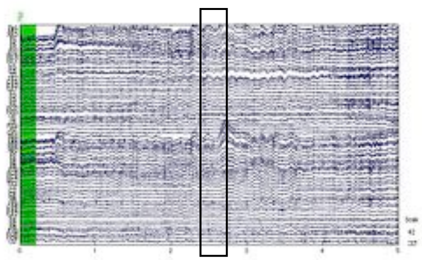
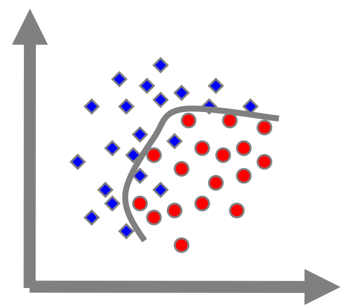
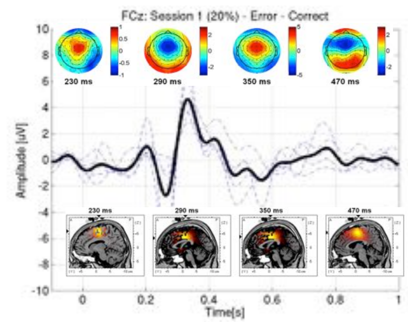






Data analysis and model classification

Introduction - Unsupervised Learning

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Ricardo Chavarriaga <ricardo.chavarriaga@epfl.ch)
José del R. Millán <jose.millan@epfl.ch>

CNBI - Defitech Chair in Brain-Machine Interface
Ecole Polytechnique Fédérale de Lausanne

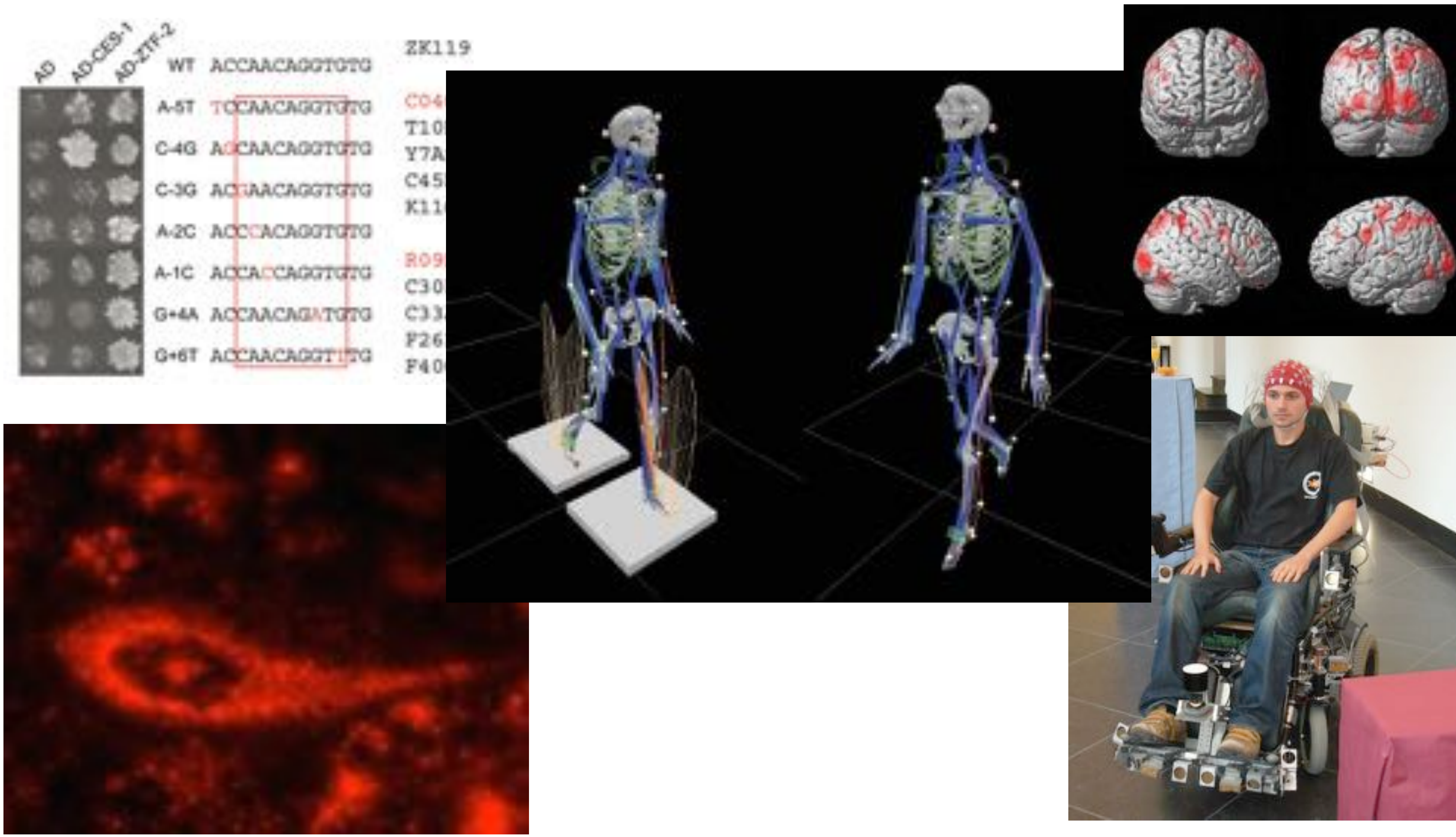
Contact: damc@listes.epfl.ch
Office hours: By appointment

Lecture: AAC137 Mon 8h15-10h00
Exercises: INF1 Mon 17h15-19h00

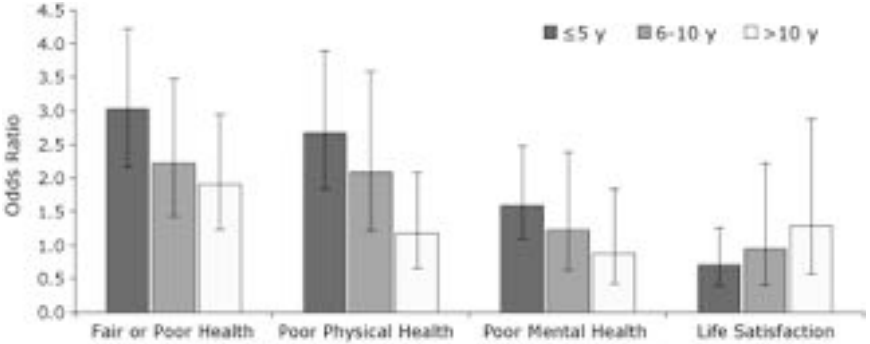
Note: For the practical work you're supposed to use your own laptop

Motivation

- Current technology allows for the acquisition of large amounts of data

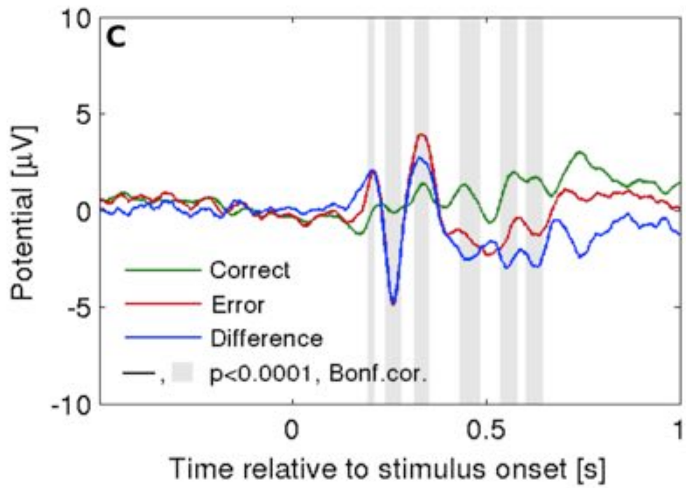


Motivation



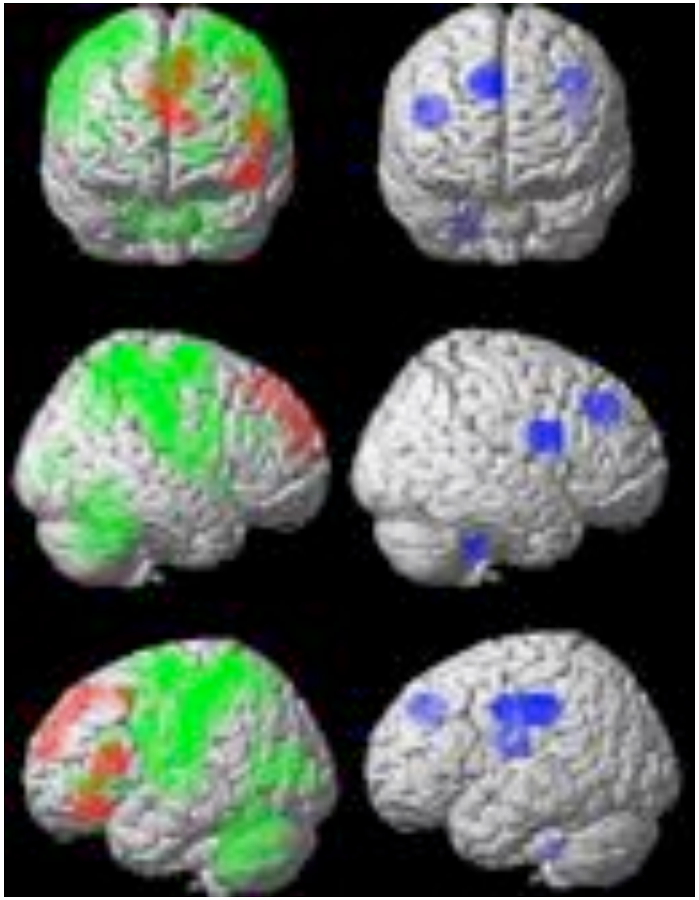
Health status and quality of life among cancer survivors, by the time since diagnosis.

Fairley et al., 2010



EEG activity (FCz) electrode after correct and erroneous feedback. Gray bars show periods where differences are statistically significant.

Chavarriaga et al., 2014

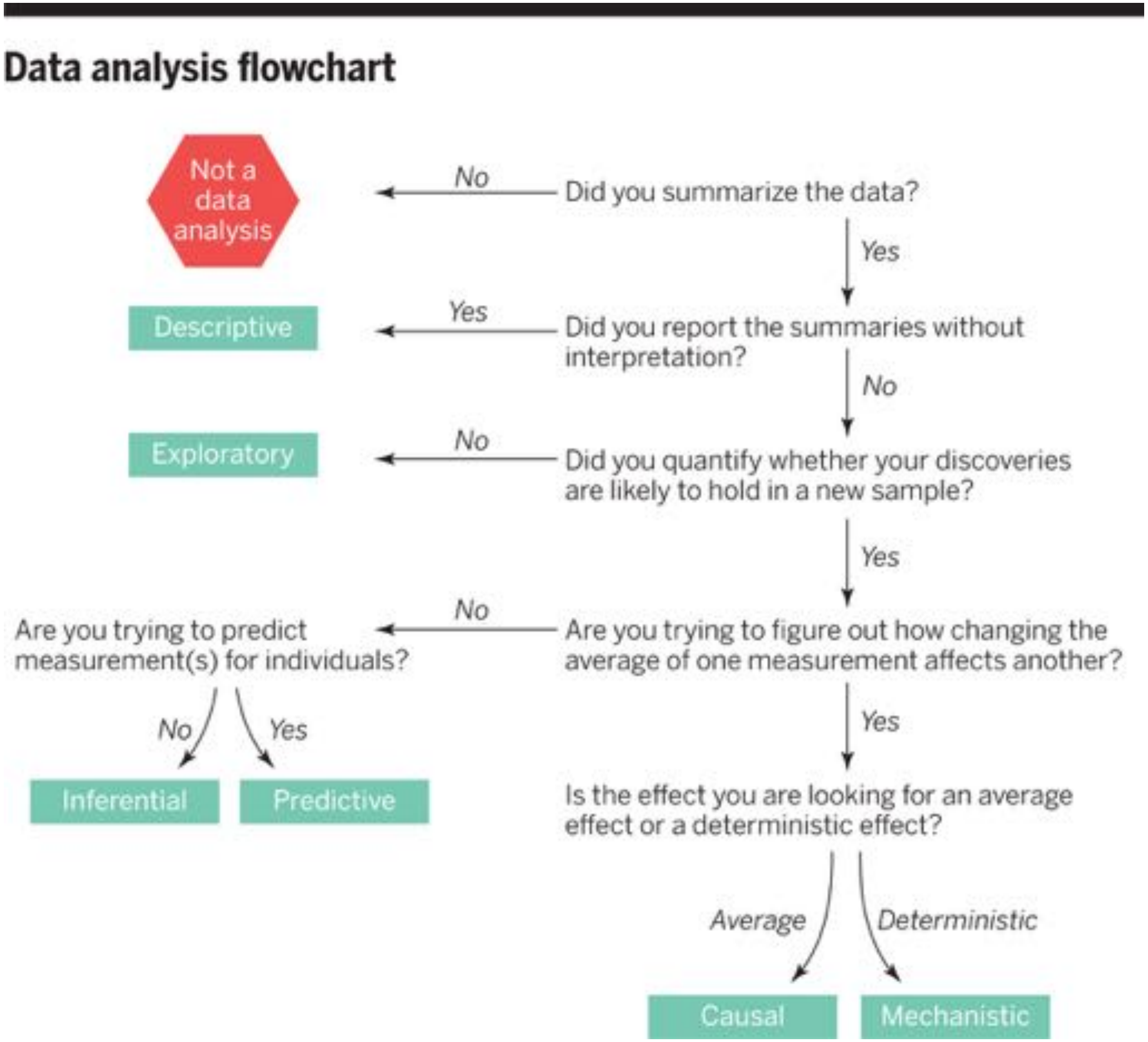


Left: Different 3D views of regions or relatively higher activity during truth-telling (green) and lying (red). Right: Areas in which functional activity was found to be most informative in terms of pattern classification.

Davatzikos et al., 2005


BRACE YOURSELF
DATA IS COMING
memegenerator.net

Data analysis: what for?



Leek, J. T. & Peng, R. D. Science, 2015

Machine learning

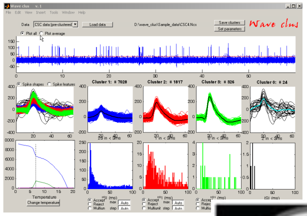
“Algorithms capable of learning to improve their performance of a task on the basis of the previous experience (empirical data)”

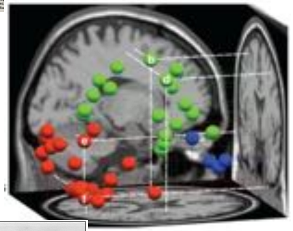
“A learner can take advantage of examples (data) to capture characteristics of interest of their unknown underlying probability distribution”

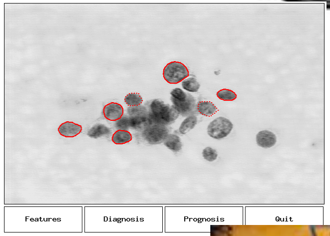
Related fields include: probability theory and statistics, data mining, pattern recognition

This course is NOT about…

- Brain-computer interfaces or neuroprothetics
- Robotics
- Bioinformatics
- Other specific applications
- Signal processing
- Programming or matlab skills








There will be work…

Course structure

- Grading
  - Written exam (2/3)
  - Exercises (1/3)

- Practical work. Teaching assistants
  - Ruslan Aydharkhanov
  - Bastien Orset
  - Julien Rechenmann

- No textbook
  - Lecture notes
  - Suggested reading material be provided
  - Moodle

- Contact: damc@listes.epfl.ch

Practical work

Three mini-projects (3 people/group)

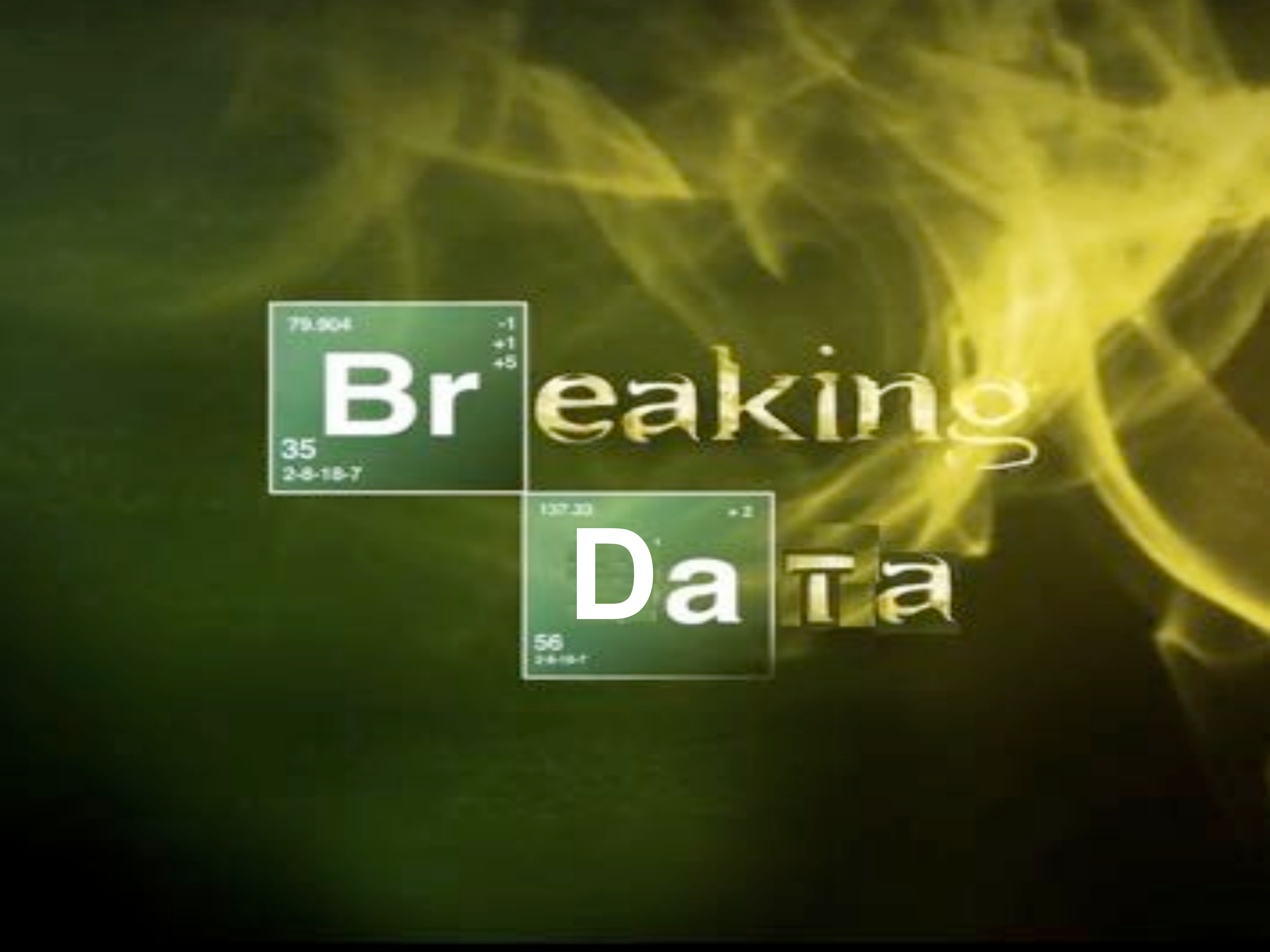
1. Unsupervised learning and model validation (25% grade). Report: 5th week
2. Supervised classification (50% grade). Report: 10th week
3. Regression (25%) 13th week

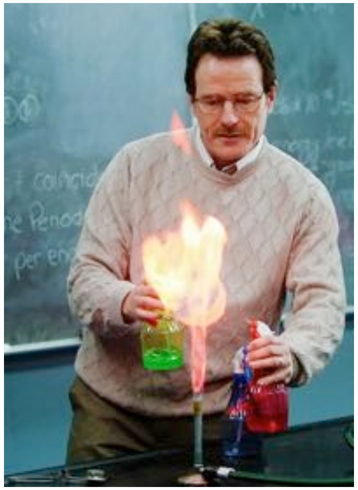
TA sessions provide support to develop the projects (guidelines provided every week). Written report (max 12 pages) handed in **at the end of each project**

Groups members should be submitted on the **2nd October**

We encourage the use of the **forum in moodle** to find solutions to your questions. When posting to the forum, mind to make clear, well-explained questions.

<u>TAs won't reply to questions before two working days after post to allow students to post their answers or hypothesis.</u>


79.904
-1
+1
+5
Breaking
35
2-8-18-7
137.33
+2
Data
56



Mr. White

J. Pinkman

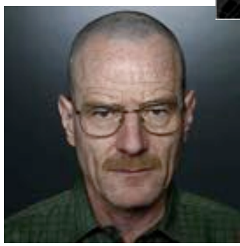
Heisenberg

Mr. White

J. Pinkman

Heisenberg

Facial hair

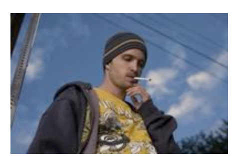




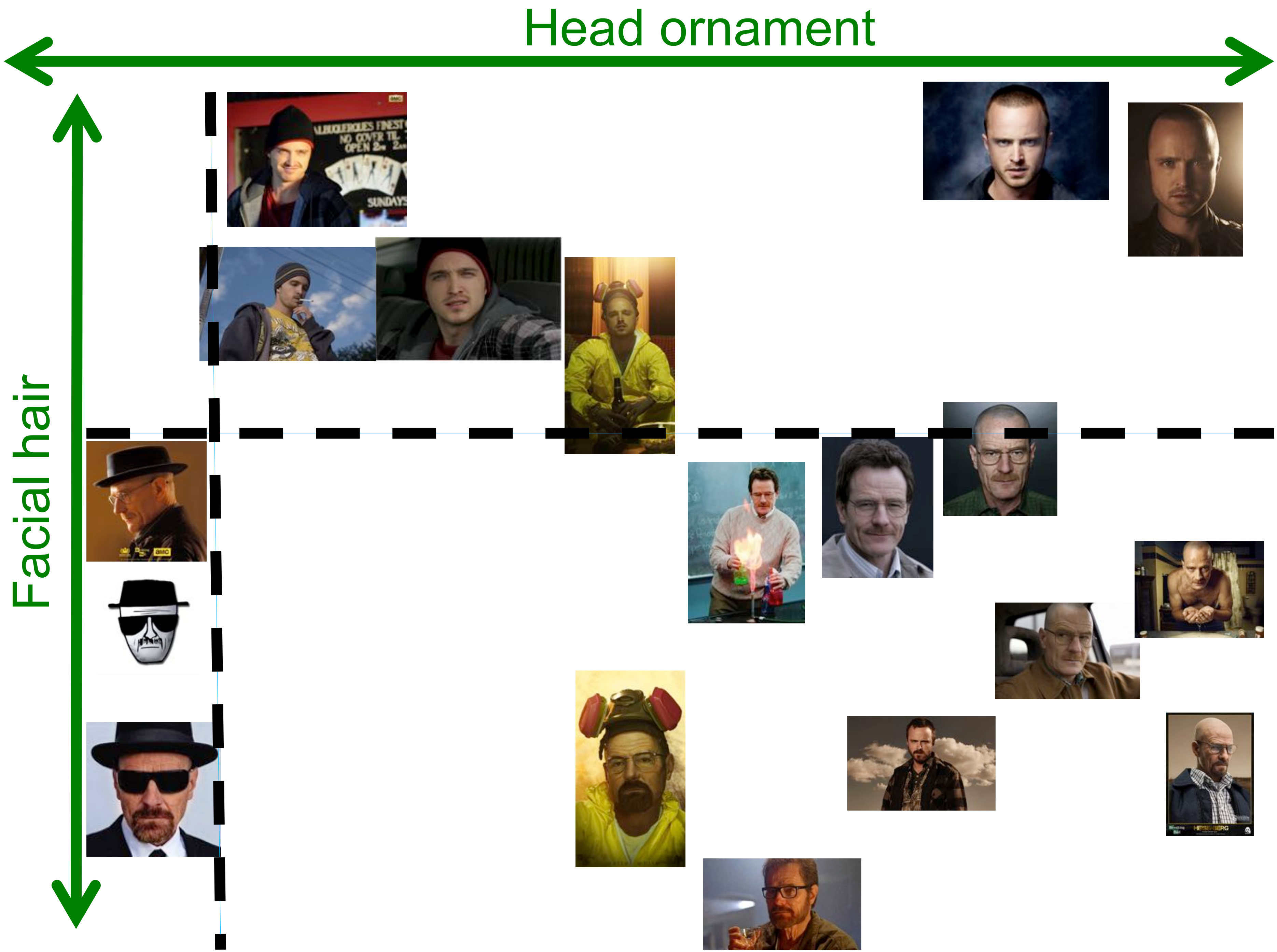


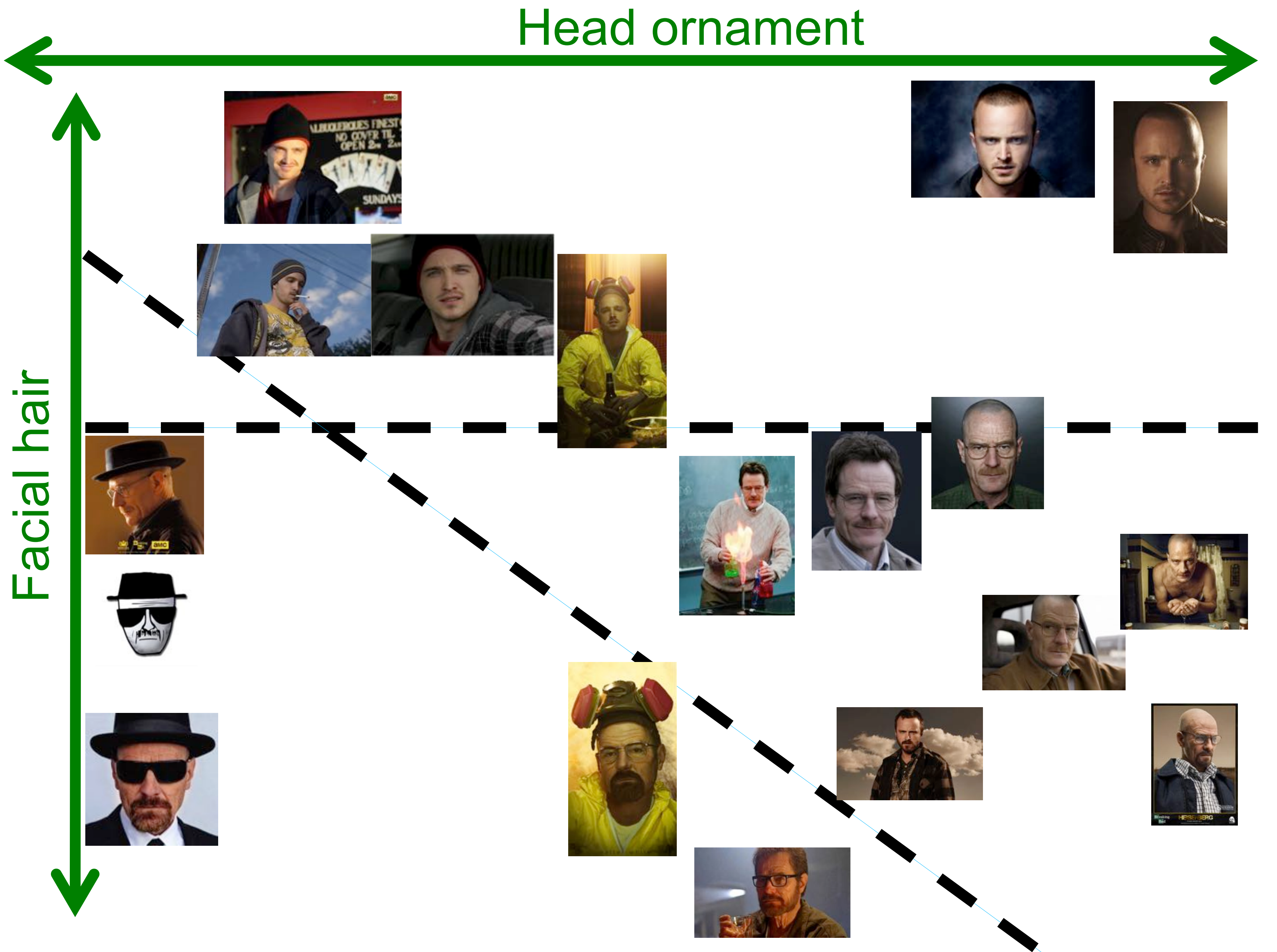


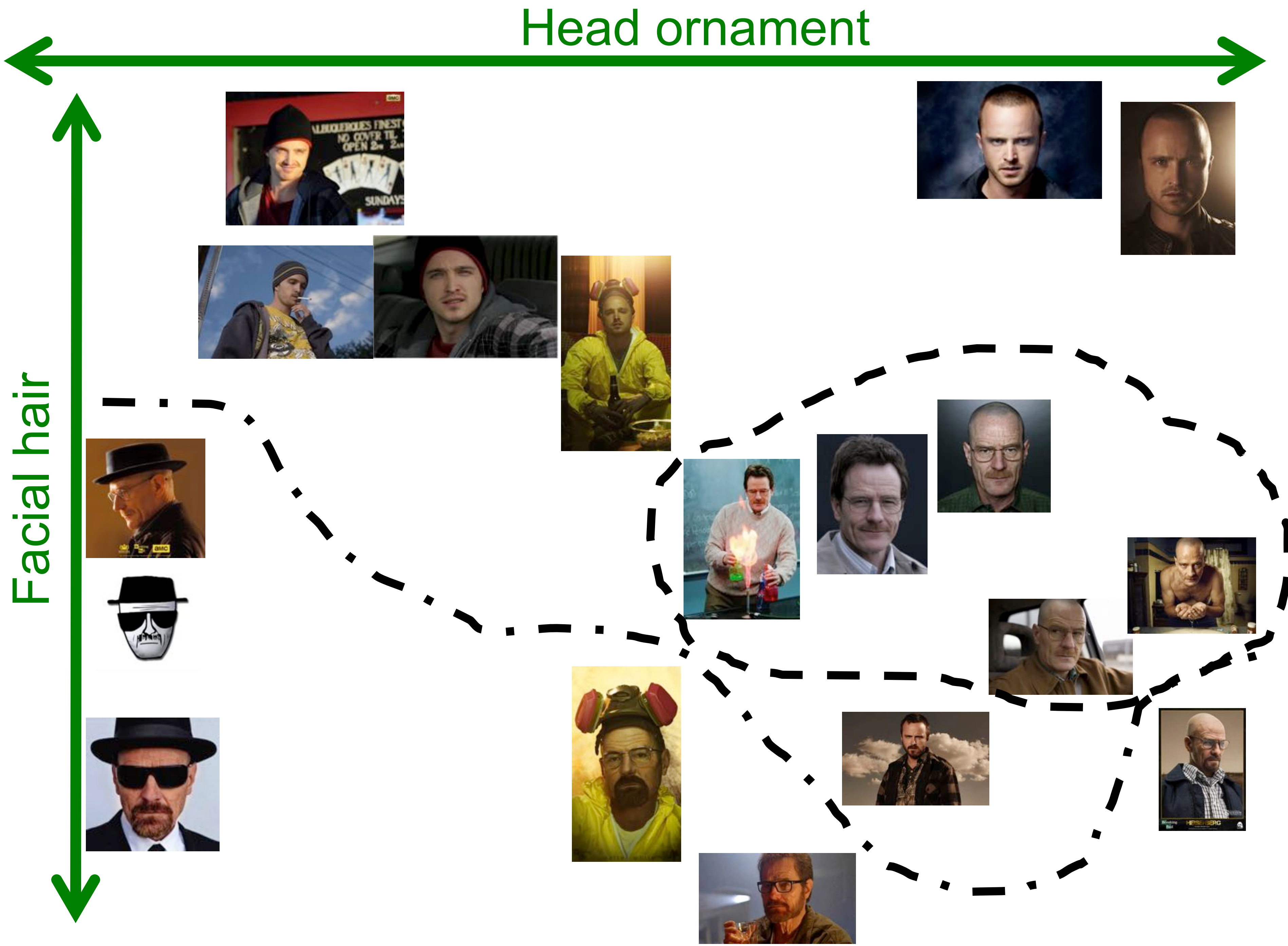


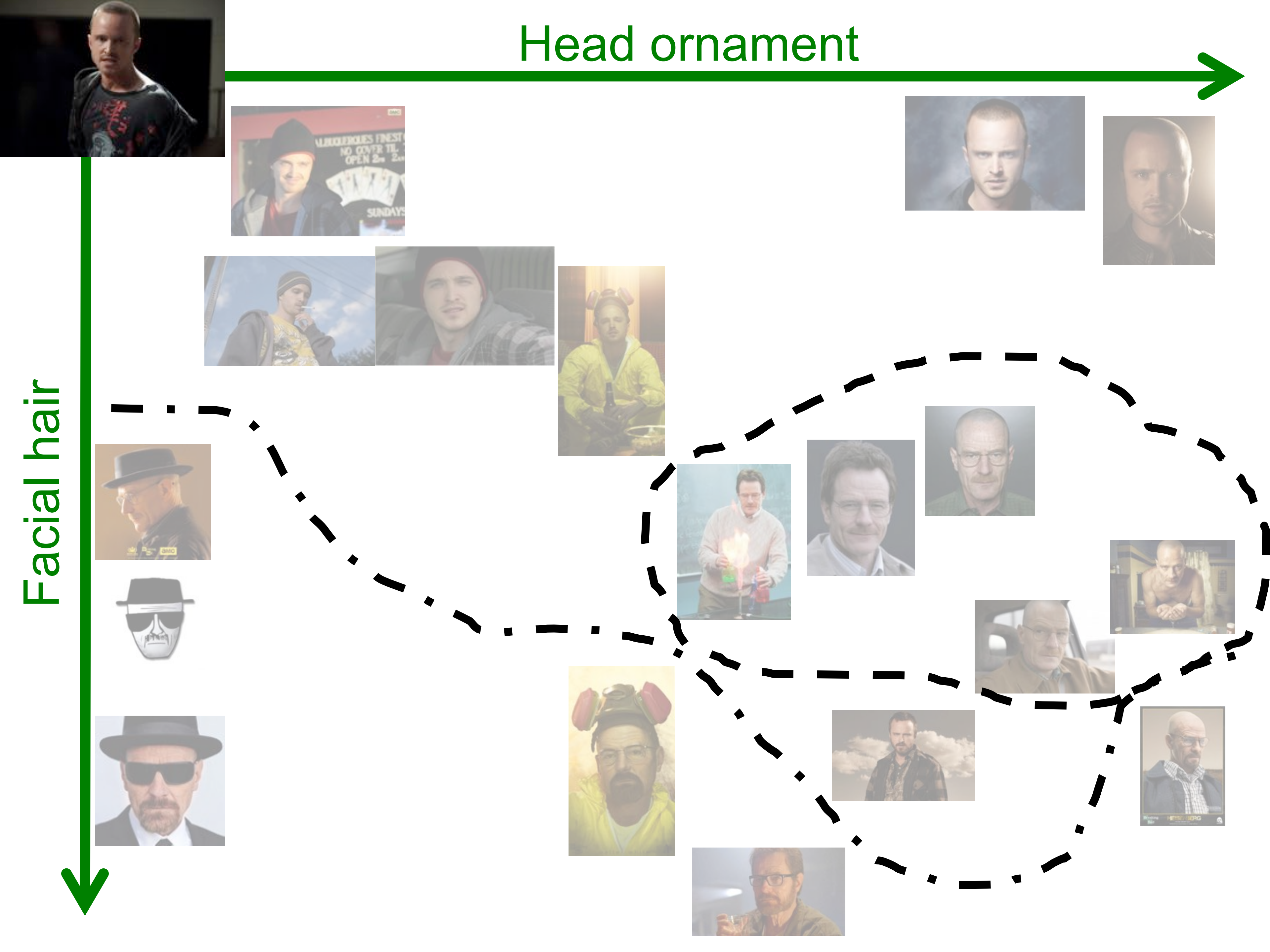



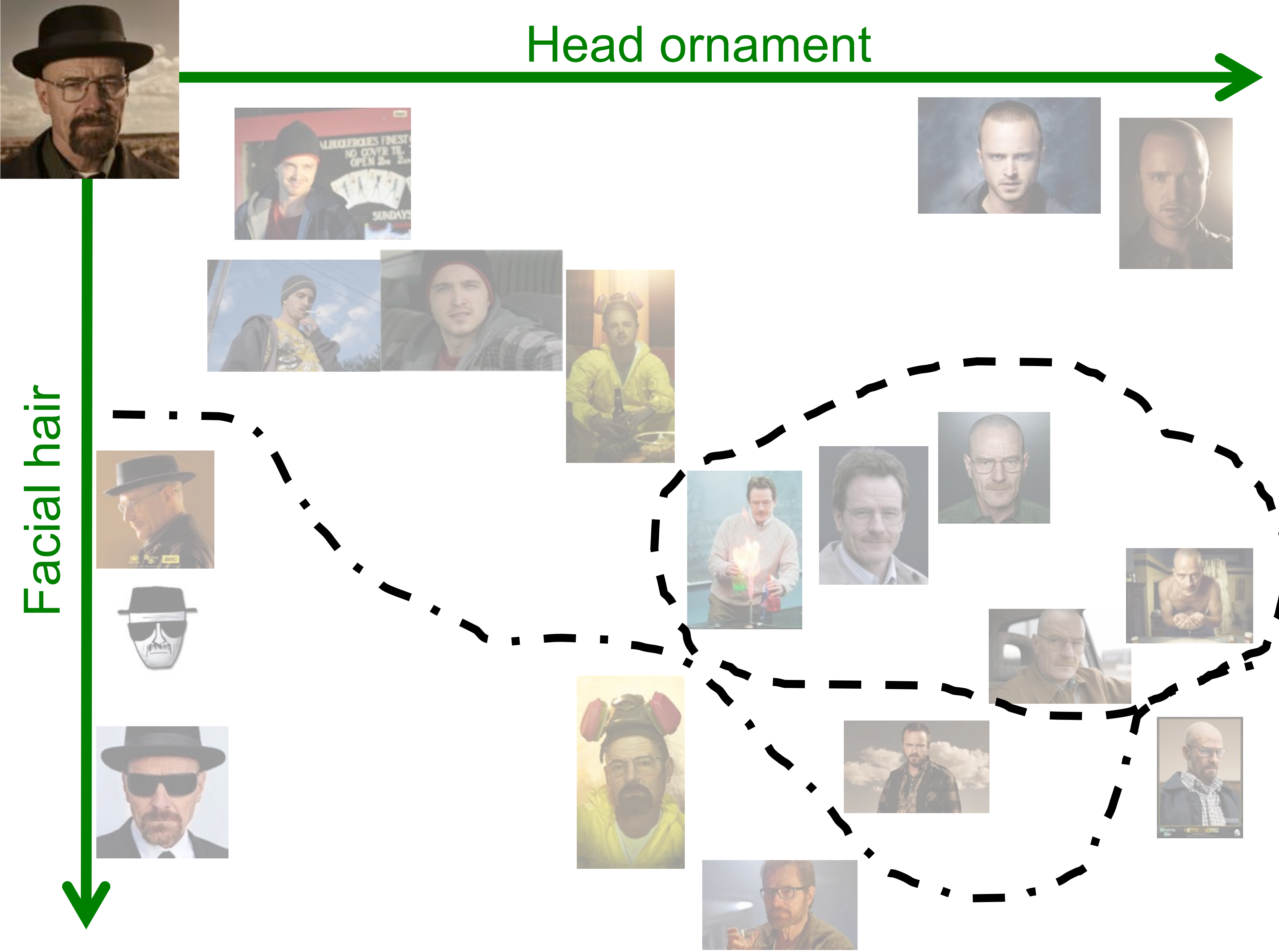

Facial hair

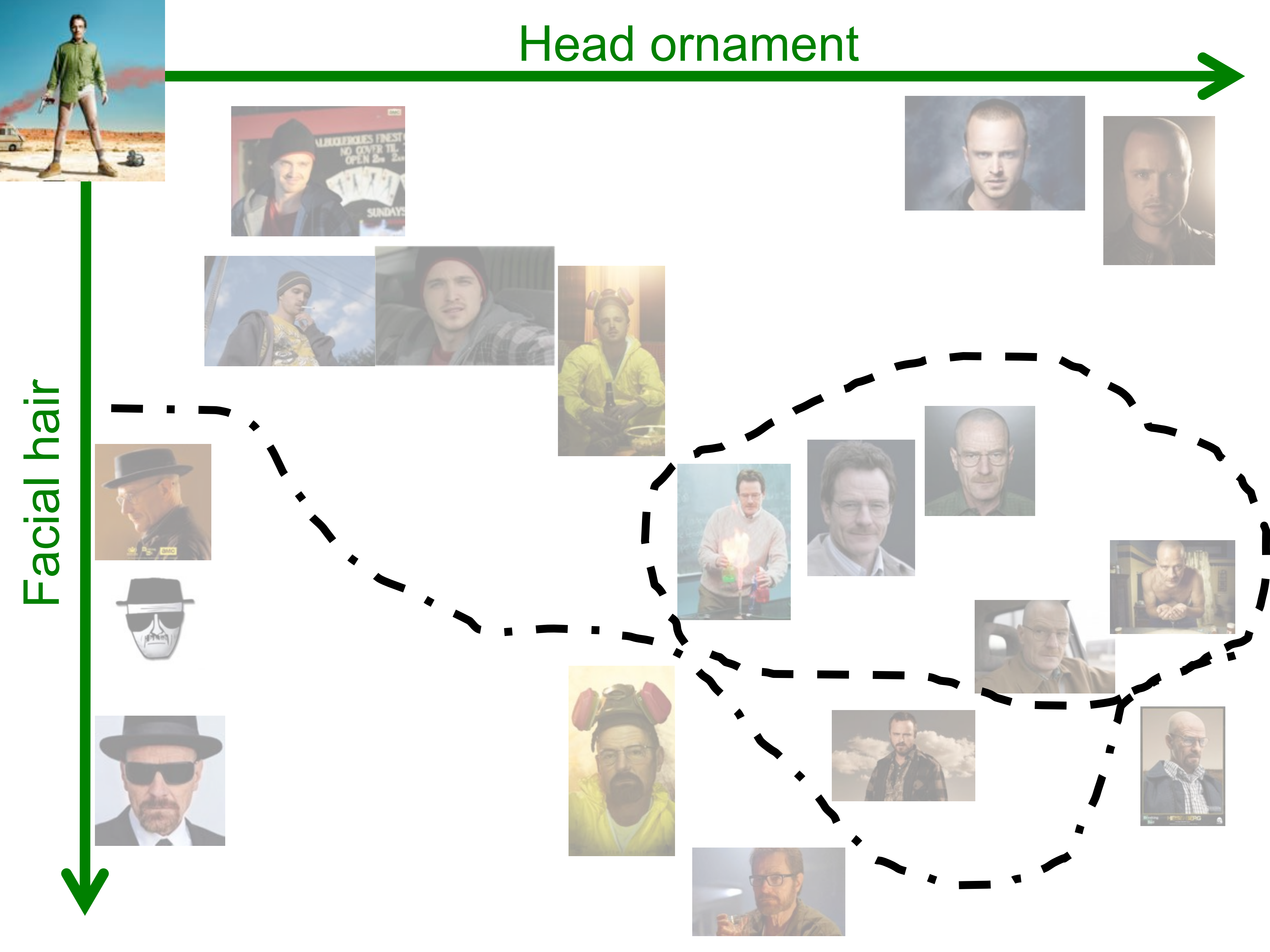

Head ornament
Facial hair

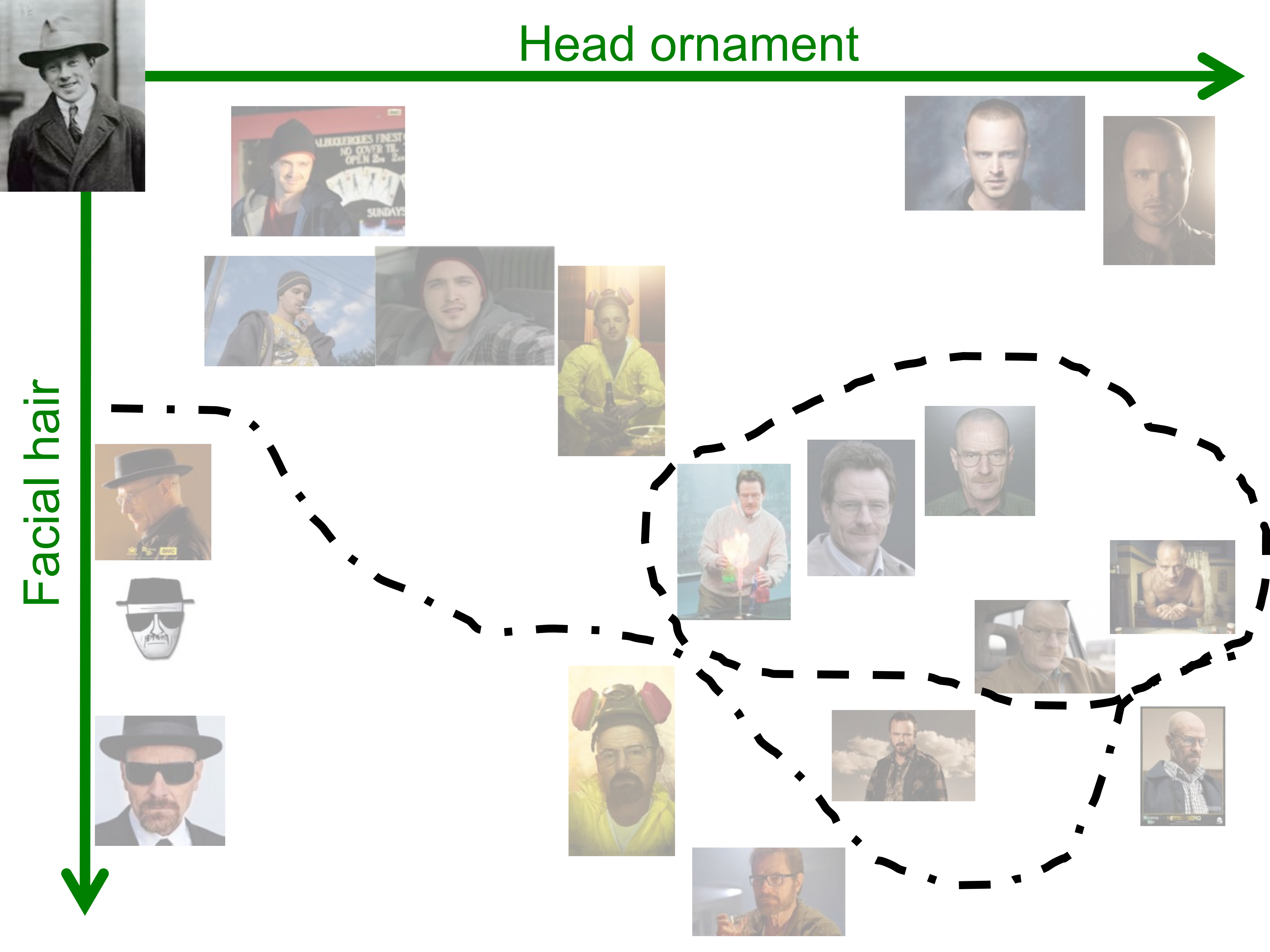

Head ornament
Facial hair


Head ornament
Facial hair


Head ornament
Facial hair

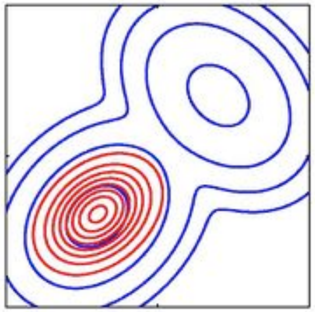

Head ornament
Facial hair


Head ornament
Facial hair


Head ornament
Facial hair

Types of learning

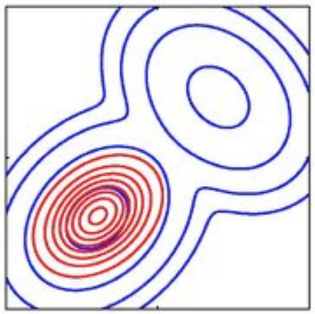
- Unsupervised learning
  - Model the data distribution without desired target values



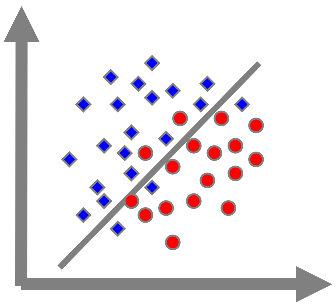
Unsupervised Learning

Types of learning

- Unsupervised learning
  - Model the data distribution without desired target values



- Supervised learning
  - Learning by examples (inputs and corresponding target values)
  - Minimization of an explicit error function



Supervised Learning

Training data

Input	Target Output
	Punk
	Glam rock
	Grunge
	Punk
	Grunge
	Glam rock

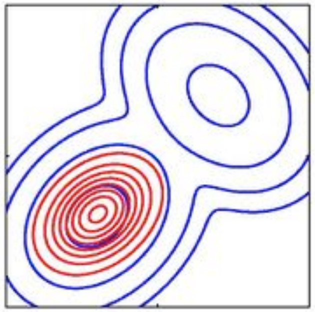
Test

Learning system

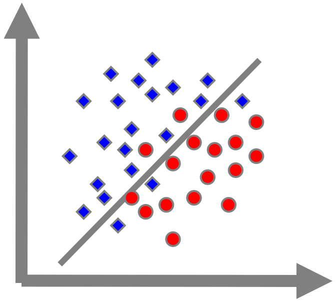
Glam rock Punk Grunge

Types of learning

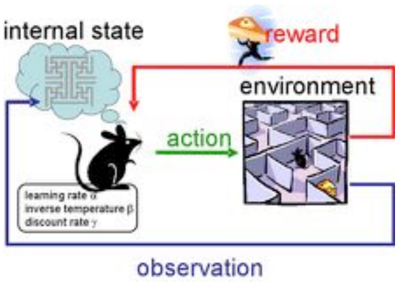
- Unsupervised learning
  - Model the data distribution without desired target values



- Supervised learning
  - Learning by examples (inputs and corresponding target values)
  - Minimization of an explicit error function



- Semi-supervised learning
  - Information about the performance is provided without explicitly providing target values



Course structure

1	
2	**Introduction**: Unsupervised machine learning, K-means
3	**Unsupervised learning**: Mixture Gaussian Models / Cross-validation
4	**Statistical pattern recognition**
5	**Supervised learning**: Linear classifiers
6	**Feature selection**: Filters, Wrappers
7	**Decision Trees**
8	**Dimensionality reduction**: Principal Component Analysis, Fisher
9	**Dimensionality reduction:** Self-Organizing Maps
10	**Regression methods**: Linear methods, Statistical approaches
11	**Regression methods**: Advanced issues
12	**Temporal pattern recognition**: Hidden Markov Models
13	**Temporal pattern recognition** : Hidden Markov Models
14	**Case study and Q&A**

Unsupervised learning

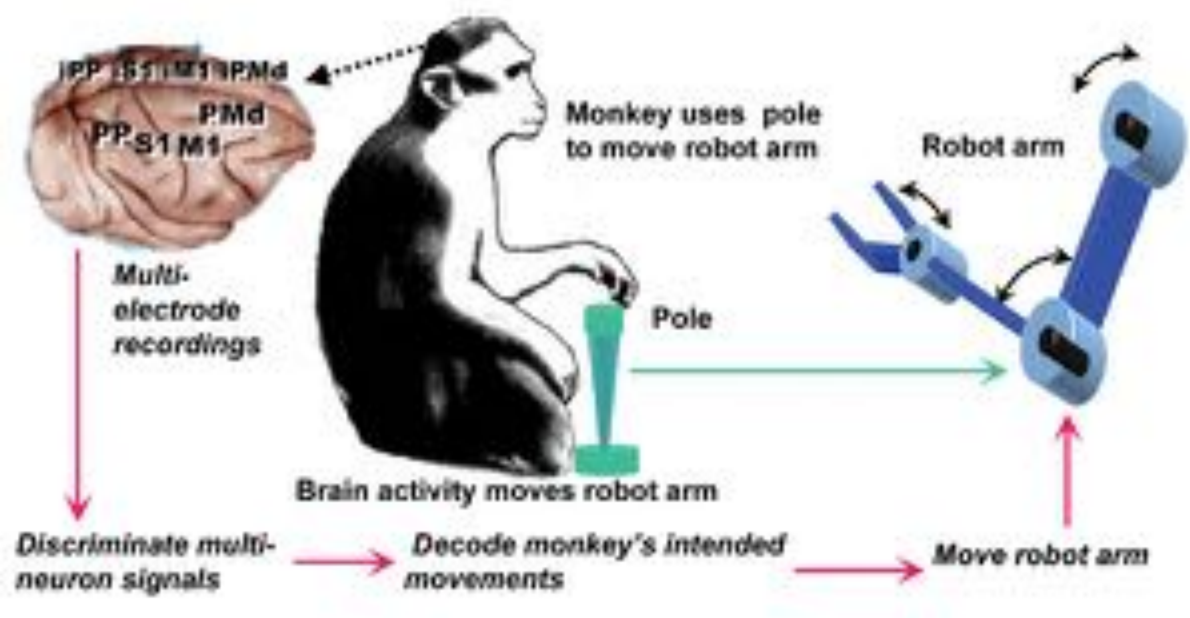
Labels are not available

Discover separate sub-classes (clusters) of data that are similar to each other

Characterize the structure of the data. Estimate its probability density function

Example: Neuroprosthetics

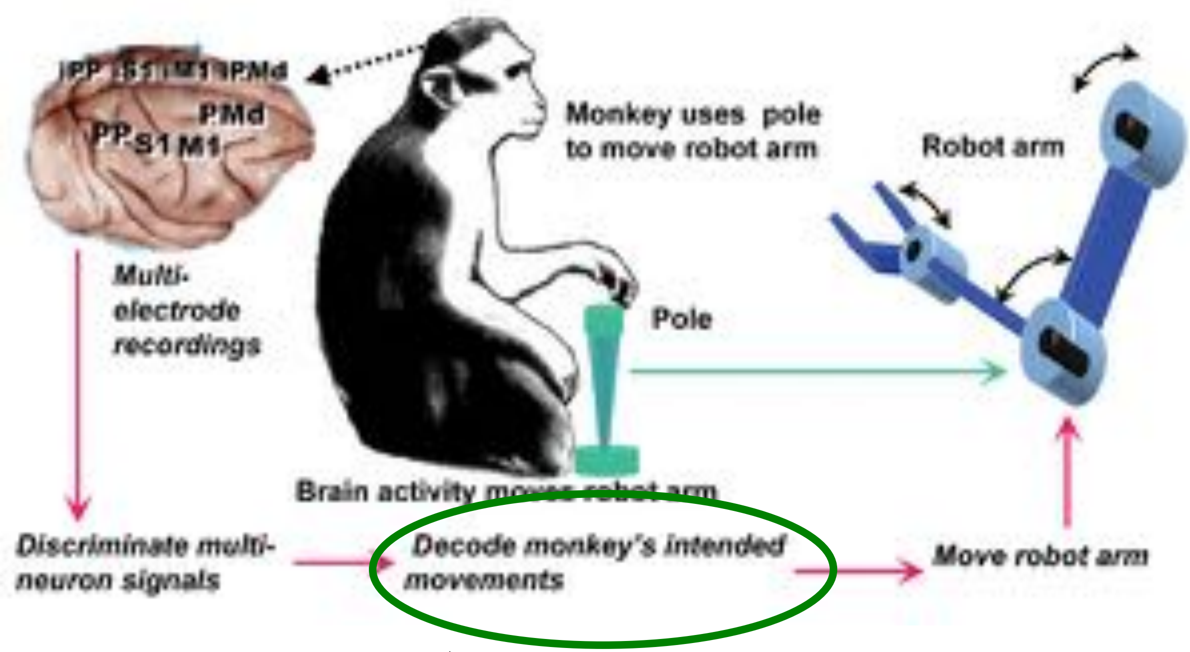
Example: Neuroprosthetics



Goal:

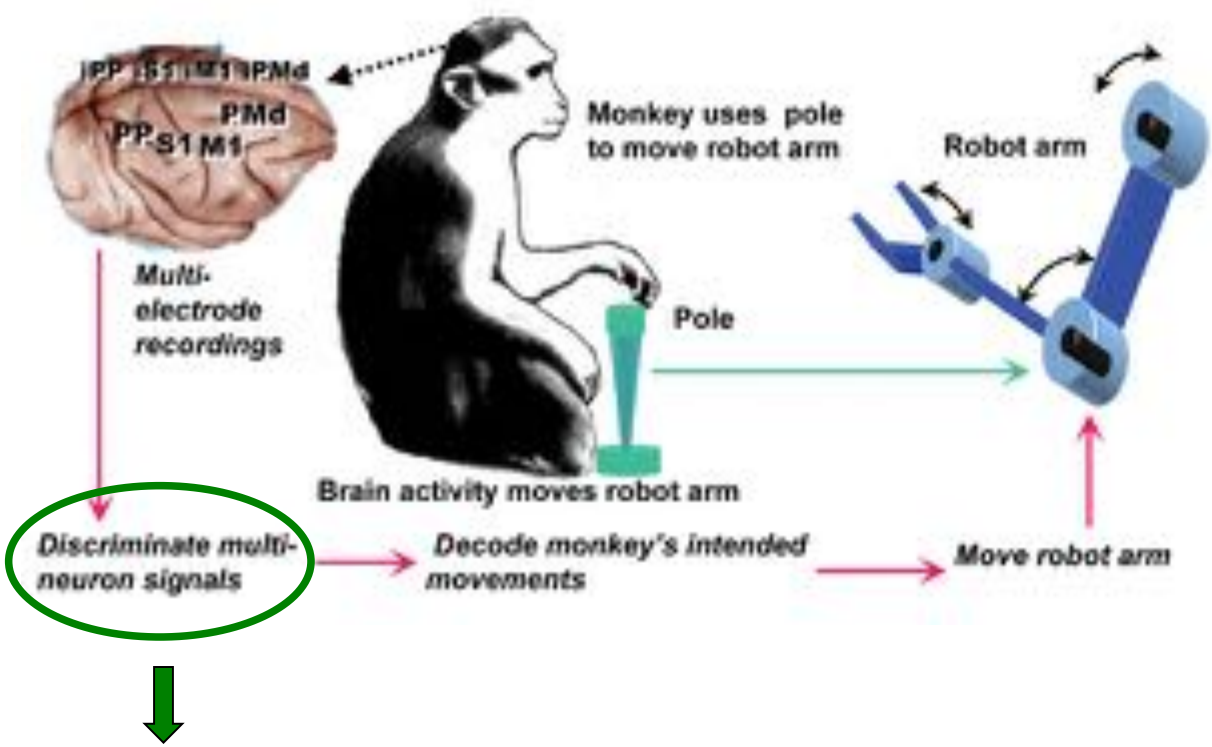
Decode arm movement direction from neural activity

Example: Neuroprosthetics



Relationship between neural activity (spikes) and arm intended movements are learned through Supervised learning

Example: Neuroprosthetics

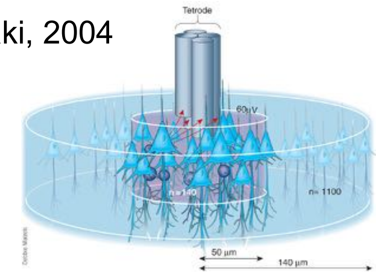


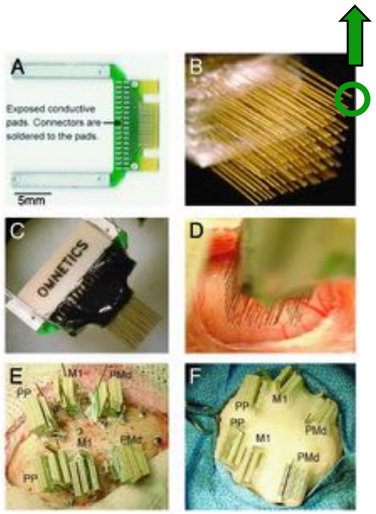
Each implanted electrode records signals from several <u>unknown</u> neurons

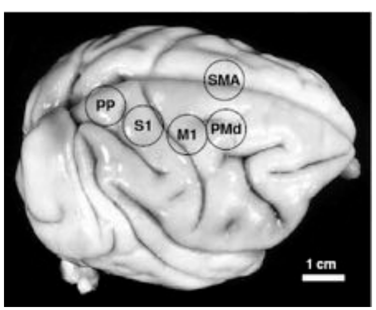
Individual neurons are identified using Unsupervised learning

Example: Neuroprosthetics

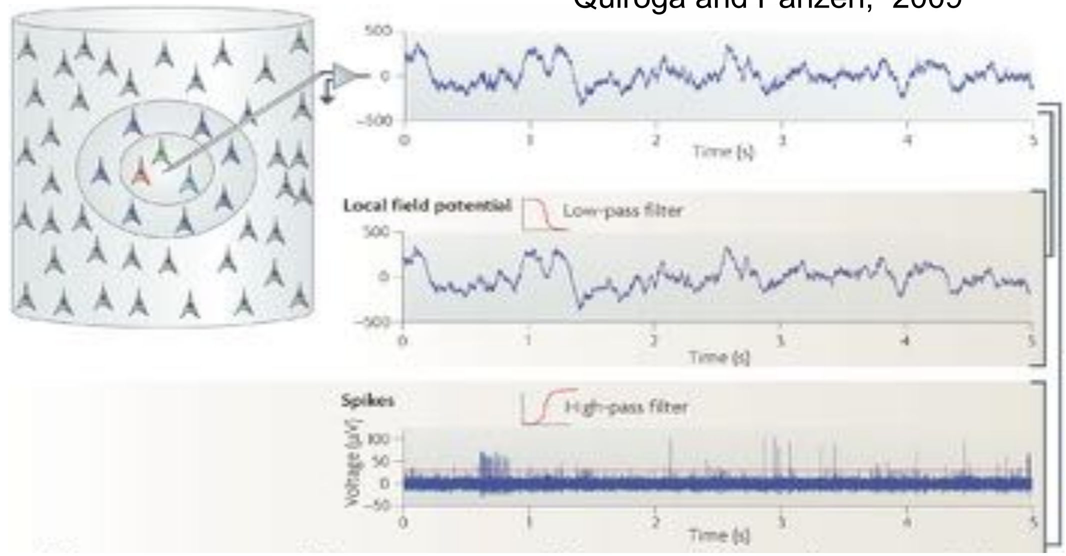
Buszàki, 2004





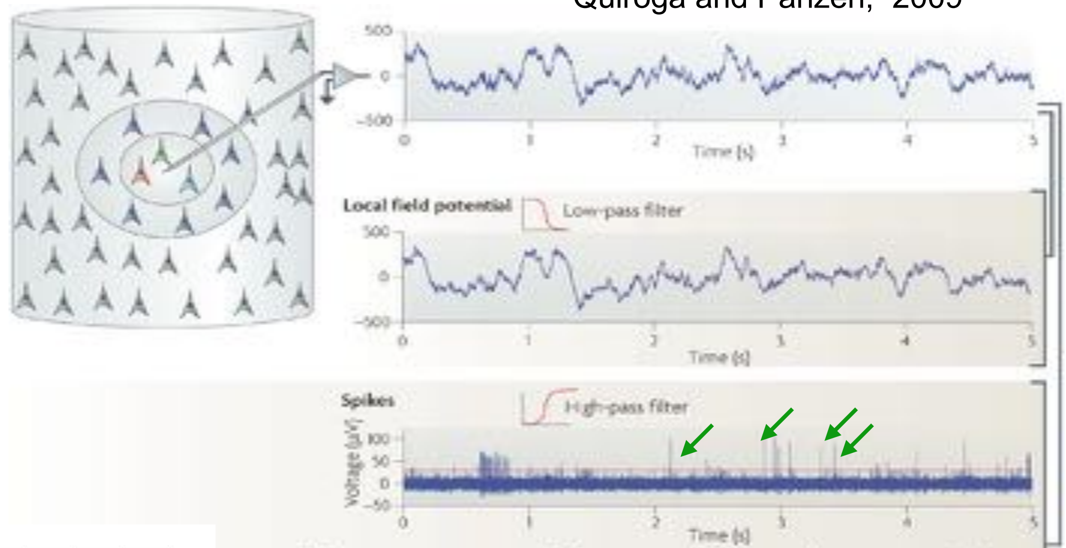


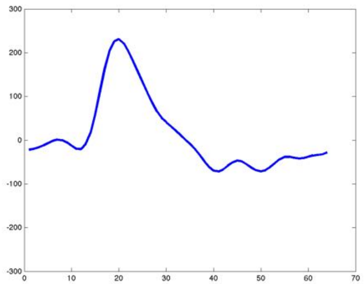
Quiroga and Panzeri, 2009

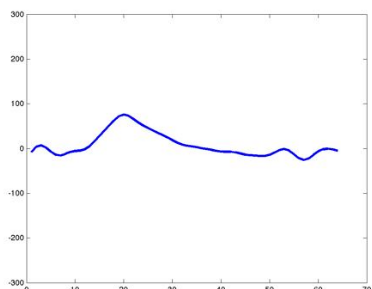


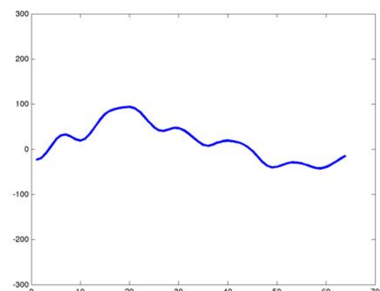
Example: Neuroprosthetics

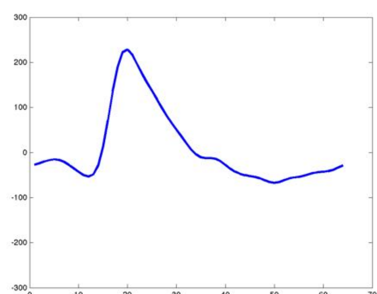
Quiroga and Panzeri, 2009

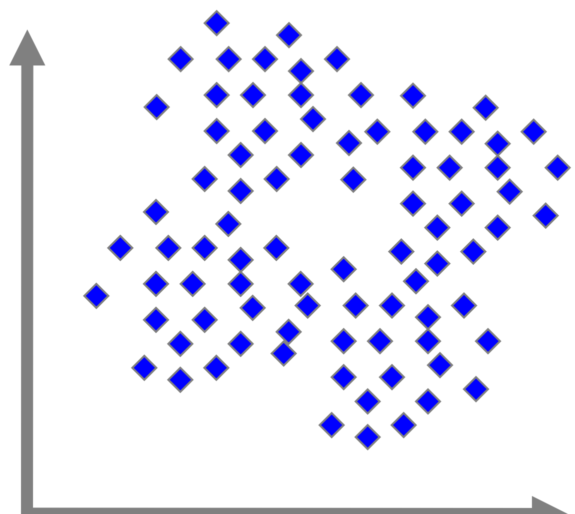






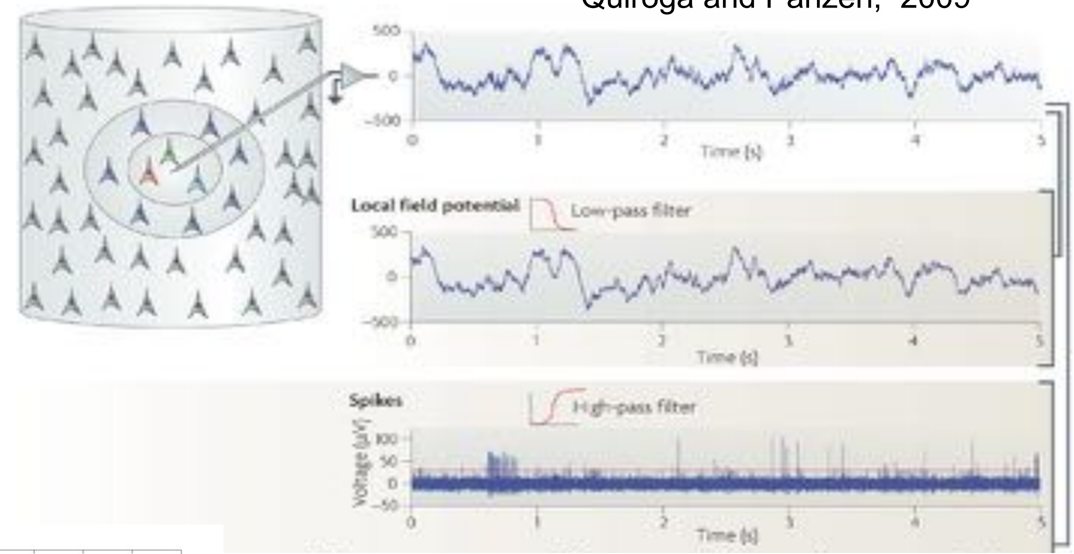


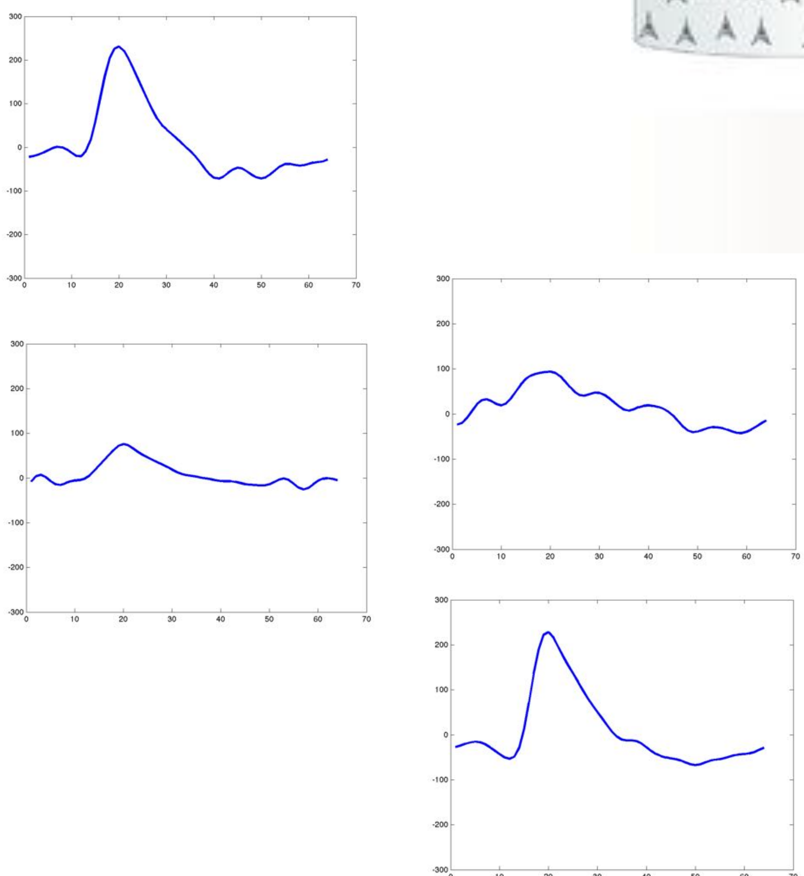


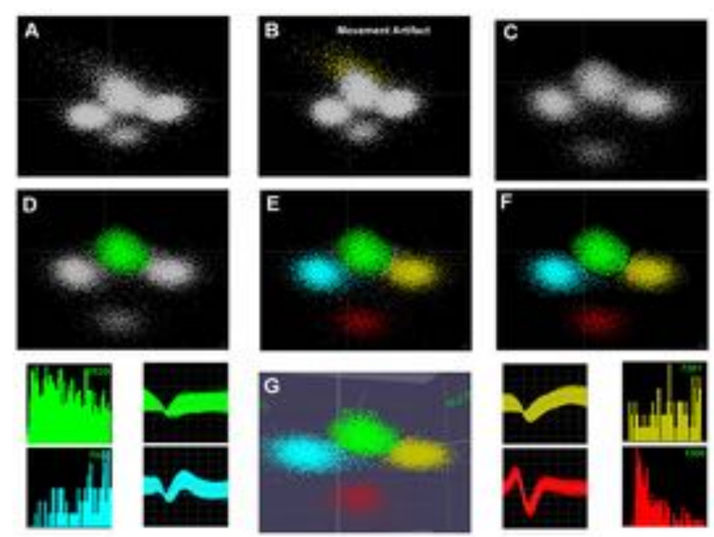


Example: Neuroprosthetics

Quiroga and Panzeri, 2009





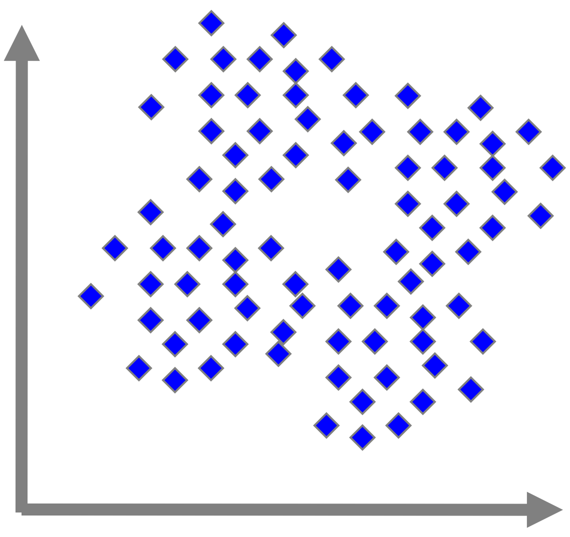


Nicolelis, 2003

Clustering

Characterize the data as an ensemble of groups of data point (clusters)

Cluster: Set of points whose inter-point distances are small compared to points outside the cluster



Clustering: K-Means

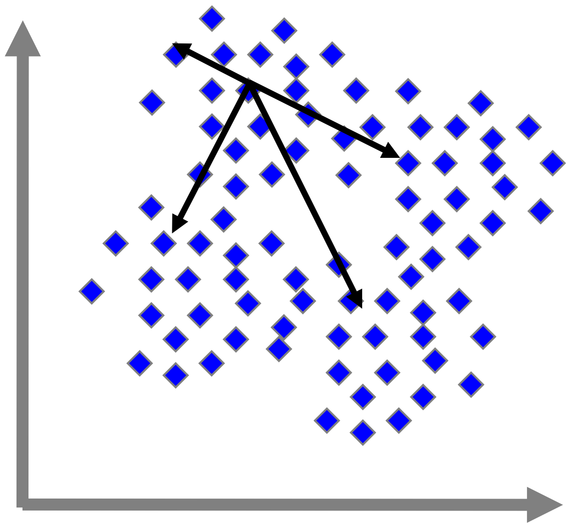
Assuming K clusters, we define:

$$\mu_k \equiv \text{Center of cluster k}$$

$$r_{nk} = \begin{cases} 1 & \text{if } \mathbf{x}_n \in \text{cluster k} \\ 0 & \text{otherwise} \end{cases}$$

Goal: minimize objective function

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} ||\mathbf{x}_n - \mu_k||^2$$



Clustering: K-Means

Assuming K clusters, we define:

$$\mu_k \equiv \text{Center of cluster k} \qquad r_{nk} = \begin{cases} 1 & \text{if } \mathbf{x}_n \in \text{cluster k} \\ 0 & \text{otherwise} \end{cases}$$

Goal: minimize objective function

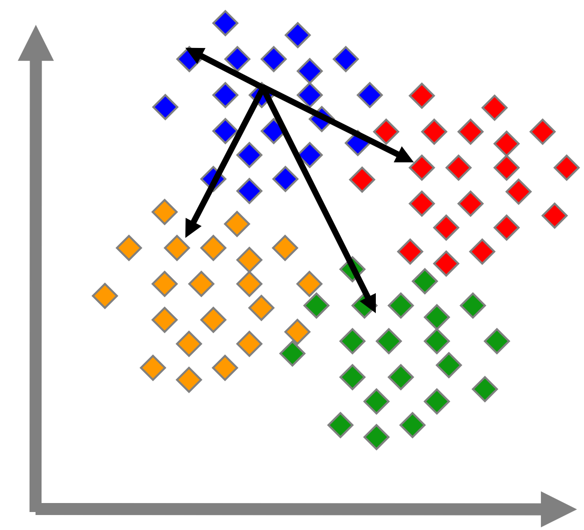
$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} ||\mathbf{x}_n - \mu_k||^2$$

Given r_{nk}, J is minimized by $\frac{\partial J}{\partial \mu_k} = 0$

$$2 \sum_{n=1}^{N} r_{nk}(\mathbf{x}_n - \mu_k) \quad = \quad 0$$

$$\boxed{\mu_k = \frac{\sum_n r_{nk} \mathbf{x}_n}{\sum_n r_{nk}}}$$

Mean of all points in cluster k



K-Means algorithm

Iterative algorithm

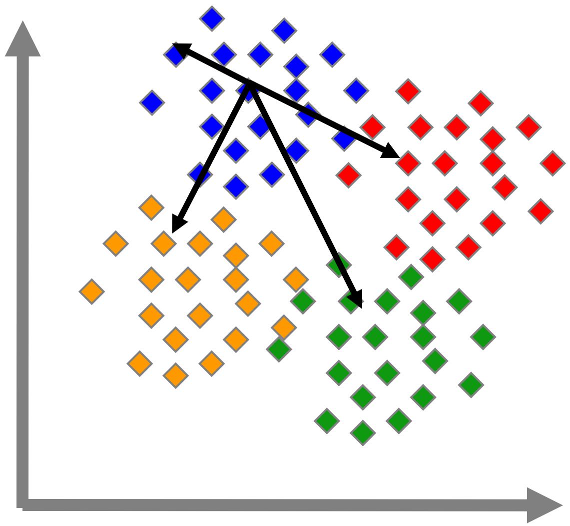
Initialize μ_k
Do
Update r_{nk}:

$$r_{nk} = \begin{cases} 1 & \text{if } k = argmin_j ||\mathbf{x}_n - \mu_j||^2 \\ 0 & \text{otherwise} \end{cases}$$

Update μ_k:

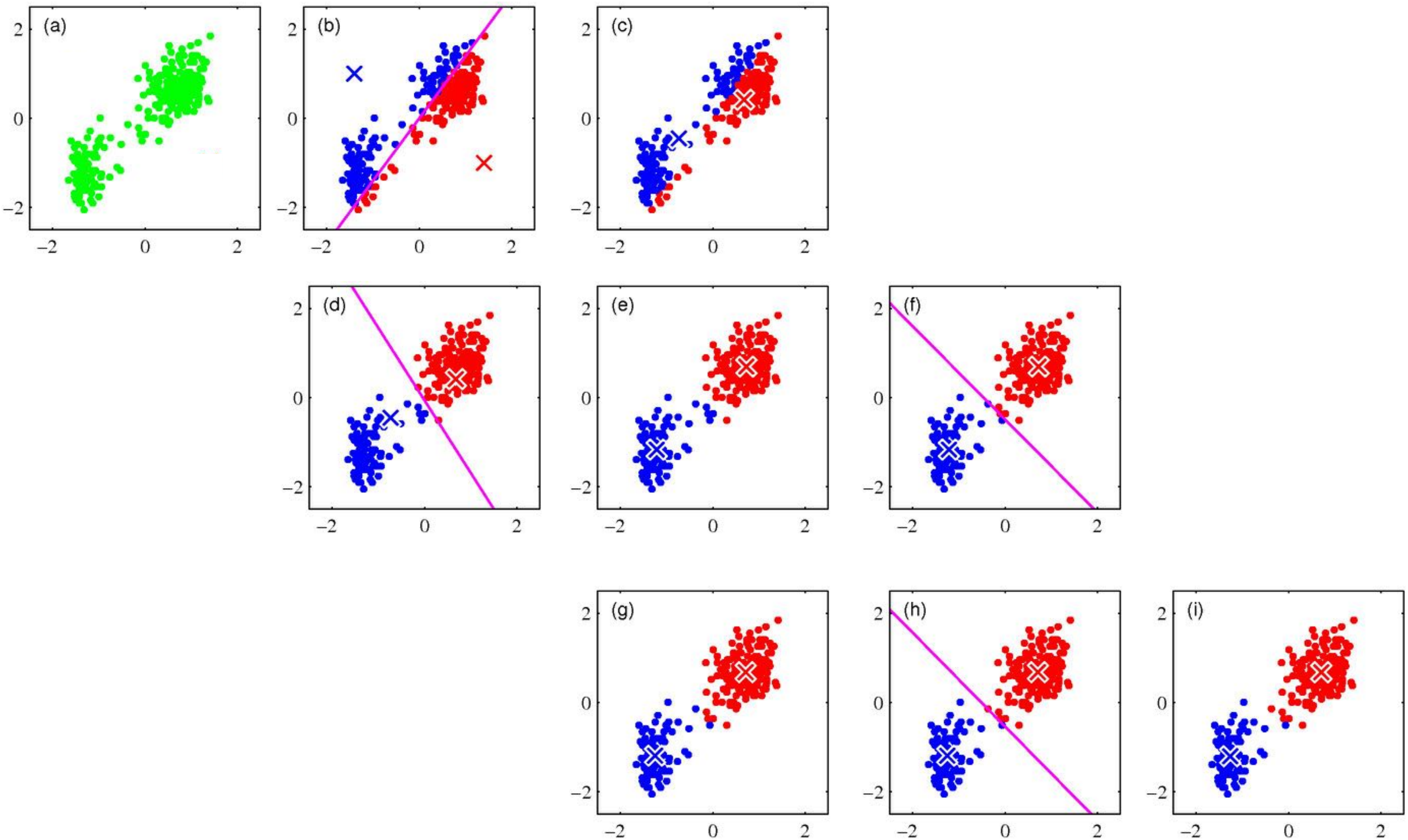
$$\mu_k = \frac{\sum_n r_{nk}\mathbf{x}_n}{\sum_n r_{nk}}$$

Until (no change)



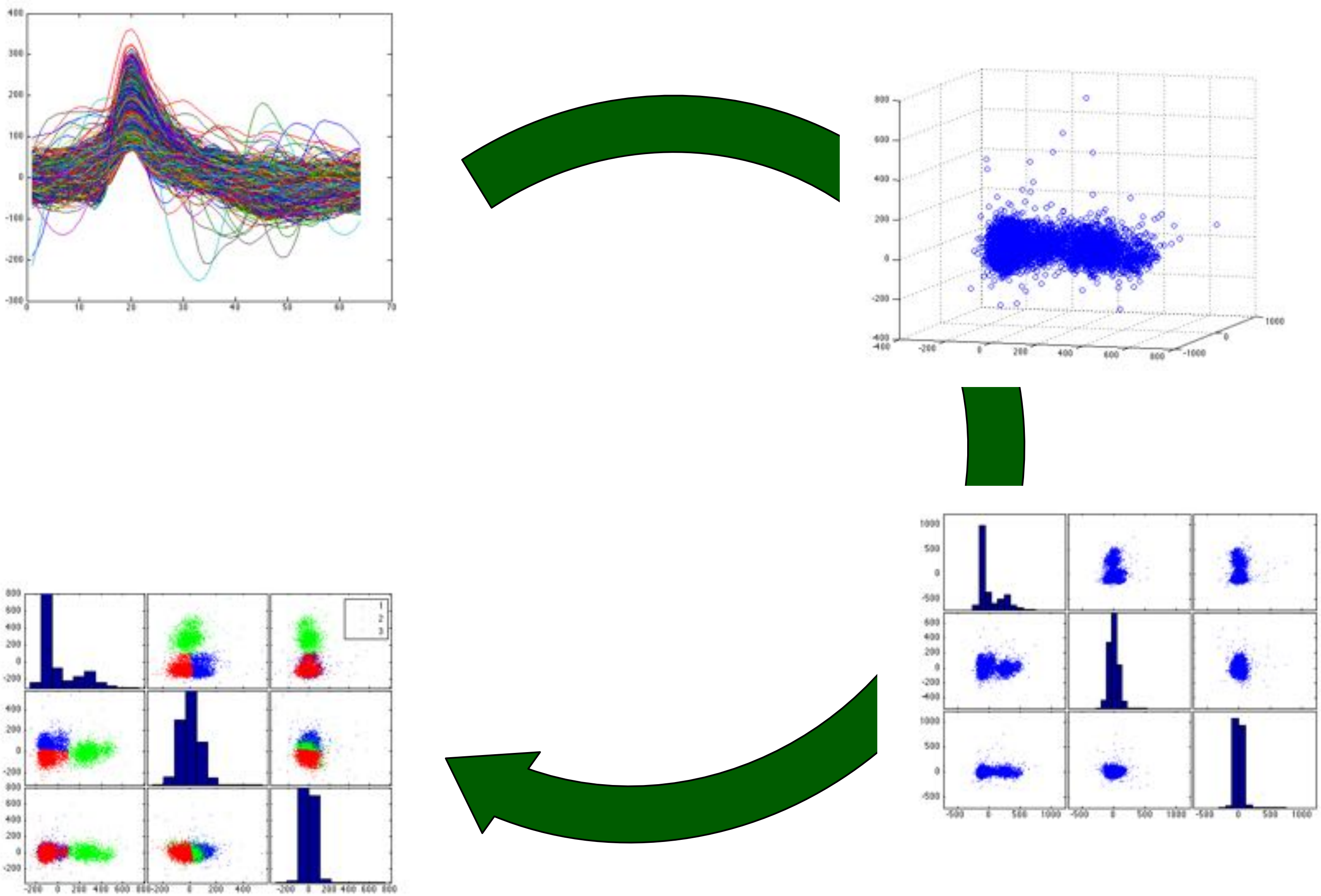
Clustering: K-means

Initialize μ_k | Update r_{nk}: | Update μ_k:



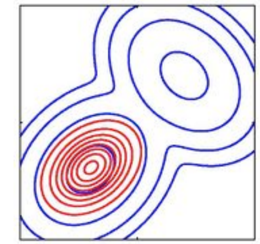
Figs: Bishop et al. Pattern recognition and machine learning, 2006

Miniproject 1: Spike sorting

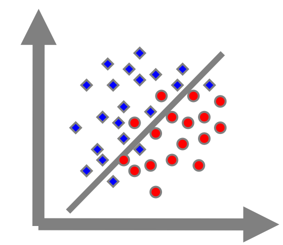


Summary

Different types of data analysis can be performed depending on the objective

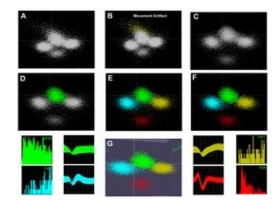


We focus on predictive analysis using machine learning

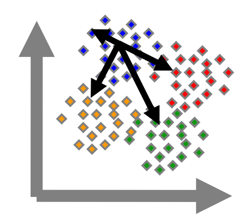


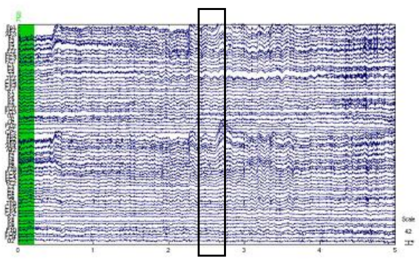
Learning can be based on examples (supervised) or on the data distribution (unsupervised)

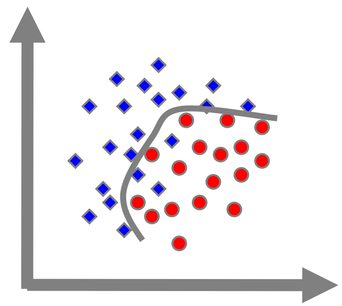
Unsupervised learning is used to process unlabelled data

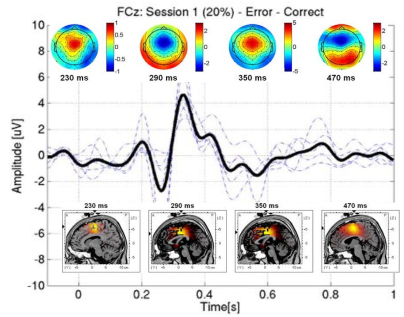


Data can be characterized by a set of different clusters of data points K-means algorithm









Data analysis and model classification

Unsupervised learning

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE