave like humans or systems that behave ideally?

What is Machine Learning?

- Machine learning is a branch of artificial intelligence.
- Concerned with algorithms for inferring a model of some physical system from observed data.
  Machine Learning is the study of computer algorithms that improve automatically through experience – Tom Mitchell.
- Very significant overlap with statistics.
- ML is often more concerned with algorithmic complexity:
  - Dealing with large datasets.
  - Forming tractable algorithms for approximate inference.
  - Non-parametric methods - large numbers of parameters.

Why Study Machine Learning

- Many industries have invested heavily in machine learning research and development, e.g. Google, IBM, Microsoft.
- Can be used to address many interesting problems in CS:
  - Machine vision.
  - Natural language processing.
  - Information retrieval.
  - Speech/face/biometric/gesture recognition.
- Good approach to many interesting scientific problems:
  - Bio/Chemoinformatics - inferring gene regulatory networks.
  - Modeling the impacts of climate change.
- Machine Learning is timely because:
  - Deluge of machine readable data.
  - Advances in computing power.
  - Rapid advances in computational theory.
- Machine learning can be fun (see NERO demo later)!

Supervised and Unsupervised Learning

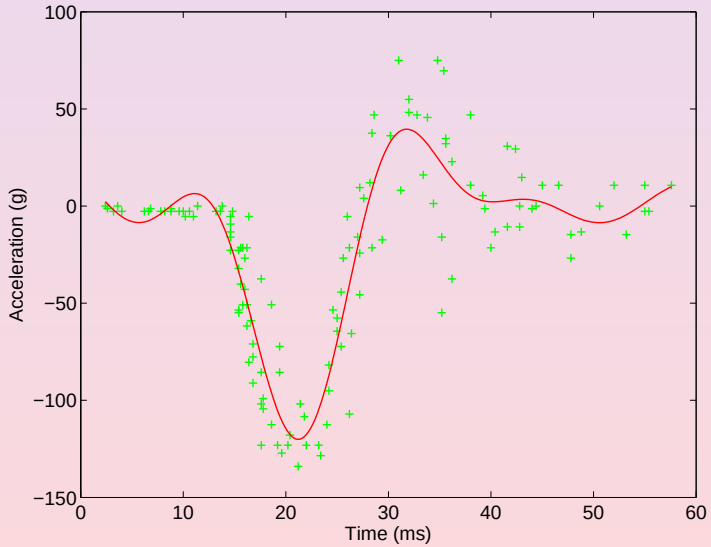
- Supervised learning is the most common form of ML.
  - We are provided with a set of labeled training data.
  - Each pattern described by attributes and response variables.
  - Attributes - descriptive features of each object.
  - Response - the desired output.
  - The aim is to predict the response as a function of the attributes.
  - Examples: regression, classification.
- Unsupervised learning:
  - We are given attributes for each object, but no response variable.
  - The aim is to reveal hidden structure of the data
  - Examples: clustering, visualization, dimensionality reduction.

Types of Machine Learning: Regression

- Model the relationship between a real valued dependent variable (response) and a set of explanatory variables
- Response is a noisy measurement of a deterministic function,

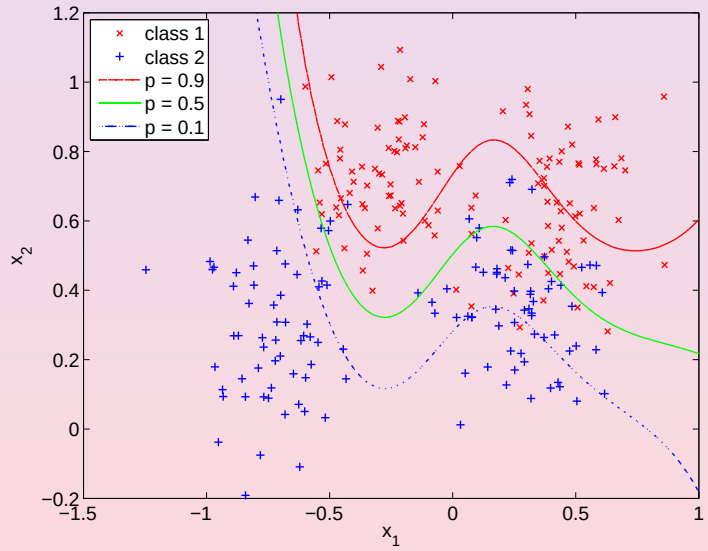
$$y = f(\boldsymbol{x}) + \epsilon, \qquad \epsilon \sim \mathcal{N}(0, \sigma^2).$$

- Linear (least-squares) regression (Legendre 1805, Gauss 1806).
- Estimating planetary orbits.
- Testing motorcycle helmets.
- Stock market forecasting.
- Estimating patient survival times following treatment.
- Predicting weld strength.



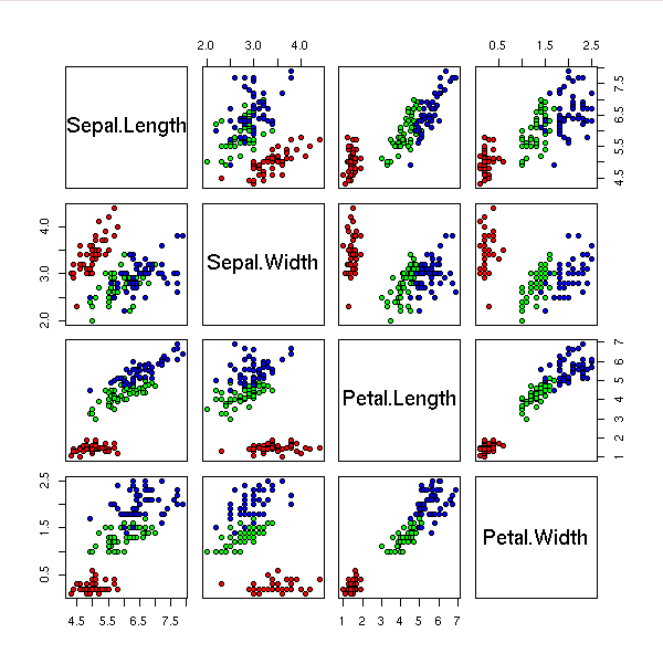
Types of Machine Learning: Classification

- Predict whether an object belongs to one of a number of pre-defined classes based on a set of attributes.
- LDA - linear discriminant analysis (Fisher 1936).
- Main focus of the first part of the unit:
  - Nearest neighbour, Naïve Bayes, decision trees, artificial neural networks, support vector machine.
- Computer aided diagnosis.
- Optical character recognition.
- Credit card fraud detection.
- Chemoinformatics.
- Machine vision.
- Spam filtering.



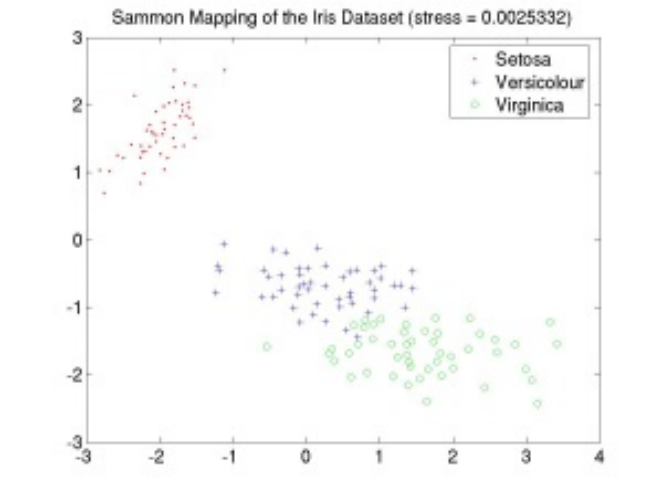
Types of Machine Learning: Clustering

- Partition the data into groups (clusters) so that the data in each cluster share some common trait.
- Generally *unsupervised* learning - desired output not specified.
- Normally assumes some kind of similarity metric.
- Identify subsets of data generated by different processes.
- Clustering may be flat or hierarchical.
- Image texture segmentation.
- Identifying sets of genes with similar patterns of expression.
- Marketing - identify different types of consumers.
- Identifying musical genres.



Types of Machine Learning: Dimensionality Reduction

- Generate a low dimensional representation of a high dimensional dataset that maximally preserves the underlying structure.
  - Most often used for visualization.
  - Pre-processing step for other machine learning algorithms.
- Feature selection - eliminate redundant attributes.
- Feature extraction - combine attributes to make new features.
- PCA (Pearson 1901).
- Multidimensional scaling.
- Artificial neural networks.
- Kernel principal component analysis (KPCA).
- Sammon's mapping.



Example: Insurance Risk

- Insurance companies use *actuaries* to predict whether or not someone will claim on an insurance policy.
  - Will this person claim or not (classification).
  - What premium should we set (regression)?
- Some machine learning could be used:
  - Insurance companies have a huge amount of historical data about their customers.
  - For each customer they have information about age, sex, postcode etc.
  - They also have a record of your previous claims.
- Could also cluster the data to find groups of similar customers
  - Targeted marketing of new products.
- Detection of credit card fraud is very similar.

Other Types of Machine Learning

- **Semi-supervised learning** - training data contains both labelled and unlabelled data.
- **Transduction** - generate predictions for a specific test set, rather than a model of the underlying process.
- **Data Mining** - where we want to extract interesting relationships from the data (but we don't necessarily know before hand what we are looking for).
- **Reinforcement Learning** - an *agent* learns to respond to stimuli through feedback received from the environment (which is often deferred).
  - Often used in robotics and economics.
  - Also used a lot in AI for computer games.

Example : Neuro Evolving Robotic Operators

- Neuro evolving robotic operators is a game-playing environment for research in a type of reinforcement learning.
- Armies of robotic soldiers learn to fight each other effectively.
- Each agent is controlled by a neural network.
- Evolution-based reinforcement learning
  - Successful agents pass on "memes" to those that follow
- "Sandbox" training phase followed by "battle" phase.
- Interesting behaviours emerge without explicit instructions.
- Download it and have a play! `nn.cs.utexas.edu/nero/`
- Lets watch the demo.

Lecture Plan

- Week 1 - Introduction and Basic Principles (GCC)
- Week 2 - k-Nearest Neighbour classifier (GCC)
- Week 3 - Linear classifiers (GCC)
- Week 4 - Neural Networks and SVMs (GCC)
- Week 5 - Decision Trees (AJB)
- Week 6 - Do different week
- Week 7 - Ensemble Classifiers (AJB)
- Week 8 - Evaluating Classifiers (AJB)
- Week 9 - Clustering (JL)
- Week 10 - Reinforcement Learning (JL)
- Week 11 - Genetic Algorithms/Learning Classifier Systems (JL)

Learning Resources

- Lectures:
  - Lecture - Monday 11am – 1pm (week 1-5, 7-12)
  - Lecture/Seminar - Tuesday 11am-12pm (week 1-5, 7-12)
- Seminar/Laboratory session:
  - Monday 1–3pm & 3pm–5pm (1-11) PC Labs 1&2
- Blackboard - all teaching materials.
- Lecturer (first part) :
  - Room BIO 2.13
  - Email: `gcc@cmp.uea.ac.uk`
- On-line resources
  - `scholar.google.com` - search for scientific papers on ML.
  - `machinelearning.org` - machine learning society.
  - `http://mlearn.ics.uci.edu/MLRepository.html`

Focus of the Course

- Students encouraged not to blindly use commercial machine learning or data mining packages for practical applications.
- Students encouraged to focus on gaining an understanding of the underlying principles of machine learning.
- Students should be able to select the appropriate machine learning method based on the nature of the application.
- Students should gain an awareness of the many pitfalls of Machine Learning in practical applications.
- Students should gain practical experience applying machine learning methods to real world problems.
- Students should gain experience implementing practical machine learning algorithms.
- Students should have fun.

Summary

- In this lecture:
  - What is machine learning.
  - Why machine learning is interesting.
  - Types of machine learning.
  - Introduction to the unit.
- In the next lecture:
  - Example of statistical pattern recognition.
  - Parametric and non-parametric modeling.
  - Machine learning and optimization.
  - Maximum likelihood.
  - Value of probabilistic modeling.
  - Generalization and over-fitting.

PhD Project Available with Prof. Bagnall

www.FindAPhD.com

Anomaly Detection in Time Series (BAGNALLAU20SCIEP) at University of East Anglia on FindAPhD.com

PhD Project - Anomaly Detection in Time Series (BAGNALLAU20SCIEP) at University of East Anglia, listed on FindAPhD.com (105 kB)

https://tinyurl.com/upbo2vj