

# Machine Learning

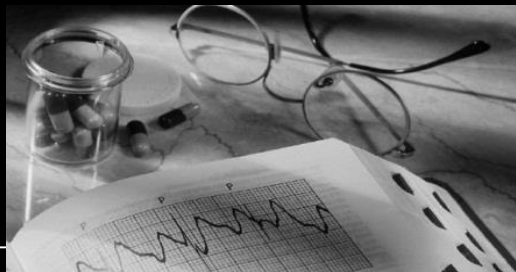
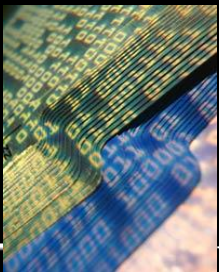
## COME REALIZZARE UN BUON PROGETTO

Fabio Stella

Associate Professor

c/o Department of Informatics, Systems and Communication

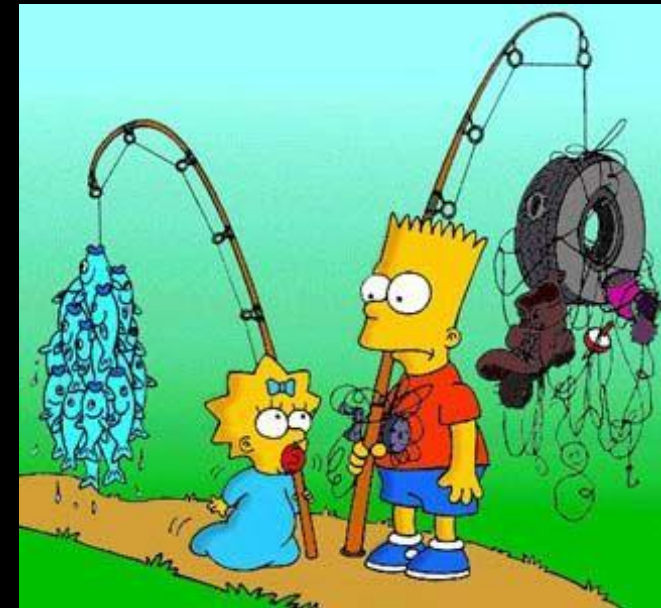
University of Milano-Bicocca



# STRUTTURA DEL PROGETTO

Un'indagine dovrebbe sempre originare da un'idea, un'intuizione, una curiosità che l'esperto di dominio suggerisce al data scientist.

- Evitare “*Fishing Expeditions*”
- Prender visione dei dataset KAGGLE suggeriti
- Cercare in KAGGLE dataset simili ai suggeriti
- Formulare le Research Questions
- Verificare se il dataset consente di rispondere alle Research Questions
- I membri del Team devono discutere, confrontarsi e concordare



# STRUTTURA DEL PROGETTO

Scelto il dataset, concordate le Research Questions (suggerisco 2/3), il Team procede alla progettazione dello studio di Machine Learning, vale a dire discute e concorda sul flusso che porta dai dati alle risposte alle Research Questions:

- Pianificare il flusso di Machine Learning (workflow)
- Implementare il flusso di Machine Learning in KNIME, possibile utilizzare tutto ciò che è naturalmente ospitato da KNIME, ma il progetto deve sviluppare un workflow KNIME
- I nodi, metanodi KNIME vanno annotati, descrivendone la finalità, eventuali parametri, assunzioni sui dati in ingresso, missingness? ...
- Sfruttare tutto il materiale disponibile, ricordarsi di citare le fonti

# STRUTTURA DEL PROGETTO

## Il Report deve

- far riferimento al workflow KNIME, ponendo in relazione le analisi condotte con le componenti del workflow KNIME
- seguire la struttura offerta dagli esempi che sono resi disponibili sulla piattaforma del corso
- non esser popolato da tabelle, figure, grafici se non strettamente necessari a comprendere e valorizzare il messaggio al lettore che intendete inviare
- non dilungarsi su dettagli algoritmici, chi legge i vostri progetti si assume che conosca la materia :-D

# STRUTTURA DEL PROGETTO

## Alcune domande

- **Report in Italiano o Inglese?**
  - Nessun premio ne penalizzazione!!!
  - Se desiderate valorizzare il vostro lavoro a livello internazionale (LinkedIn, Google Scholar, ResearchGate, ...) in inglese, astenetevi se non siete confidenti circa il vostro Inglese, potete valorizzare il lavoro anche in Italiano
- **Quanti membri del Team?**
  - Suggerisco da 3 a massimo 5, confrontarsi, discutere, concordare, dissentire sono opzioni importanti per crescere, imparare, sviluppare un progetto robusto.
  - Possibile ma sconsigliato fare progetto da soli.
- **Posso fare domande durante lo sviluppo del progetto?**
  - SI, se si tratta di domande su ambiguità lasciate dal docente
  - NO, se si tratta di avere indicazioni di conferma o critica da parte del docente

# STRUTTURA DEL PROGETTO

E la consegna?

- Il report ed il modello KNIME vanno caricati sulla piattaforma MOODLE.
- Si comprimono il file PDF del Report e il workflow KNIME in un file zip
- Il nome dello zip file deve essere “TEAM\_num”
- Solo un membro del TEAM effettua la consegna
- Il workflow va esportato escludendo il file dei dati utilizzati
- La valutazione del progetto è resa nota dopo la prova in laboratorio

...allora...

**Coraggio e  
buon lavoro!**

...