



User Profiling through Deep Multimodal Fusion

Golnoosh Farnadi

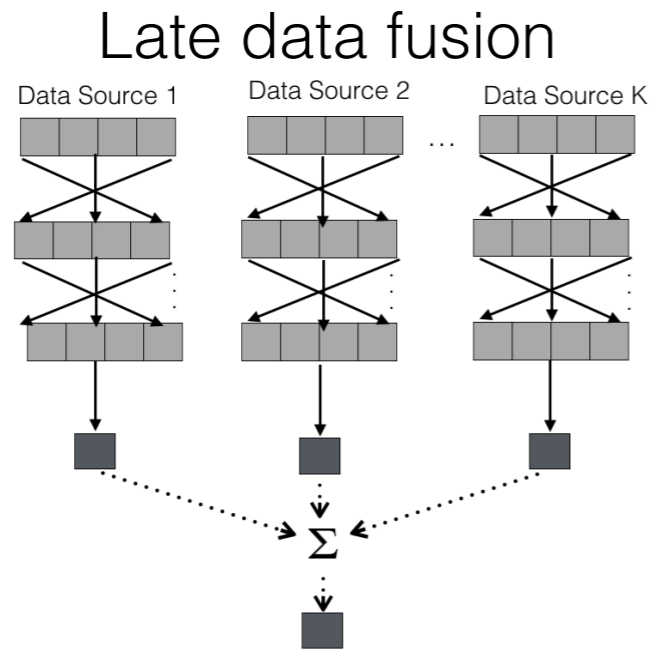
Jie Tang

Martine De Cock

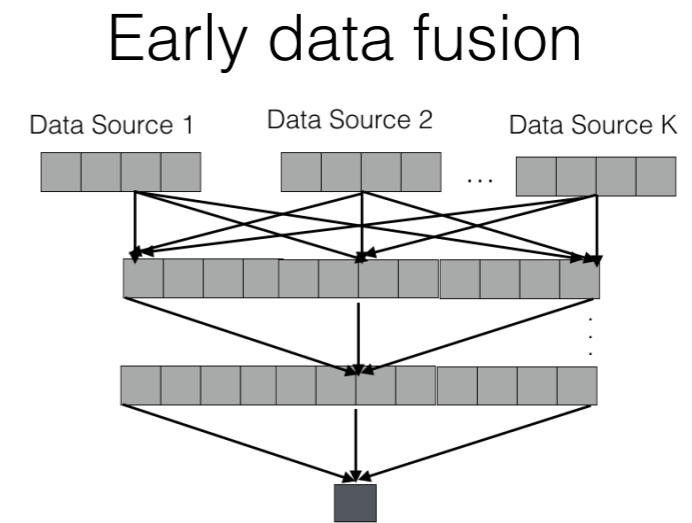
Sien Moens



Data Fusion in DNNs



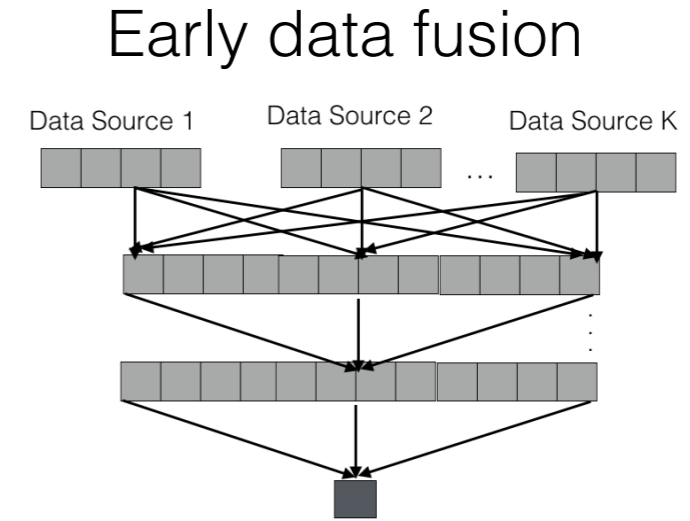
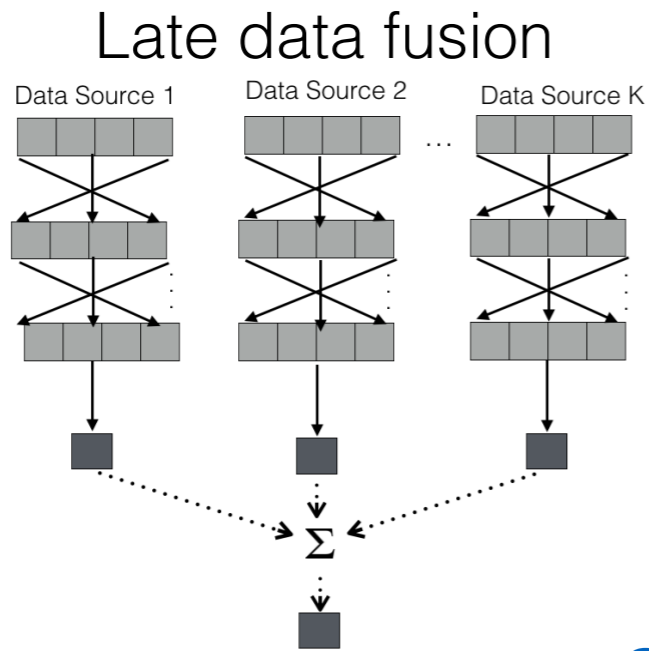
no dependency among data sources



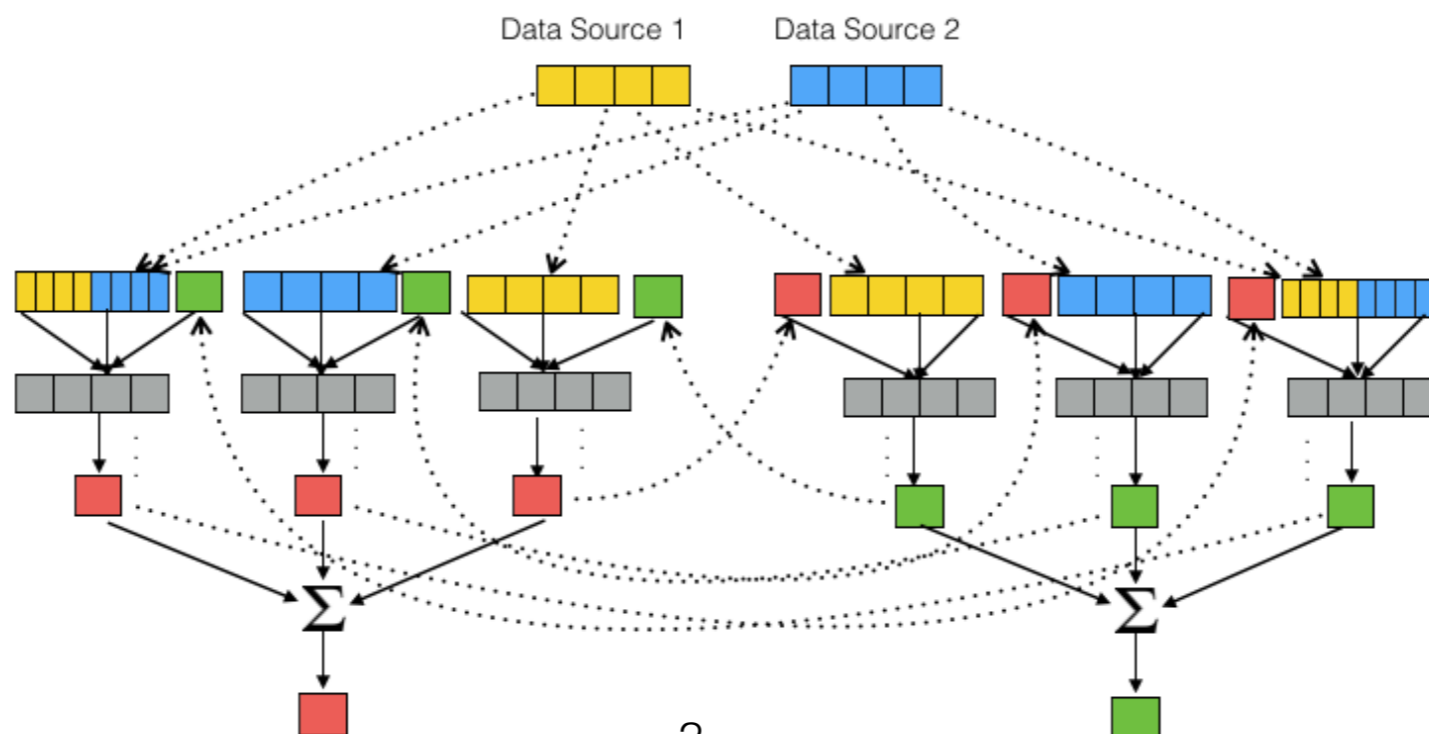
correlation between features



Hybrid Deep Multimedia Fusion



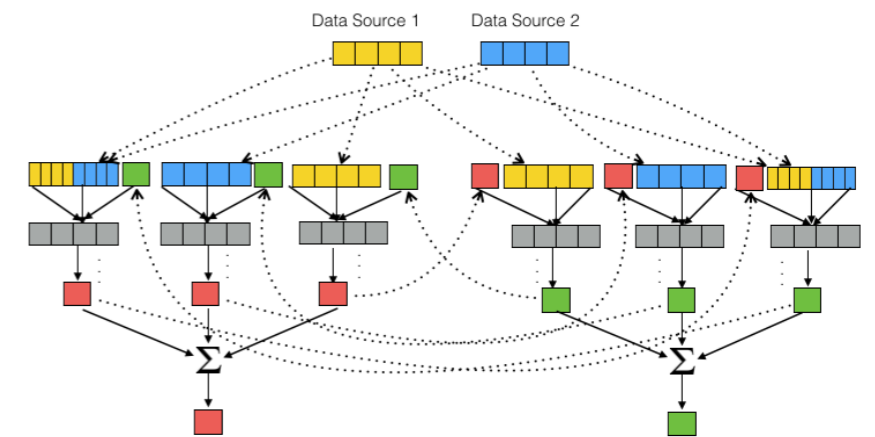
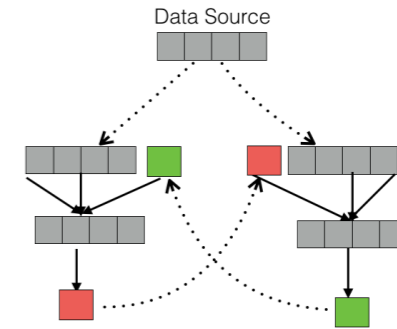
Our hybrid approach



Hybrid Deep Multimedia Fusion

The **stacking** is suitable for **multi-task learning**.

The **powerset combination** incorporates **correlations among features and data sources**



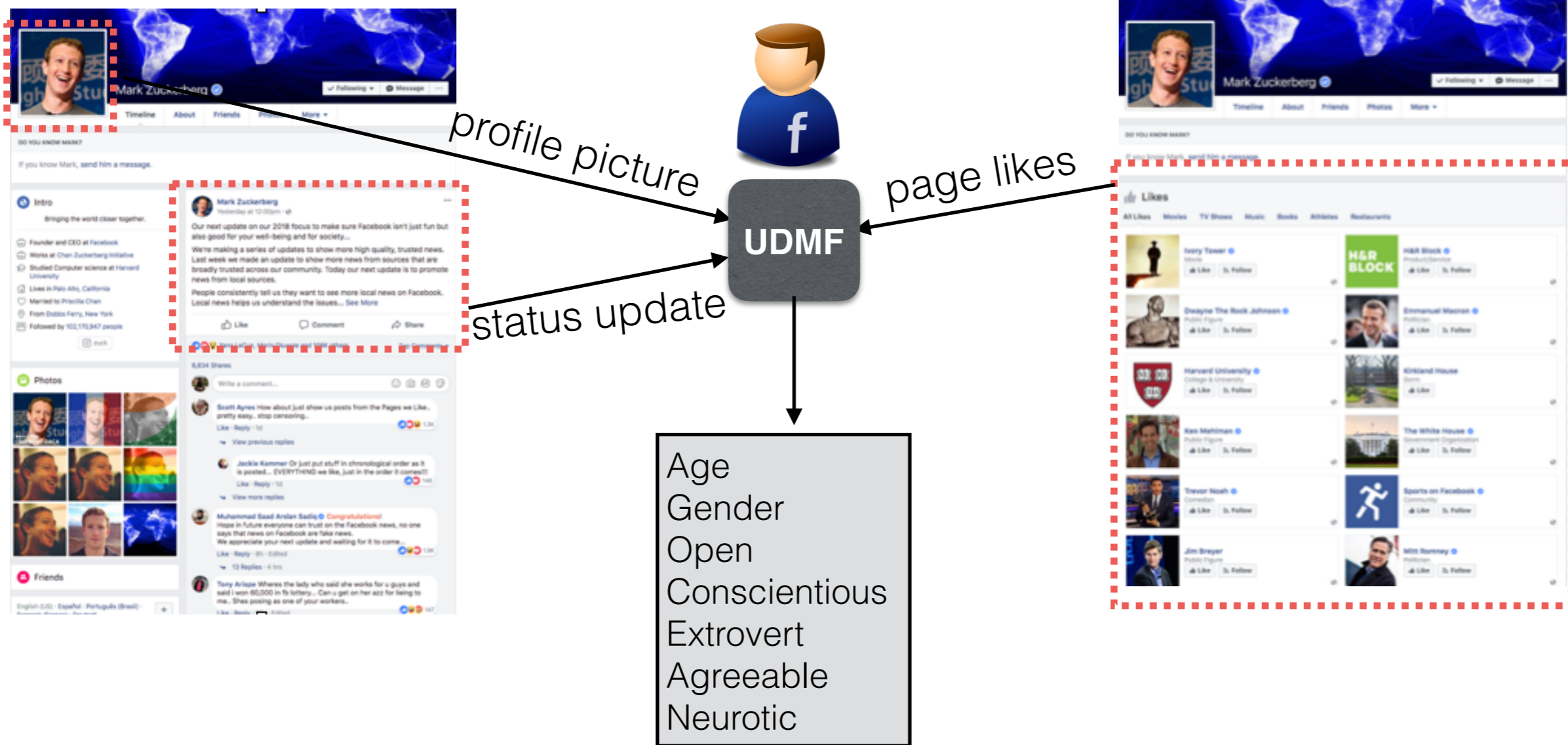
powerset combination

stacking

$$U_i^{0q}(\mathcal{D}) = f\left(\sum_{D \in \mathcal{D}} \sum_j w_{ij} \cdot D_j + \sum_z w_{iz} \cdot \alpha_z \cdot t_z^{q-1}\right)$$

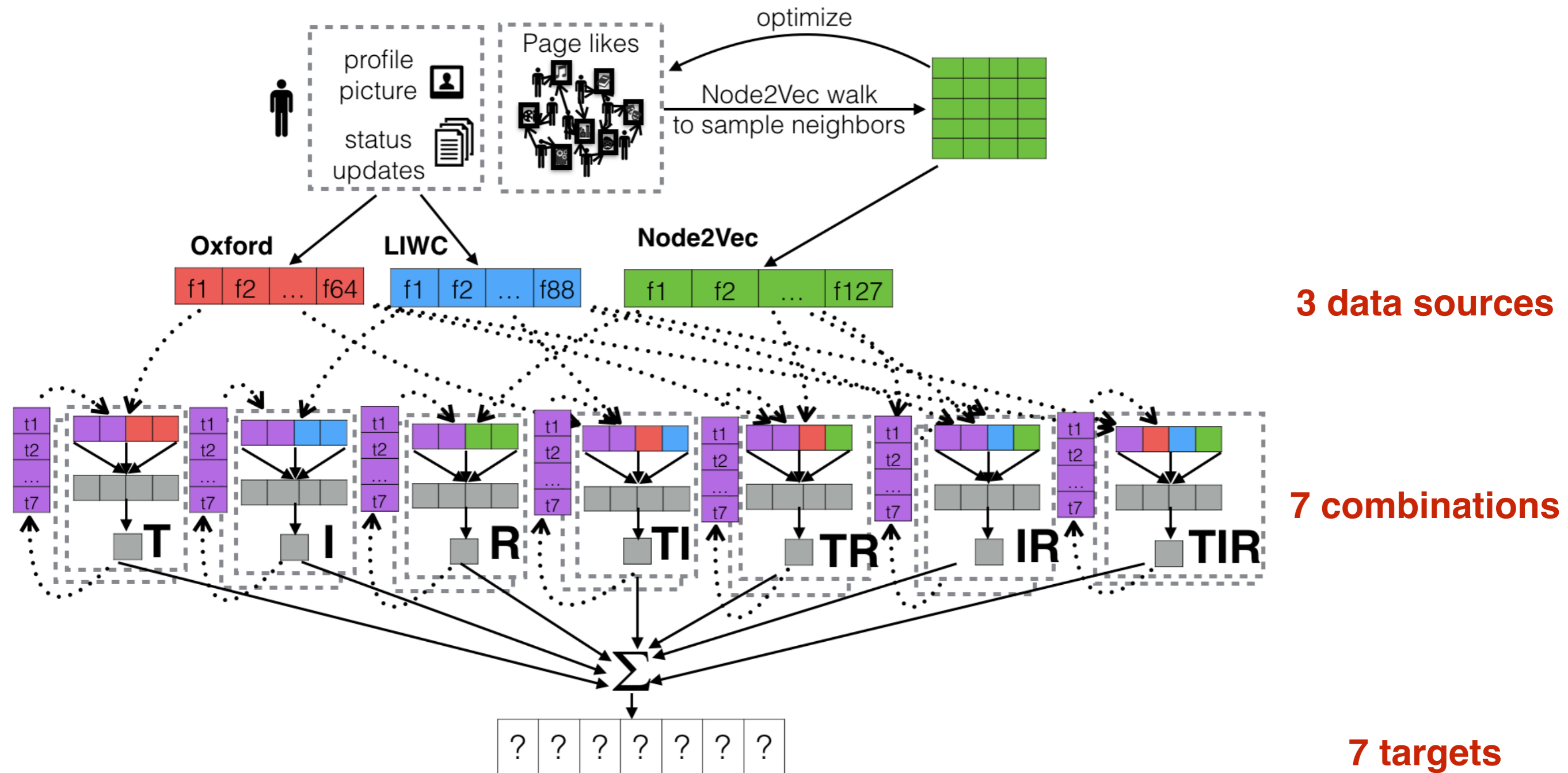
epoch: $U_i^{0q}(\mathcal{D})$
 neuron: $U_i^{0q}(\mathcal{D})$
 data sources: \mathcal{D}
 activation function: $f(\cdot)$
 weight: w_{ij}
 input neuron: j
 weight: w_{iz}
 gate: z
 target variable: t_z^{q-1}
 subset of the input data sources: \mathcal{D}

User profiling



Infer Facebook users' age, gender and personality traits

User Profiling in Facebook using UDMF



Experimental results

- **Data fusion** improves **user profiling**
- **Stacking** works w/o data fusion
- HDMF **outperform** state-of-the-art user profiling and data fusion frameworks
- We accurately predict **age with AUC 0.90** and **gender with 0.96**

Model	Stack	Age	Gender	Opn	Con	Ext	Agr	Neu
Baseline		0.488	0.492	0.502	0.502	0.506	0.506	0.486
One source								
Text	✗	0.741±0.022	0.668±0.020	0.550±0.016	0.575±0.017	0.536±0.016	0.547±0.016	0.523±0.016
	✓	0.748±0.022	0.668±0.020	0.553±0.017	0.574±0.017	0.545±0.016	0.550±0.016	0.524±0.016
Image	✗	0.552±0.016	0.915±0.027	0.502±0.015	0.500±0.015	0.504±0.015	0.512±0.015	0.520±0.016
	✓	0.550±0.016	0.897±0.027	0.516±0.015	0.511±0.015	0.518±0.015	0.519±0.015	0.541±0.016
Relation	✗	0.875±0.026	0.886±0.027	0.601±0.018	0.571±0.017	0.567±0.017	0.525±0.016	0.558±0.017
	✓	0.893±0.027	0.898±0.027	0.622±0.018	0.589±0.018	0.573±0.017	0.533±0.016	0.563±0.016
Two sources								
Early approach TI	✗	0.734±0.022	0.873±0.026	0.569±0.017	0.588±0.018	0.536±0.016	0.545±0.016	0.547±0.016
	✓	0.746±0.022	0.864±0.026	0.546±0.016	0.568±0.017	0.542±0.016	0.546±0.016	0.536±0.016
Early approach TR	✗	0.878±0.026	0.896±0.027	0.610±0.018	0.586±0.018	0.567±0.017	0.535±0.016	0.554±0.017
	✓	0.891±0.027	0.899±0.027	0.627±0.019	0.601±0.019	0.572±0.017	0.551±0.016	0.574±0.017
Early approach IR	✗	0.878±0.026	0.951±0.028	0.606±0.018	0.574±0.017	0.569±0.017	0.524±0.016	0.562±0.017
	✓	0.895±0.027	0.951±0.028	0.633±0.019	0.592±0.018	0.577±0.017	0.537±0.016	0.564±0.017
Three sources								
Ensemble (Late approach)	✗	0.876±0.026	0.952±0.028	0.603±0.018	0.587±0.018	0.569±0.017	0.537±0.016	0.562±0.017
	✓	0.893±0.027	0.949±0.028	0.626±0.019	0.606±0.018	0.582±0.017	0.549±0.016	0.570±0.017
Early approach TIR	✗	0.887±0.027	0.947±0.028	0.617±0.018	0.577±0.017	0.567±0.017	0.541±0.016	0.566±0.017
	✓	0.899±0.027	0.934±0.028	0.635±0.019	0.607±0.018	0.560±0.018	0.551±0.016	0.572±0.017

Model	Age	Gender	Opn	Con	Ext	Agr	Neu
One/Two sources							
Page likes	0.743±0.020	0.699±0.022	0.605±0.017	0.516±0.016	0.555±0.016	0.540±0.016	0.527±0.016
LR (T)	0.711±0.021	0.654±0.020	0.564±0.017	0.568±0.017	0.551±0.016	0.548±0.016	0.530±0.016
LR (I)	0.584±0.017	0.858 ±0.026	0.514±0.015	0.520±0.015	0.528±0.016	0.528±0.016	0.525±0.016
LR(T,I)	0.711±0.017	0.852 ±0.025	0.555±0.017	0.564±0.017	0.551±0.016	0.550±0.016	0.542±0.016
UDMF(T,I)	0.756±0.023	0.886±0.027	0.569±0.017	0.575±0.017	0.552±0.017	0.552±0.016	0.539±0.016
UDMF(T,R)	0.879±0.026	0.943±0.028	0.628±0.019	0.607±0.018	0.580±0.017	0.564±0.017	0.575±0.017
UDMF(I,R)	0.892±0.027	0.955±0.029	0.630±0.019	0.607±0.018	0.587±0.018	0.551±0.016	0.571±0.017
Three sources							
Weighted Soft Voting	0.656±0.019	0.861±0.026	0.523±0.016	0.0523±0.016	0.508±0.015	0.507±0.015	0.518±0.015
Random Forest (100)(T,I,R)	0.786 ±0.023	0.900±0.027	0.588±0.018	0.564 ±0.017	0.544±0.016	0.549±0.016	0.538±0.016
LR(T,I,R)	0.808 ±0.024	0.888±0.027	0.603±0.018	0.585 ±0.018	0.550±0.017	0.550±0.016	0.572±0.017
UDMF(T,I,R)	0.903±0.027	0.956±0.029	0.647±0.019	0.615±0.018	0.592±0.018	0.556±0.017	0.580±0.017

Thank you!

See you soon

Take away messages

- UDMF is a **hybrid multimodal data fusion**
- Deep learning fusion** improves the accuracy of predictions
- Stacking** works when we have multi-task learning (e.g., user profiling)
- Powerset combination** enhances the fusion power
- We are able to predict your **age, gender and personality traits** given your Facebook activities!

User Profiling through Deep Multimodal Fusion

Golnoosh Farnadi^{1,2}, Jie Tang³, Martine De Cock^{1,4}, Marie-Francine Moens⁵
¹ Ghent University, ² University of California, Santa Cruz, ³ Tsinghua University, ⁴ University of Washington Tacoma, ⁵ KU Leuven

User profiling in social media has gained a lot of attention due to its varied set of applications in advertising, marketing, recruiting, and law enforcement. In this study, we propose a deep learning approach that extracts and fuses information across different modalities of user data – such as text, images, and relations.

Data Fusion in DNNs

The main goal of modeling multiple data sources is to integrate two or more sources of data/knowledge and create a single representation that provides a more accurate description of the data sources than any of the individual ones.

Stacking: A neural network where multiple data sources are processed by separate networks, and their outputs are stacked and fed into a final layer.

Power-set Combination: A neural network where data sources are combined at various stages of the network, creating a power-set of possible fusion points.

Early vs. late approaches of integrating multiple data sources in a deep neural network architecture.

$$U_i^{out}(D) = f\left(\sum_{D \in \mathcal{D}} \sum_j W_{ij} \cdot D_j + \sum_z W_{iz} \cdot \alpha_z \cdot f_z^{t-1}\right)$$

epoch, activation function, weight, input neuron, gate, target variable, subset of the input data sources.

Experimental Results: User Profiling

Our sample Facebook dataset includes 49,372 pages, and 724,948 page like relations for 5,670 users.

- Textual data source embedding**: 88 Linguistic Inquiry and Word Count (LIWC).
- Visual data source embedding**: 64 facial features using the Oxford Face API.
- Relational data source embedding**: To represent users with pages that they like, we train an unsupervised deep neural network approach called Node2Vec on our relational graph.

Model	Stack	Age	Gender	Open	Con	Ext	Aggr	Neu
Text sources								
Text	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Image	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Relation	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Visual sources								
Early approach	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
TS	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Early approach	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
TS	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Early approach	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
TS	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Text + Visual sources								
Ensemble	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Late approach	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Early approach	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
TS	✓	0.748±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016

Mean and standard deviation of AUC scores in inferring age, gender and personality traits by fusing two and three data sources in UDMF. All results are averaged over a 10-fold CV. In each column, the highest results are typeset in bold.

Node2Vec features extracted from users' page likes outperform using only pages that users like as features for the tasks of inferring gender, age, and Big Five personality traits.

$$\max_f \sum_{u \in V} \log Pr(NS(u)|f)$$

network neighbor, feature representation

Model	Stack	Age	Gender	Open	Con	Ext	Aggr	Neu
Page Likes								
Page Likes	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
LR (T)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
LR (I)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
LR (C)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
UDMF(TL)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
UDMF(TI)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
UDMF(LR)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Text + Visual sources								
Weighted Soft Voting	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
Random Forest (100) (LR)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
LR (LR)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016
UDMF(TLR)	✓	0.743±0.022	0.888±0.028	0.535±0.017	0.576±0.017	0.536±0.016	0.547±0.016	0.527±0.016

For Further Information
Please contact g.farnadi@ucsc.edu

CENTER FOR DATA SCIENCE
UNIVERSITY OF WASHINGTON TACOMA
Institute of Technology

LEUVEN
UNIVERSITY OF GENT
SANTA CRUZ