Machine Learning per Health & Medicine

Gianvito Pio, Michelangelo Ceci,

Paolo Mignone, Emanuele Pio Barracchia and Donato Malerba

Dipartimento di Informatica, Università degli Studi di Bari "Aldo Moro"

Via Orabona, 4 - 70125 Bari - Italy

KDDE Projects and Research Activities

- Current research activities: study of the relationships among sub-cellular entities and their possible correlations with diseases

Homogeneous

gene-gene, miRNA-miRNA
Impact: gene regulation analysis

Heterogeneous

miRNA-mRNA, ncRNA-disease
Impact: discovery of gene-disease relationships

- Current research projects:

 – *UE CA18131 ML4Microbiome*

    Identification of the relationships between the human microbiome and pathologies

  – *PON TALIsMAn* «Tecnologia per gli ambienti della vita»

    Tecnologie di **A**ssistenza persona**LI**zzata per il **M**iglioramento della qu**A**lità della vita

Link prediction & Co-clustering

- **Task**: identification of (single & groups of) miRNA-mRNA associations

- **Method**: semi-supervised meta-learning approach & overlapping and hierarchical co-clustering
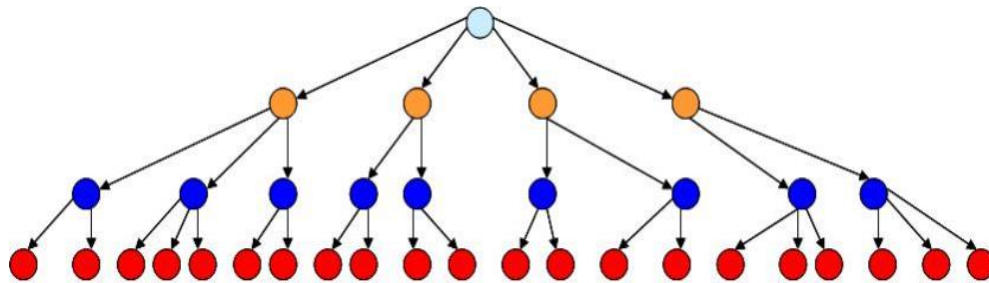
Labeled data (known links)

		Score$_1$	score$_2$	...	...	score$_k$	label
g_1	m_1						Yes
g_1	m_2						?
..							?
							Yes
							?
							?
g_M	m_N						Yes

Learns to combine the prediction scores of multiple prediction methods

G. Pio, M. Ceci, D. D'Elia, C. Loglisci, D. Malerba, *A novel biclustering algorithm for the discovery of meaningful biological correlations between miRNAs and mRNAs*, BMC Bioinformatics 14 (Suppl 7), S8, 2013

G. Pio, D. Malerba, D. D'Elia, M. Ceci *Integrating microRNA target predictions for the discovery of gene regulatory networks: a semi-supervised ensemble learning approach*, BMC Bioinformatics 15 (Suppl 1), S4, 2014

G. Pio, M. Ceci, D. Malerba, D. D'Elia, *ComiRNet: a Web-based System for the Analysis of miRNA-gene Regulatory Networks*, BMC Bioinformatics 16 (Suppl 9), S7, 2015
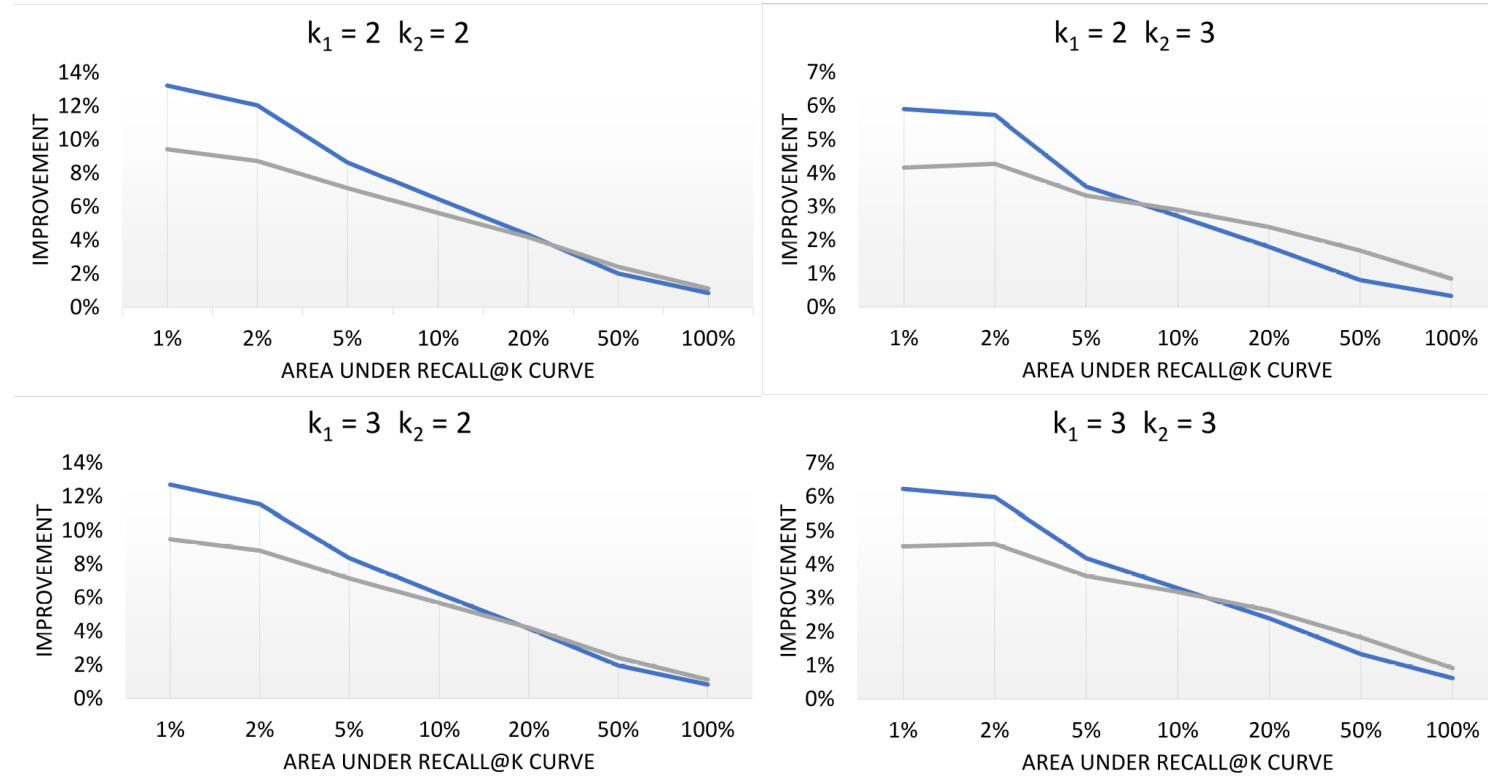
Multi View Meta Learning

- **Task**: gene regulatory network reconstruction
- **Method**: *semi-supervised multi-view learning* approach which combines the output of multiple prediction methods

M. Ceci, G. Pio, V. Kuzmanovski, S. Dzeroski *Semi-Supervised Multi-View Learning for Gene Network Reconstruction* PLoS One 10(12): e0144031, 2015
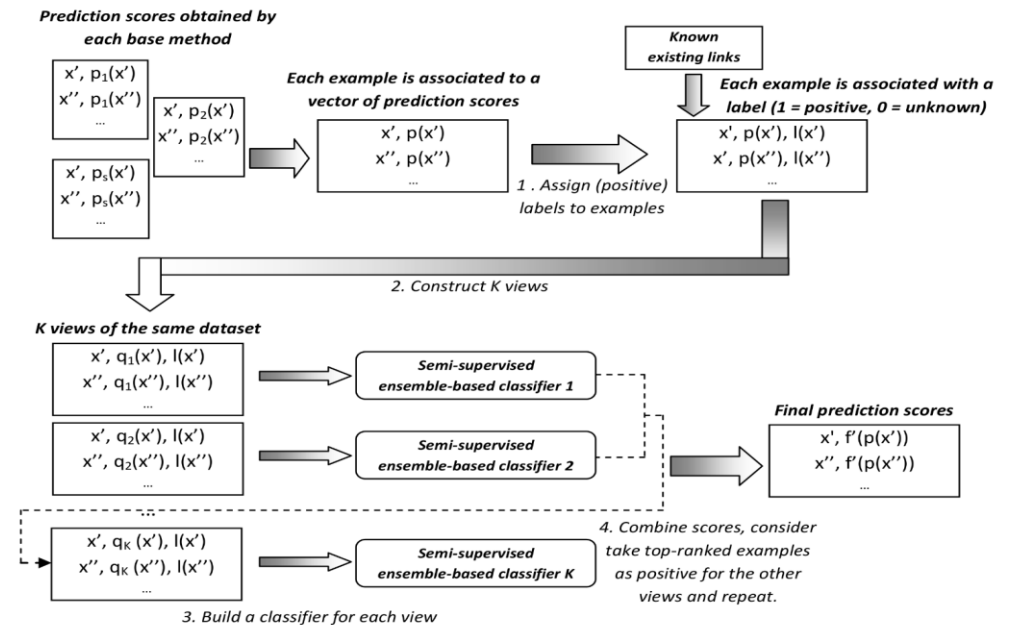
Transfer Learning

- **Task**: gene regulatory network reconstruction

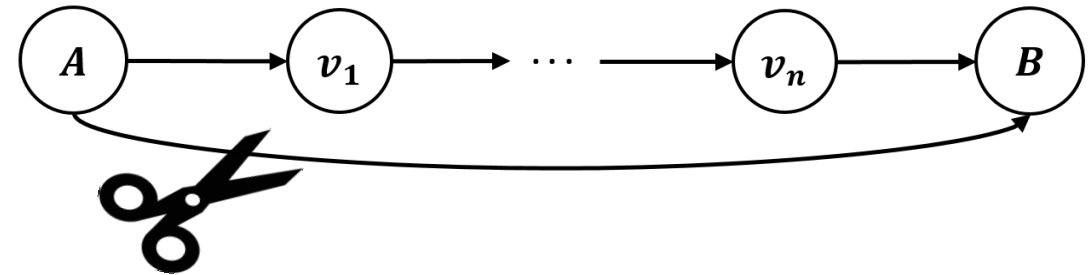- **Method**: *transfer learning* based on *clustering, instance weigthing* and *weighted SVM.*

We exploit information about the *mouse* gene regulations to better reconstruct the *human* gene regulatory network

$k_1 = 2$ $k_2 = 2$

$k_1 = 2$ $k_2 = 3$

$k_1 = 3$ $k_2 = 2$

$k_1 = 3$ $k_2 = 3$

IMPROVEMENT

AREA UNDER RECALL@K CURVE

Blue: *our method*
Gray: *union of instances*

P. Mignone, G. Pio *Positive Unlabeled Link Prediction via Transfer Learning for Gene Network Reconstruction* ISMIS 2018: 13-23
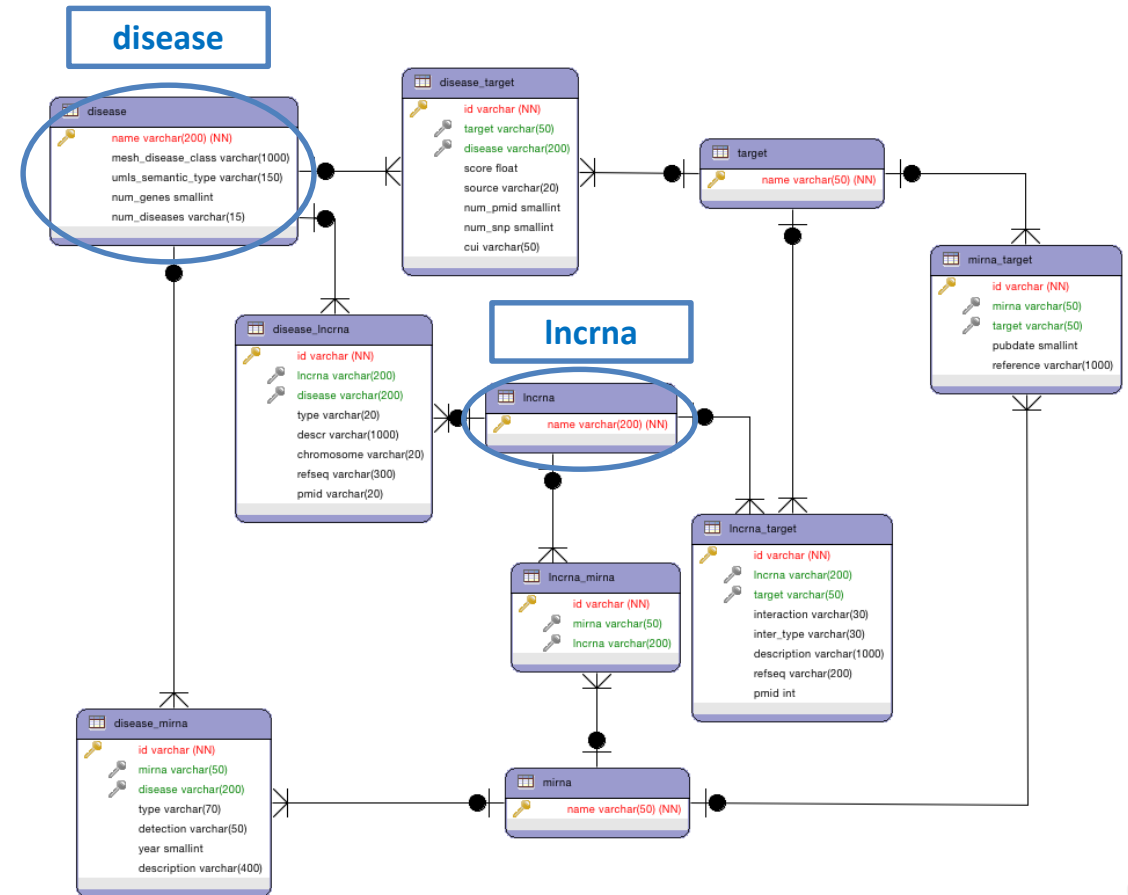
Causality Analysis

- **Task**: gene regulatory network reconstruction
- **Method**: identification of indirect relationships and their exploitation to identify incorrect gene regulatory associations. We exploit the concept of common *causes* and common *effects.*

G. Pio, M. Ceci, F. Prisciandaro, D. Malerba, *LOCANDA: Exploiting Causality in the Reconstruction of Gene Regulatory Networks. Discovery Science 2017*, LNCS 10558, pp. 283-297, 2017
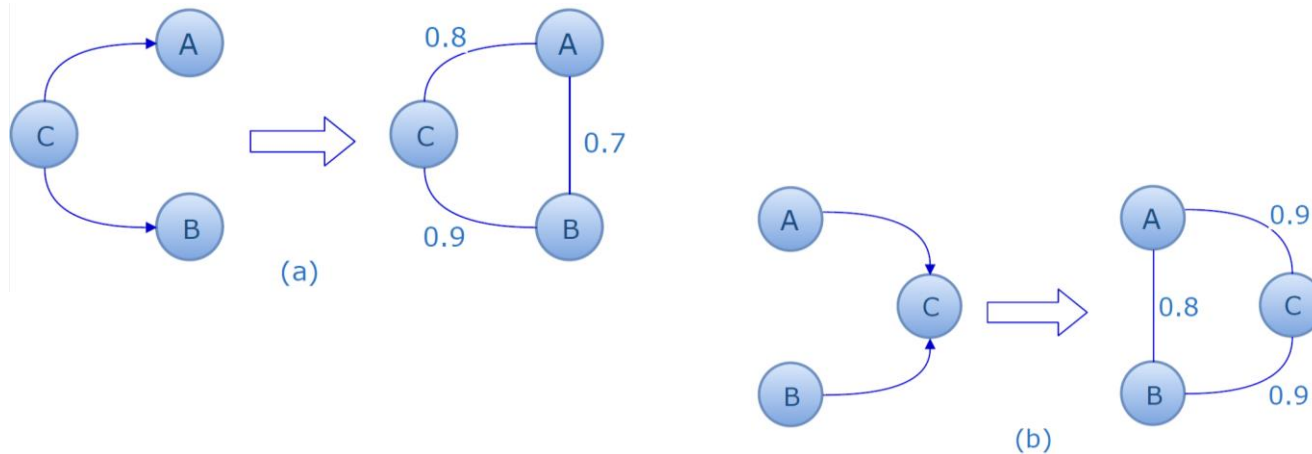
Heterogeneous Clustering

- **Task**: identification of ncRNA-disease associations

- **Method**: identification of ncRNA-disease clusters from *heterogeneous networks* and identification of previously unknown associations

E. P. Barracchia, G. Pio, D. Malerba, M. Ceci, *Identifying lncRNA-Disease Relationships via Heterogeneous Clustering*, NFMCP@ECML/PKDD 2017, pp. 35-48, 2018

TALIsMAn Project

Development of prevention systems, monitoring and diagnosis, personalized on the individual.

KDDE activities:

- definition of analytics tools to identify behavioural patterns and recognize anomalous situations in order to enrich the health profiling.

- support the discovery of statistically interesting sequences of activities, which represent optimal treatment protocols, useful to detect possible anomalies.

ML4Microbiome Project

KDDE activities:

- analysis and evaluation of ***the state of the art***, in terms of Machine Learning methods

- definition of the priority areas for the development of ***new Machine Learning methods*** that can solve specific problems for the study of the microbiome

- identification of the best practices for the application of Machine Learning for the study of the ***microbiome***.

KDDE Staff working in the area of AI for Health and Medicine

Prof. **Michelangelo Ceci**
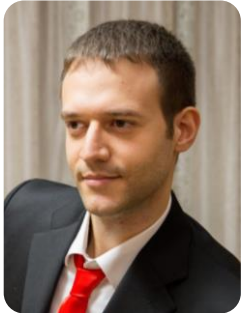
Dott. **Emanuele Pio Barracchia**

Dr. **Gianvito Pio**

Dott. **Paolo Mignone**

Prof. **Donato Malerba**