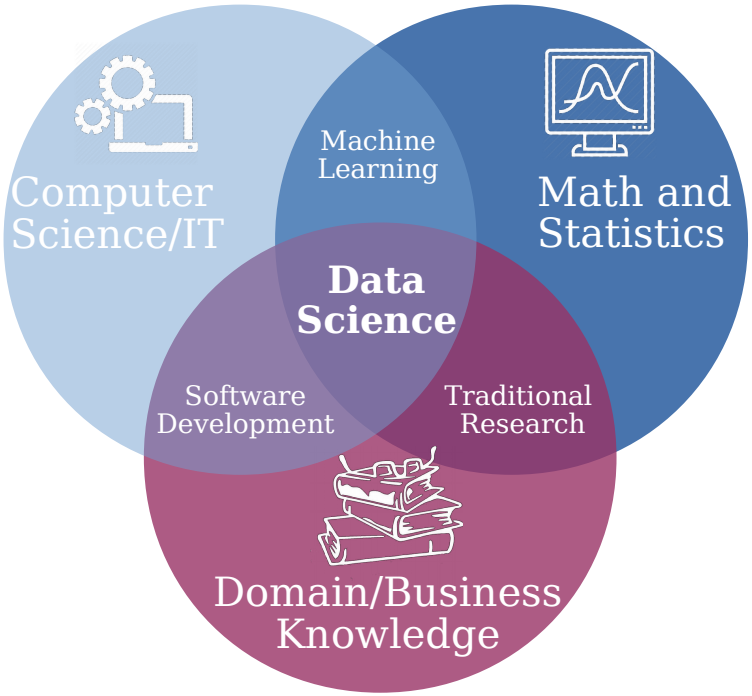
Complex Networks and Machine Learning

Silvia Giordano, Luca Luceri, Kamini Garg
silvia.giordano@supsi.ch

Data Science: Machine Learning

In a wide range of domains:

- In automatic manner
- Data structure and features

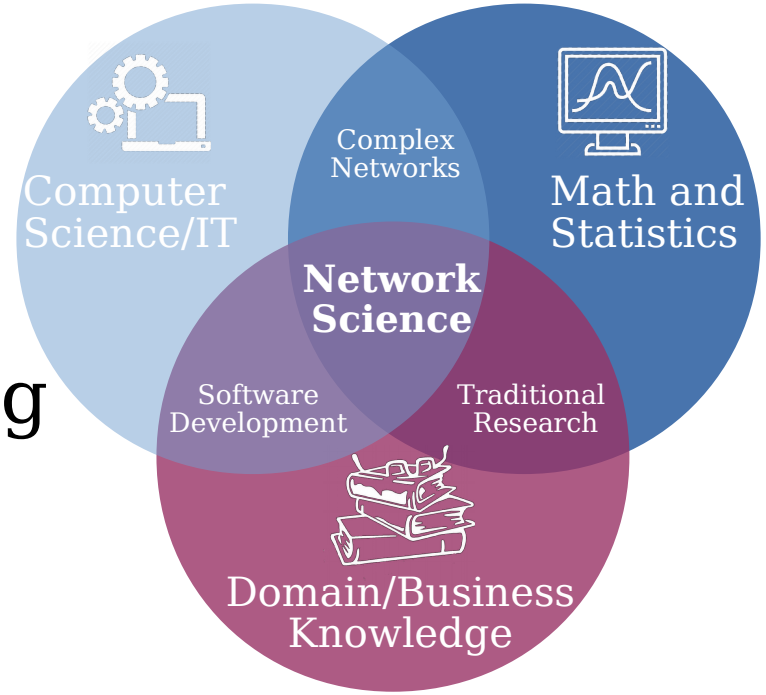


generate models apt to organize the existing knowledge or mimic the behaviour of experts

Complex system

In a wide range of domains:

- In automatic manner
- Objects relationships/networking

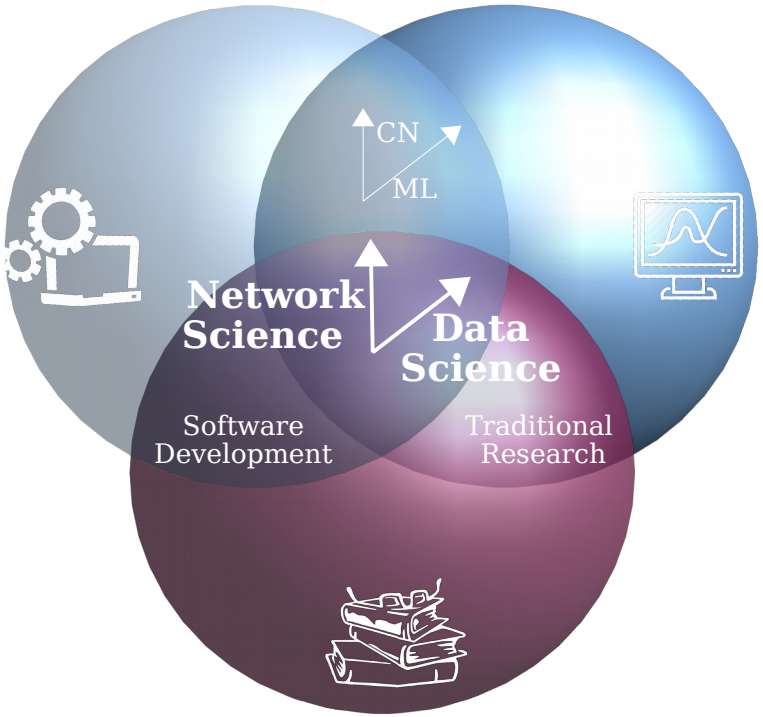


extract a knowledge from the way objects are structured and evolve, and if and how they are connected

Why combine them

In a wide range of domains:

- Objects relationships/networking
- (Objects's) data structure and features



extract a knowledge from the way objects are structured and if and how they are connected
generate models apt to organize the extracted knowledge or mimic the behaviour of experts

Related Work: ML in CN

- development of new machine learning techniques based on complex networks:
  - unsupervised learning technique based on particle competition in complex networks;
  - supervised learning extention of the particle competition model for classification tasks
  - semi-supervised learning with combination of low (physical) and high (patterns identification) level classifiers

Thiago Christiano Silva
Machine learning in complex networks: modeling, analysis, and applications
PhD dissertation (2012)

Related Work: ML meets CN

- characterization of the large variety of real network structures, which are originating from the Big Data
- network geometry for understanding their complex network structures by revealing their hidden metric:
  - intelligent machines for unsupervised recognition and visualization of similarities in big data can be reconducted to hyperbolic model
  - on such hyperbolic model we can apply complex networks theory to discover the hidden metric

Alessandro Muscoloni, Josephine Maria Thomas, Sara Ciucci, Ginestra Bianconi & Carlo Vittorio Cannistraci
Machine learning meets complex networks via coalescent embedding in the hyperbolic space
Nature Communications volume 8, Article number: 1615 (2017)

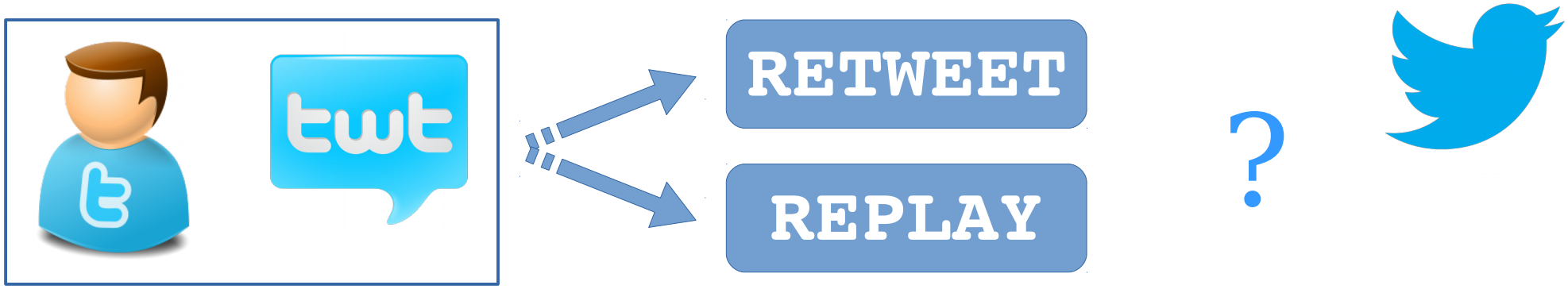
CN to empower ML

- characterization of the large variety of real network structures, which are originating from the Big Data
- understand their complex network structures
- build on top of such structure some “layered” machine learning solution:
  - tradeoff between dynamics control and performance
  - possibility to take into consideration some metrics
  - possibility to have multiple scales

Garg Kamini, Arnaboldi Valerio, Giordano Silvia
“*A Novel Approach to Predict Retweets and Replies Based on Privacy and Complexity-Aware Feature Planes*”
Complex Networks 2016, 30th Nov. - 2nd Dec. 2016, Milano, Italy

Luceri Luca, Torsten Braun, and Giordano Silvia.
"*Social Influence (Deep) Learning for Human Behavior Prediction.*"
CompleNet’18, Mar. 2018, Boston.

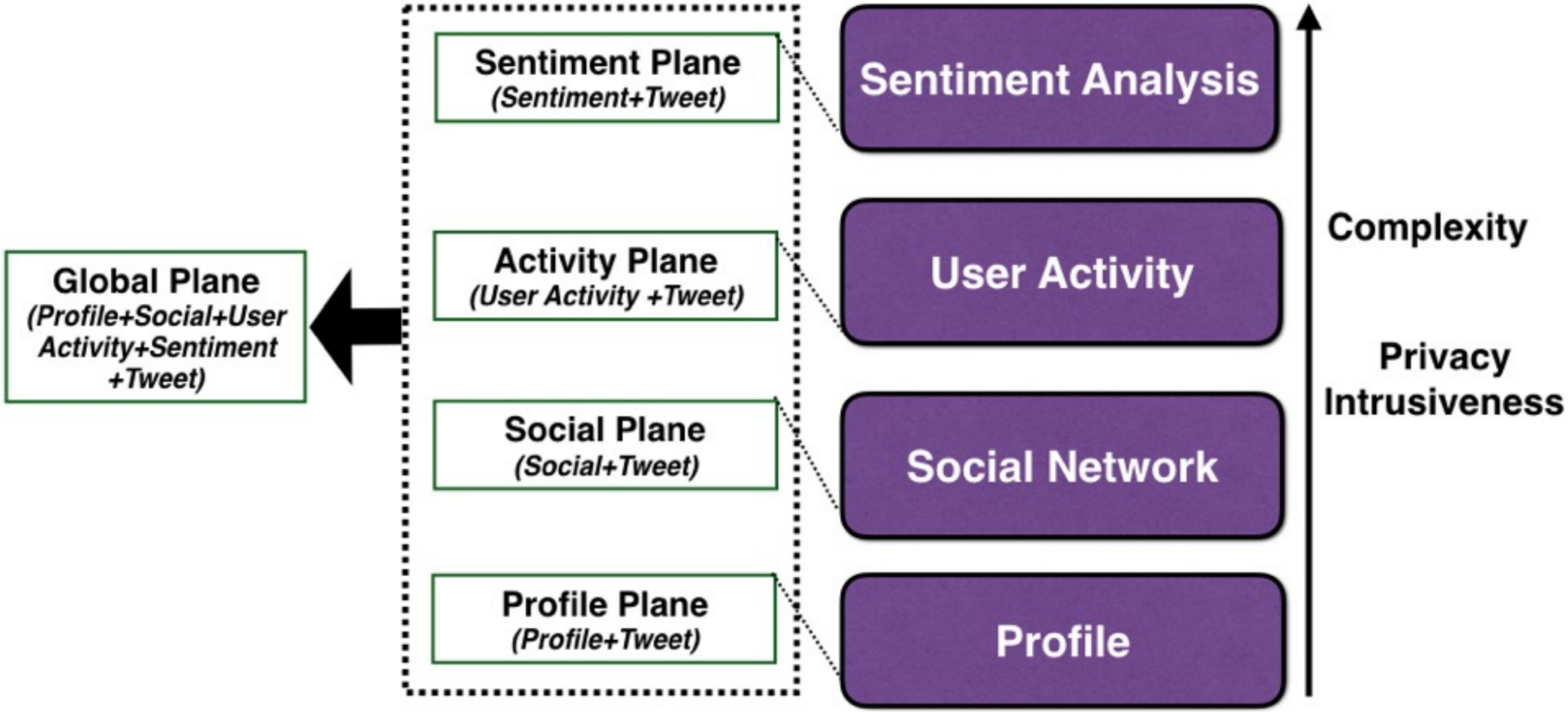
Twitter prediction



- Predict the likelihood to retweet and reply by using different feature planes based on:
  - Complexity to acquire
  - Privacy intrusiveness.
- Predicts the likelihood of tweets received from friends and/or others as well.

Garg Kamini, Arnaboldi Valerio, Giordano Silvia
"*A Novel Approach to Predict Retweets and Replies Based on Privacy and Complexity-Aware Feature Planes*"
Complex Networks 2016, 30th Nov. - 2nd Dec. 2016, Milano, Italy

Twitter prediction: features classification

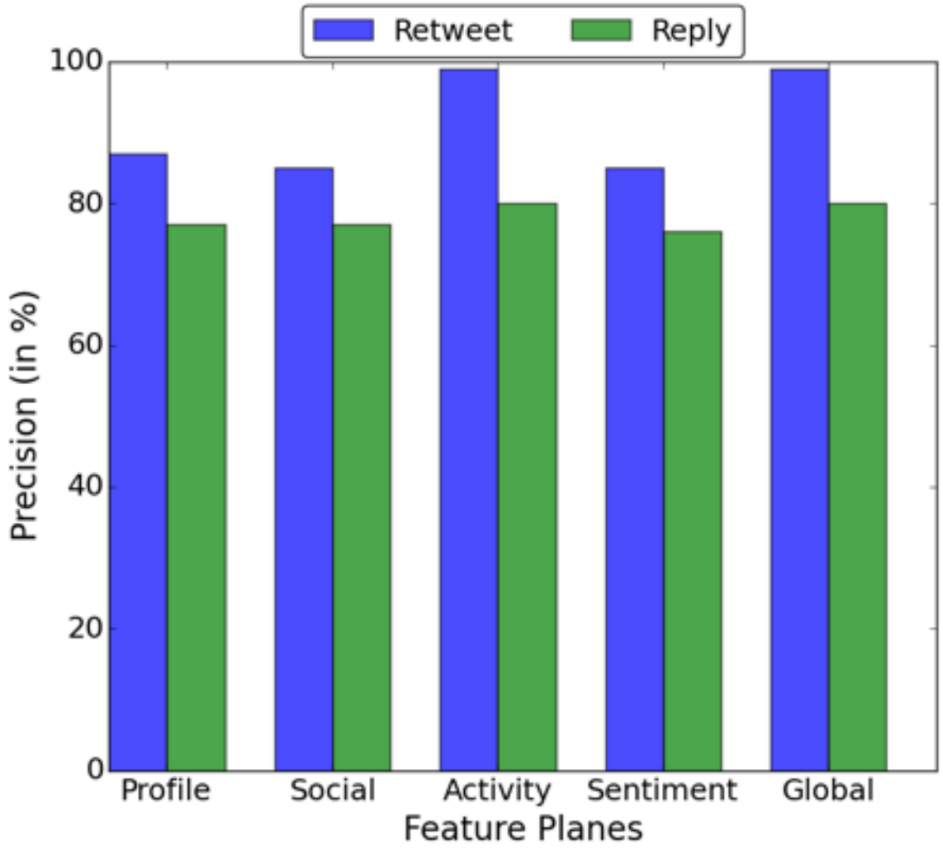


Tweet Features: Time of tweet, retweet count, length of the tweet, # urls, #hashtag

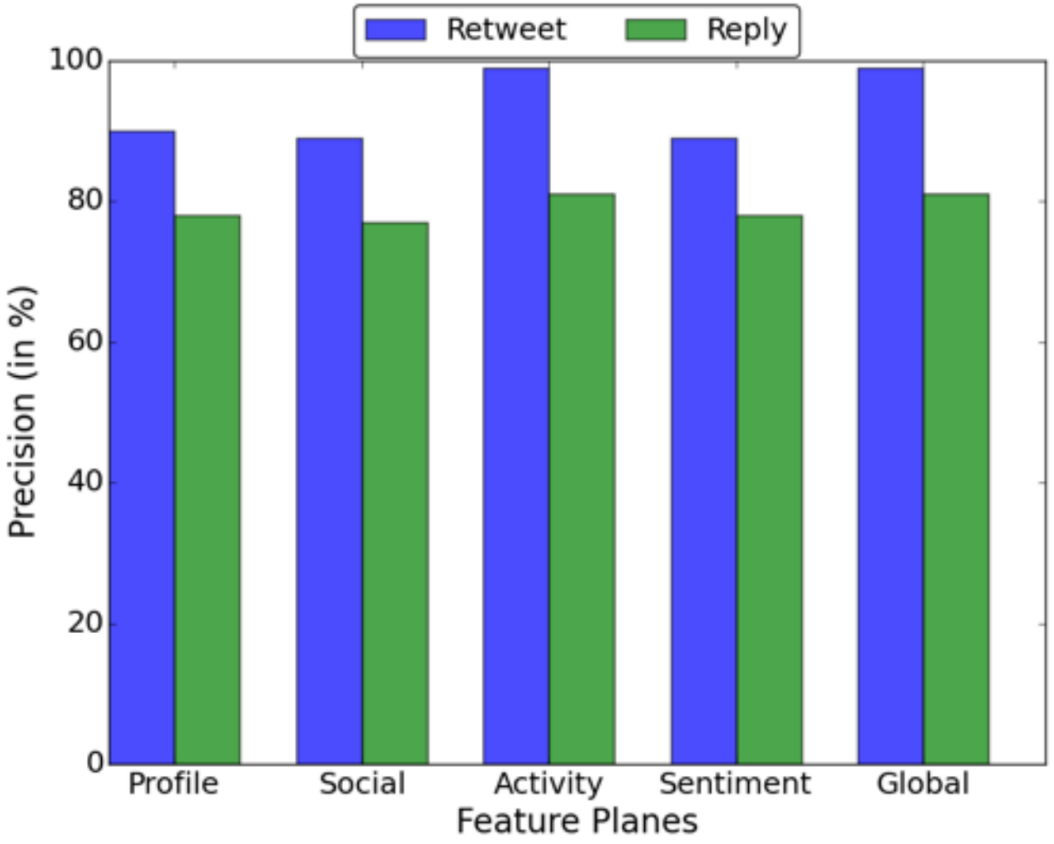
Twitter prediction: machine learning

- Utilized Gradient boosting method: XGBoost Model
- Tested model on two different samples of dataset based on different time-intervals with 673,858 and 1,031,116 tweets respectively.
- Model accuracy was tested for each plane starting from Profile plane to Global plane.

Twitter prediction: results

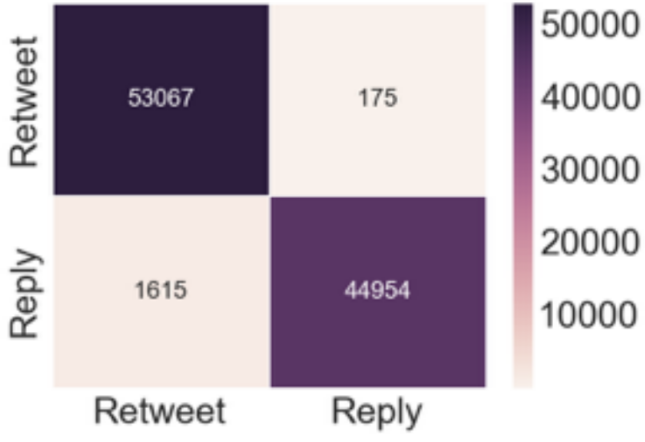


Dataset 1

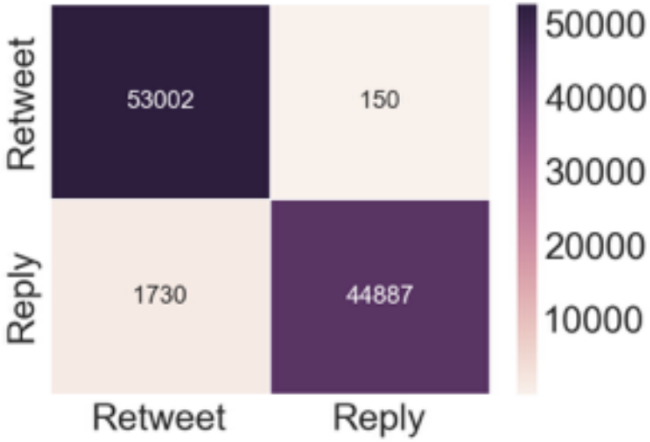


Dataset 2

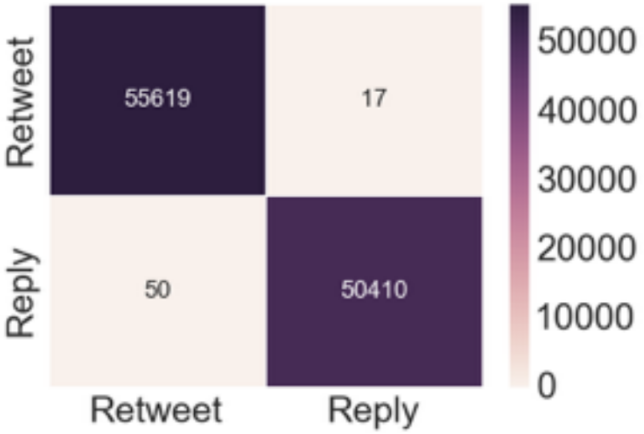
Twitter prediction



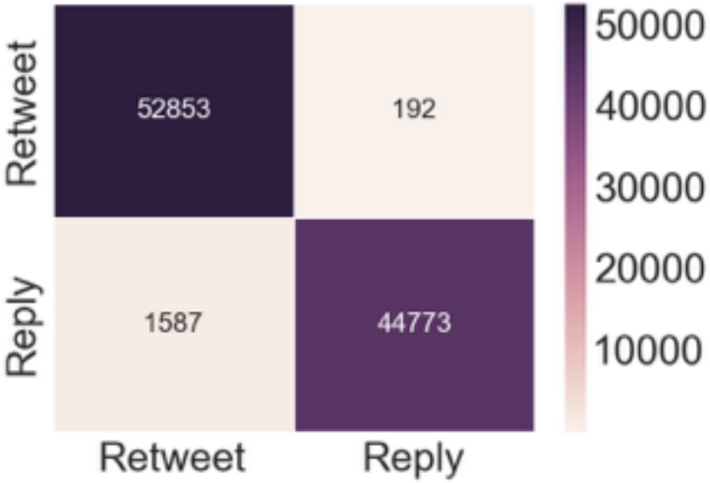
Profile



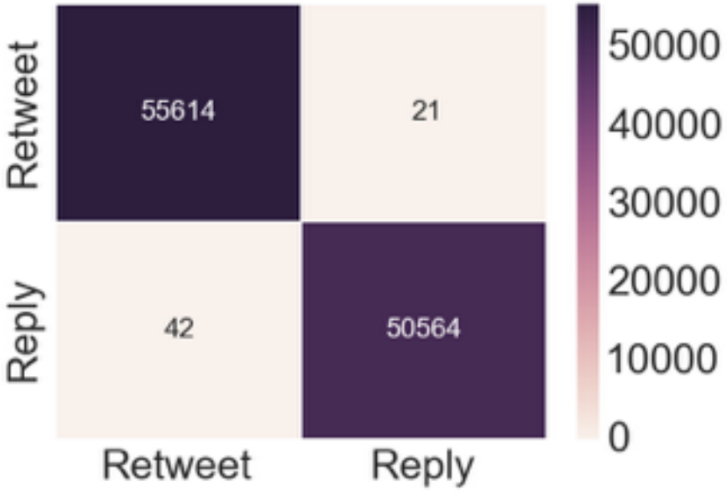
Social



Activity



Sentiment

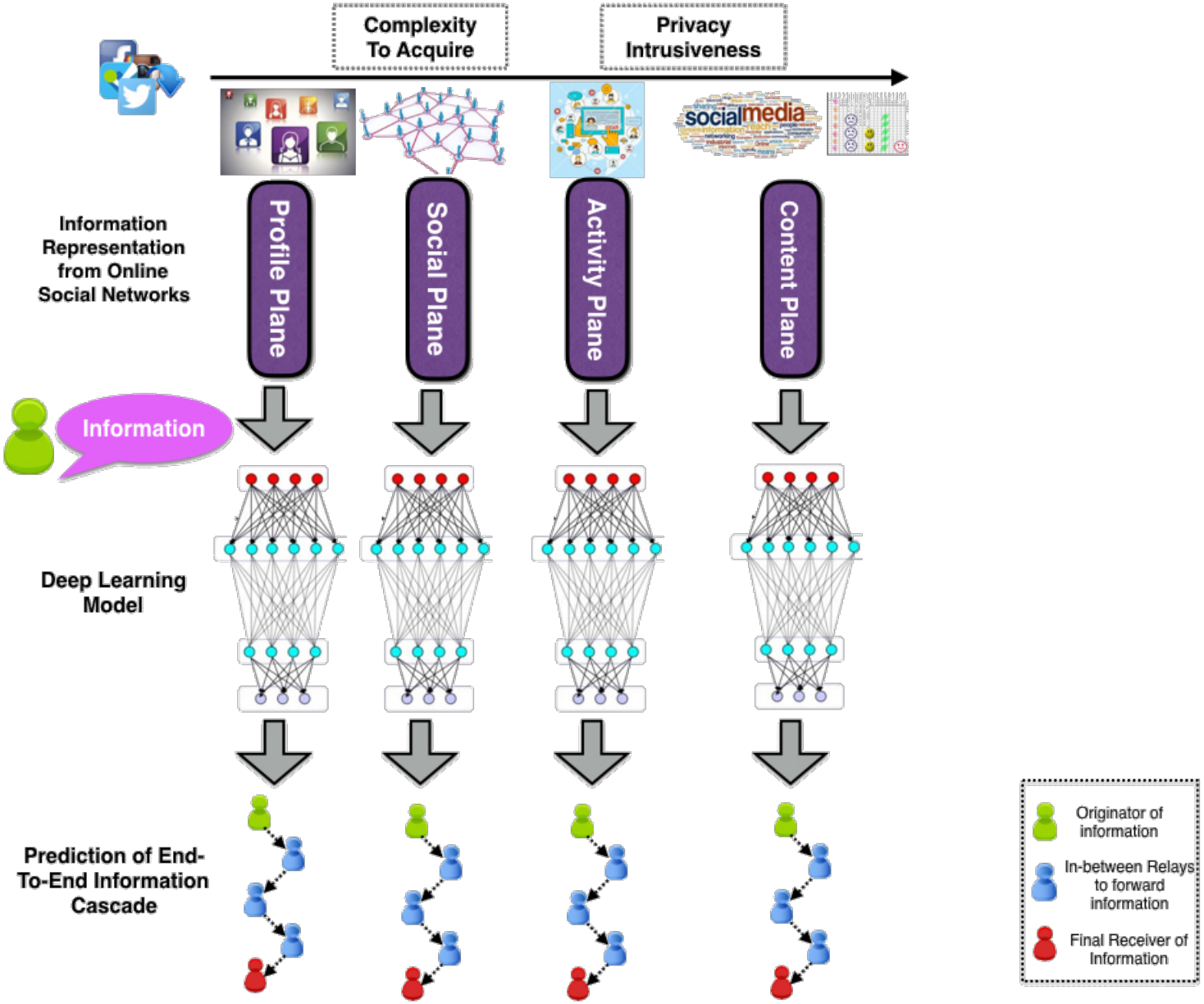


Global

Twitter prediction

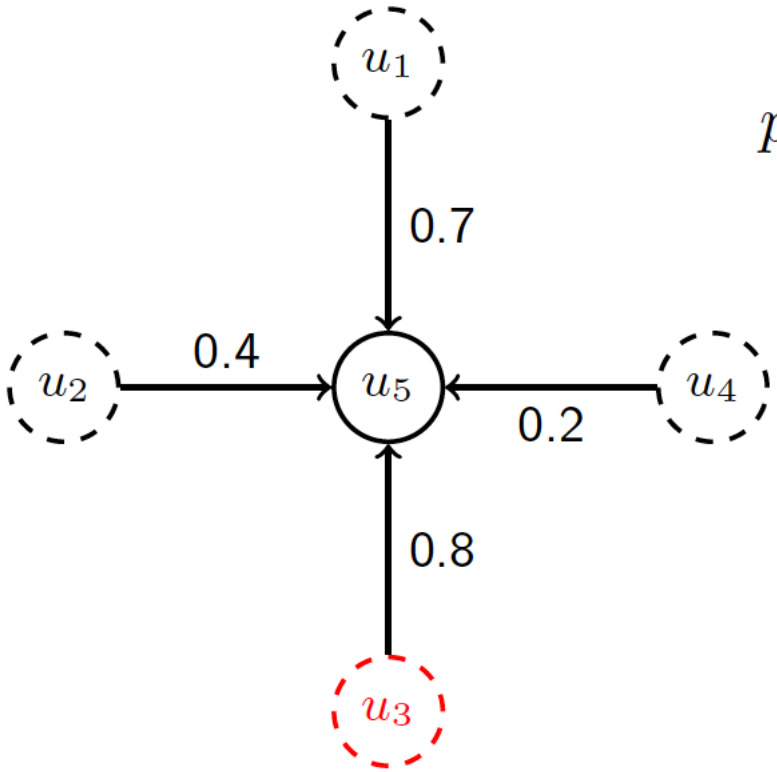
- Predicted the single step diffusion in social networks based on different feature planes (multiple network topologies).
- Our model enables retweet and reply prediction for tweets received from friends and/or others as well.
- Activity plane performs the best. Adding more information does not improve prediction.
- Move toward cascade prediction

Twitter prediction



Event participation prediction

Study the social influence



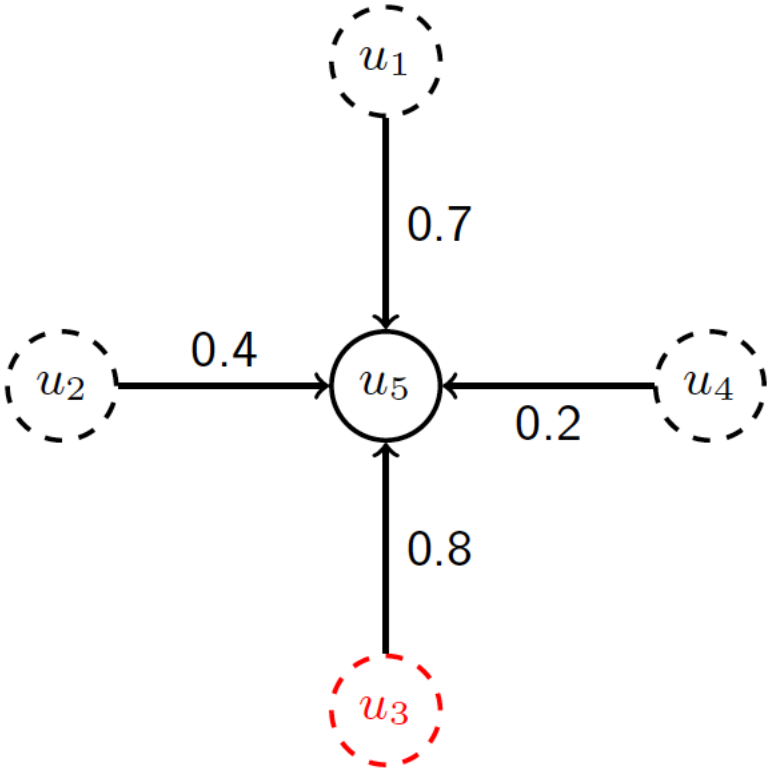
$p_{u_i}(S_{u_i,a})$: probability that subject u_i is influenced by the active friends $S_{u_i,a}$

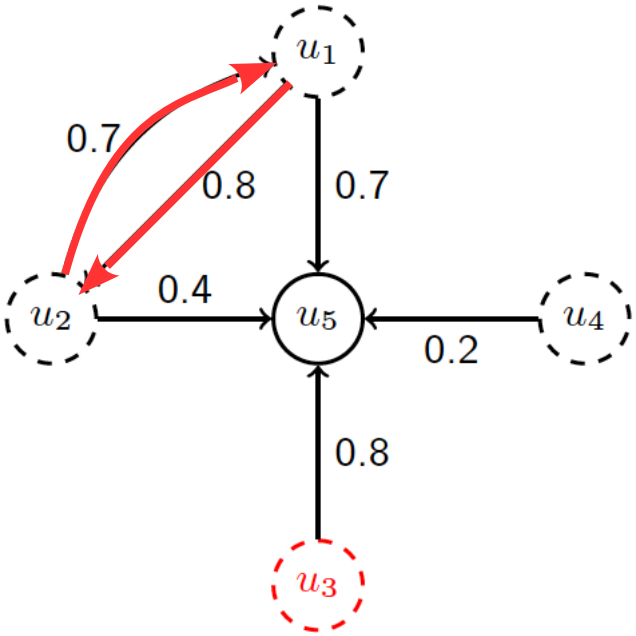
$$p_{u_i}(S_{u_i,a}) = 1 - \prod_{u_j \in S_{u_i,a}} (1 - p_{u_j,u_i})$$

where $p_{u_j,u_i} \in [0,1]$ is the influence probability of u_j on u_i

Luceri Luca, Torsten Braun, and Giordano Silvia.
"*Social Influence (Deep) Learning for Human Behavior Prediction.*"
CompleNet'18, Mar. 2018, Boston.

Event participation prediction

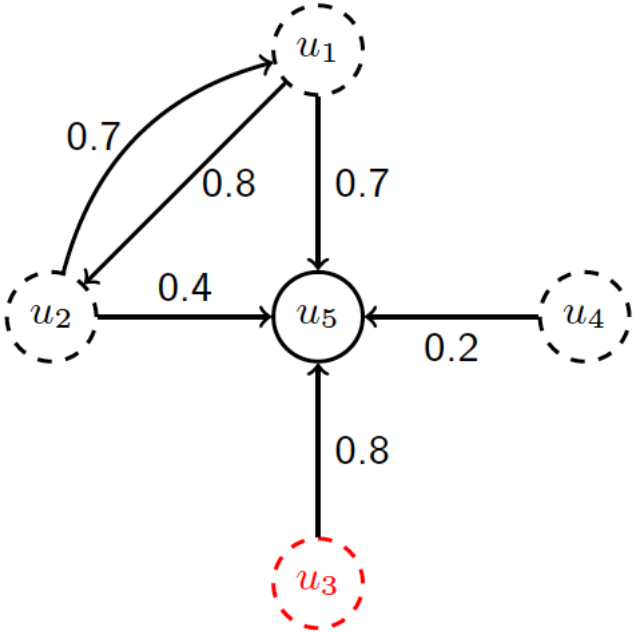


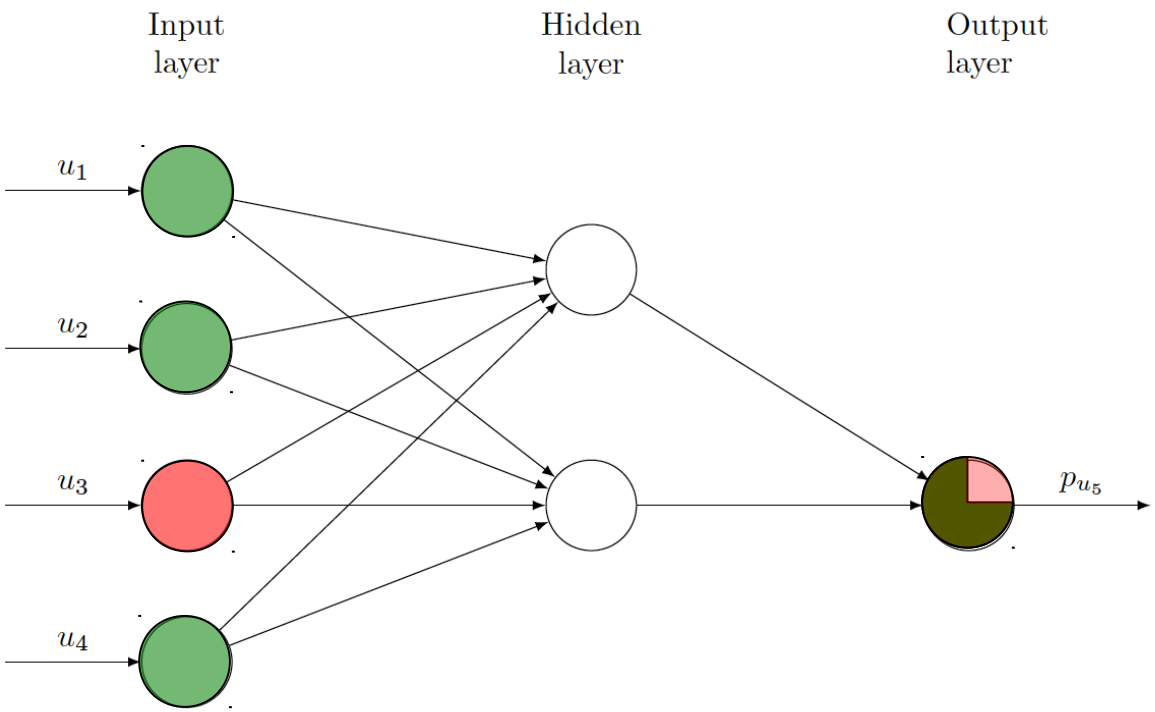


$$p_{u_i}(S_{u_i,a}) = 1 - \prod_{u_j \in S_{u_i,a}} (1 - p_{u_j,u_i})$$

where $p_{u_j,u_i} \in [0,1]$ is the influence probability of u_j on u_i

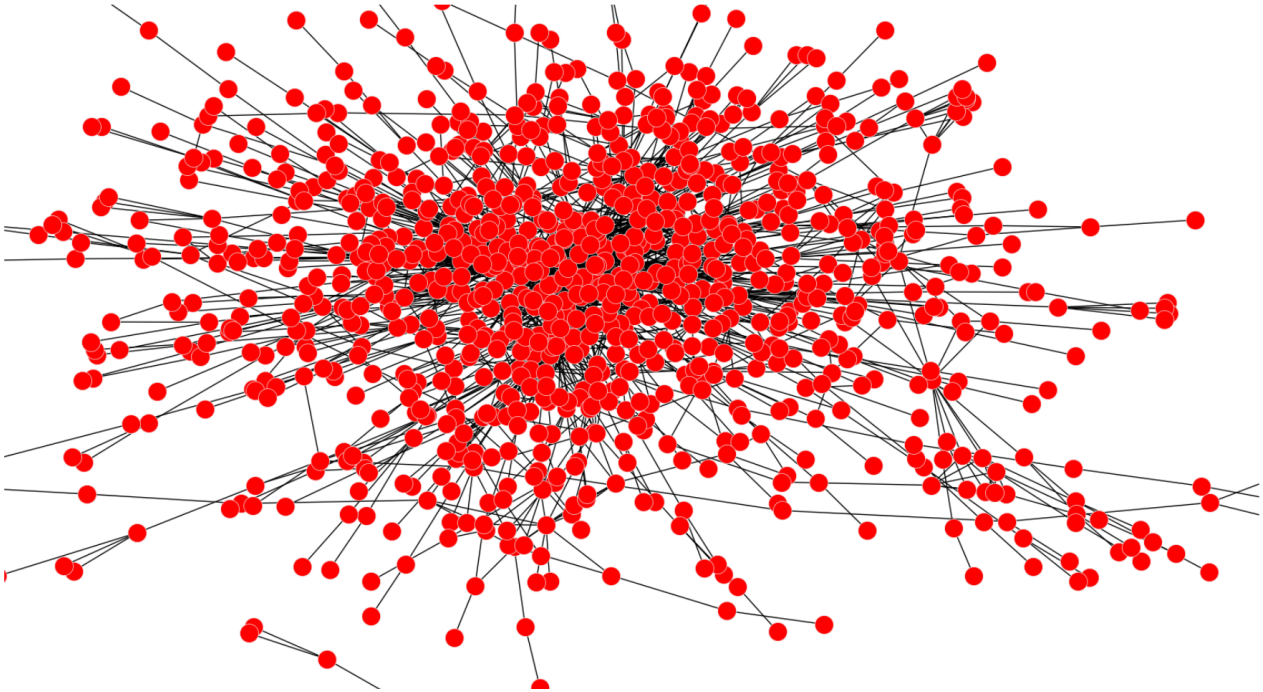
Event participation prediction





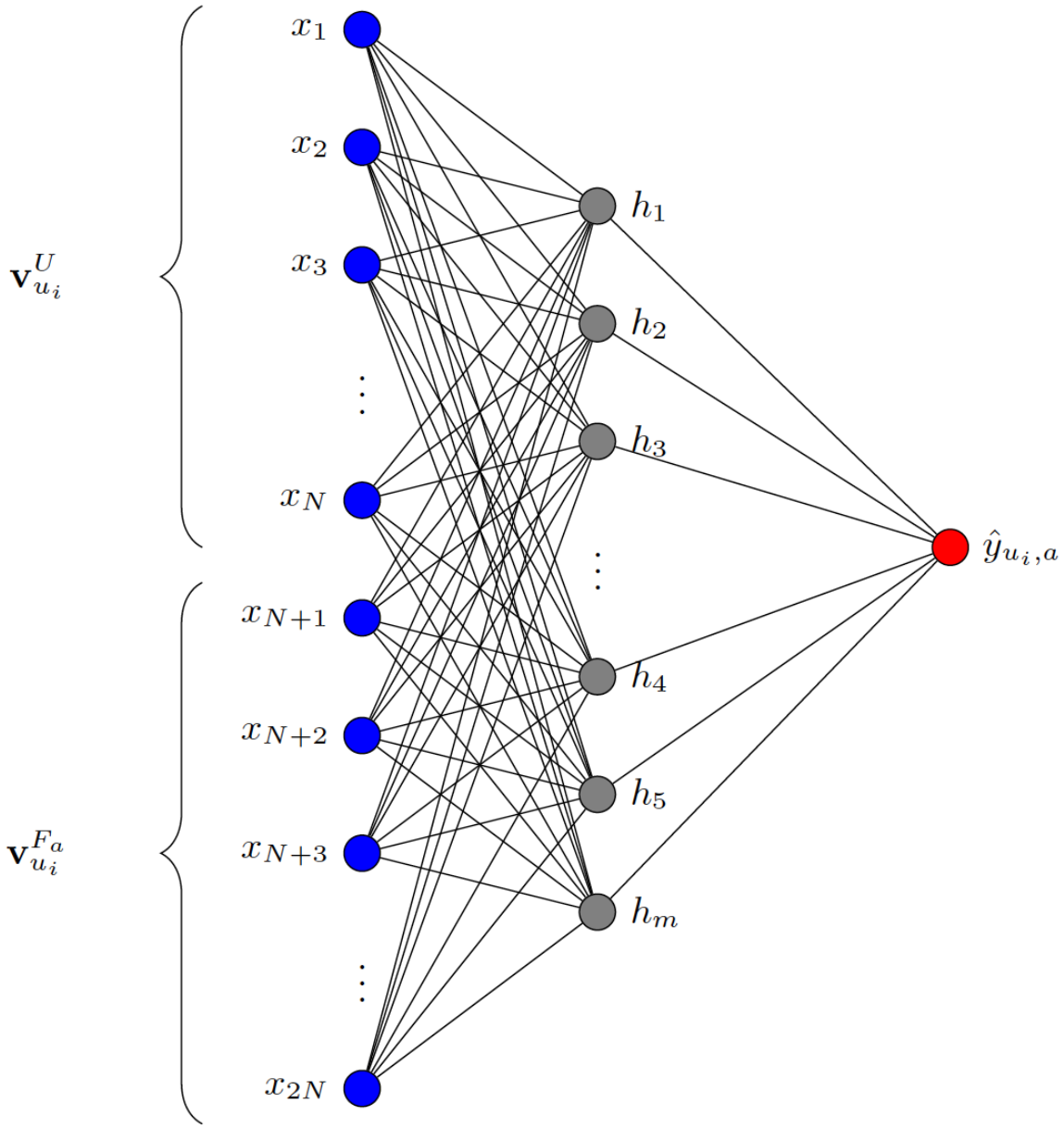
Deep Neural Network (DNN)

Event participation prediction: CN+ML



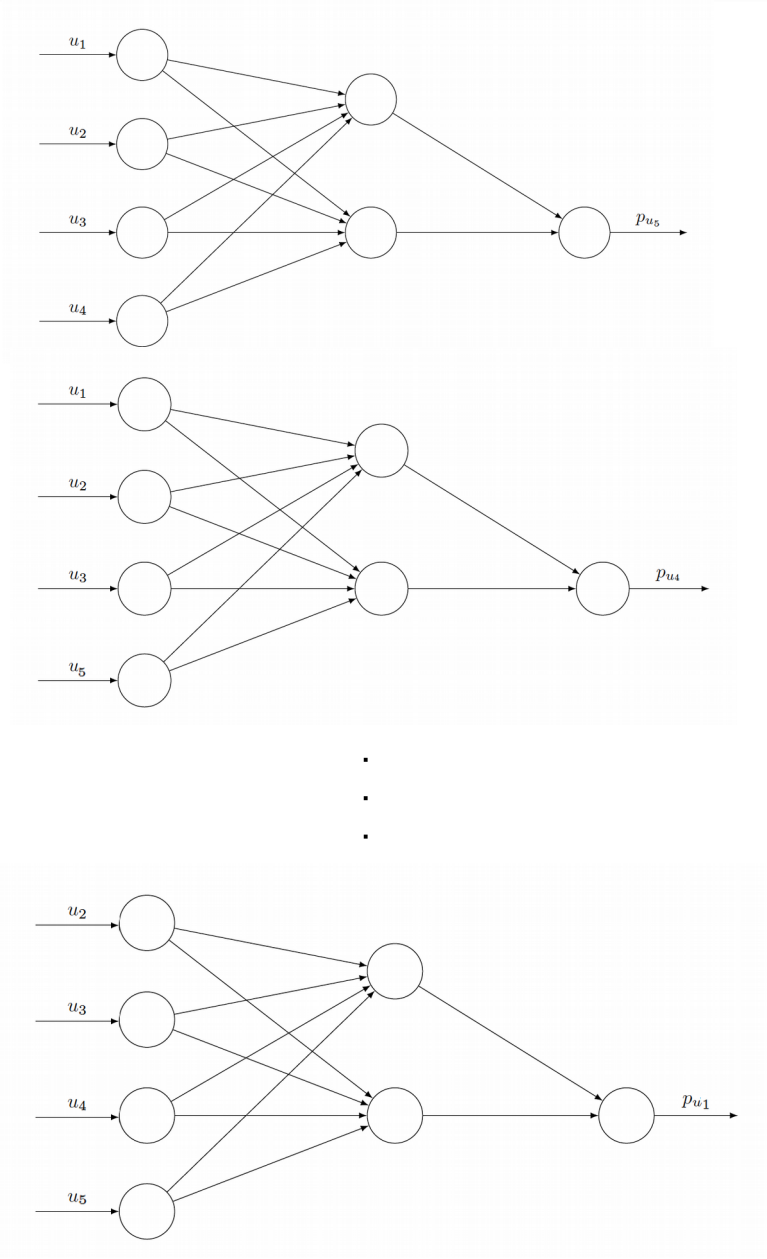
- Monolithic DNN
- Local-DNN (L-DNN)
- Community-based DNN (C-DNN)

Event participation prediction



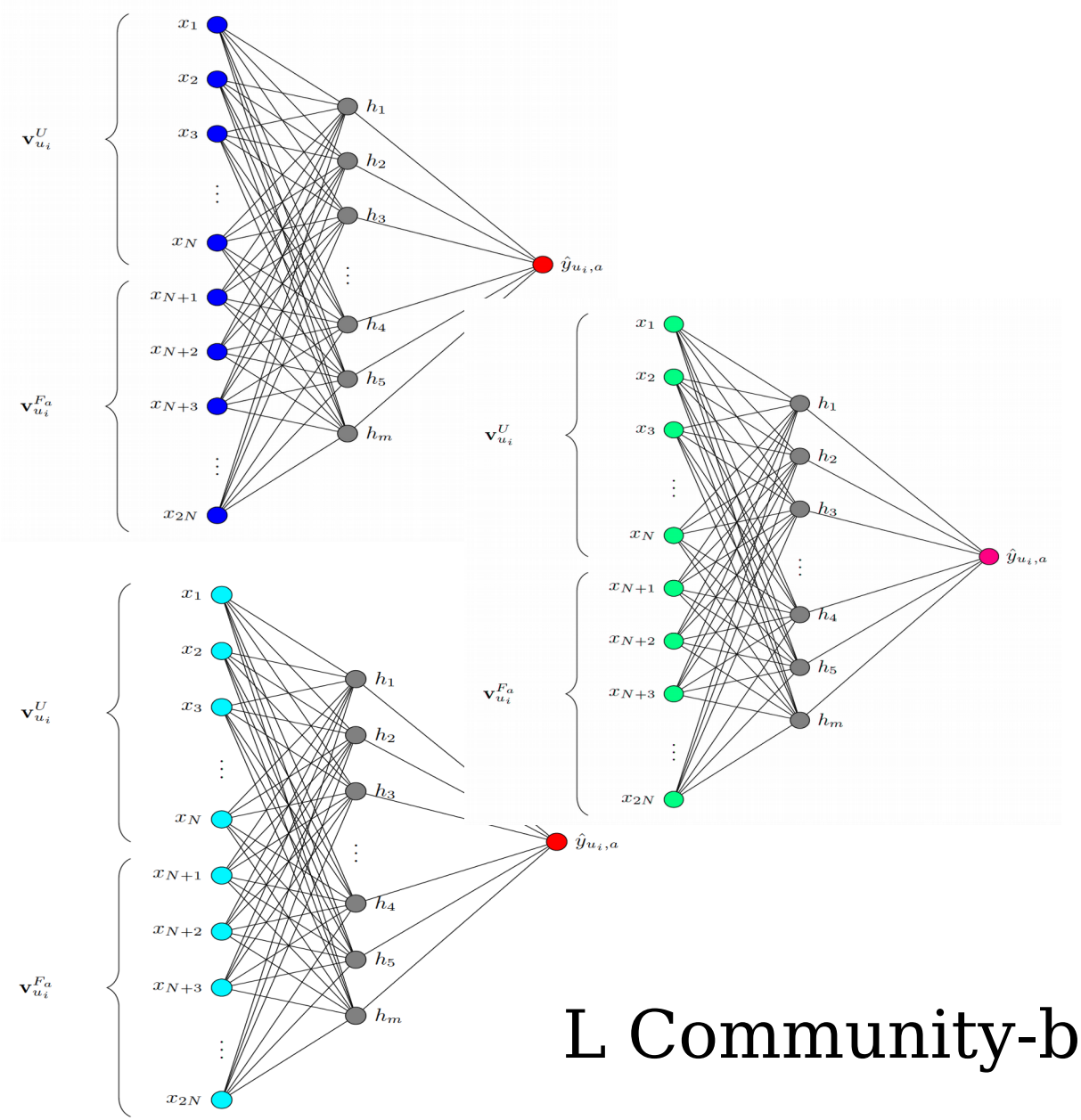
Monolithic DNN

Event participation prediction



N Local-DNN (L-DNN)

Event participation prediction



L Community-based DNN (C-DNN)

Event participation prediction: results

	DNN	L-DNN	C-DNN
Accuracy	85%	80%	82%
# of DNNs	*1*	*N*	*1<L<N*

Event participation prediction

- C-DNN much “understandable” and robust
- C-DNN performance just slightly below M-DNN
- We want to study C-DNN with more types of communities:
  - physical
  - social
  - interest