User Profiling through Deep Multimodal Fusion

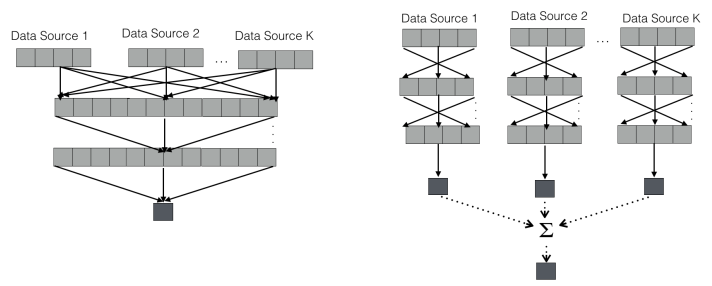
Golnoosh Farnadi[1,2], Jie Tang[3], Martine De Cock [1,4], Marie-Francine Moens[5]

[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

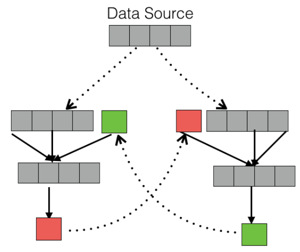
The main goal of modeling multiple data sources is to integrate two or more sources of data/knowledge and create a single representation that provides a more accurate description of the data sources than any of the individual ones.



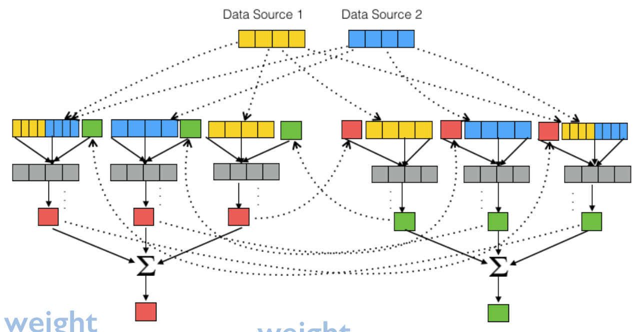
Early vs. late approaches of integrating multipledata sources in a deep neural network architecture.

UDMF

Stacking



Power-set Combination

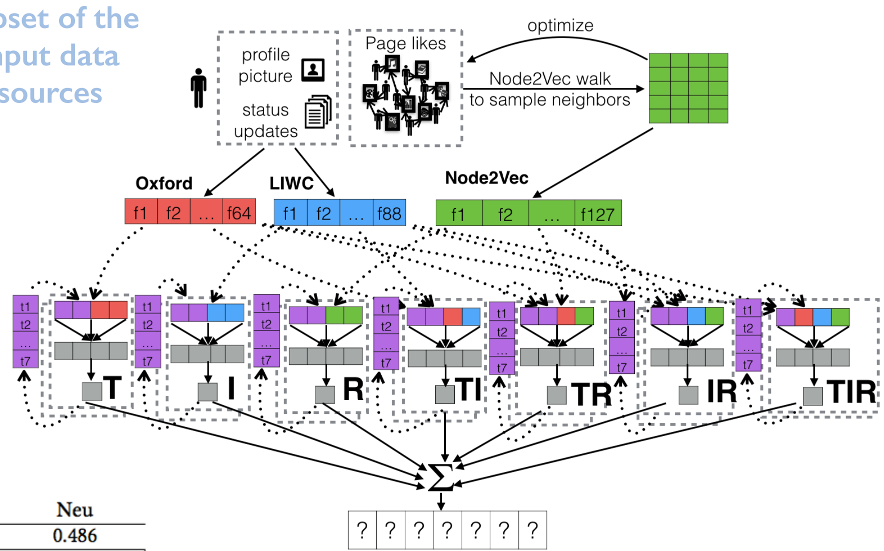


epoch activation function weight weight

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

neuron data sources 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.

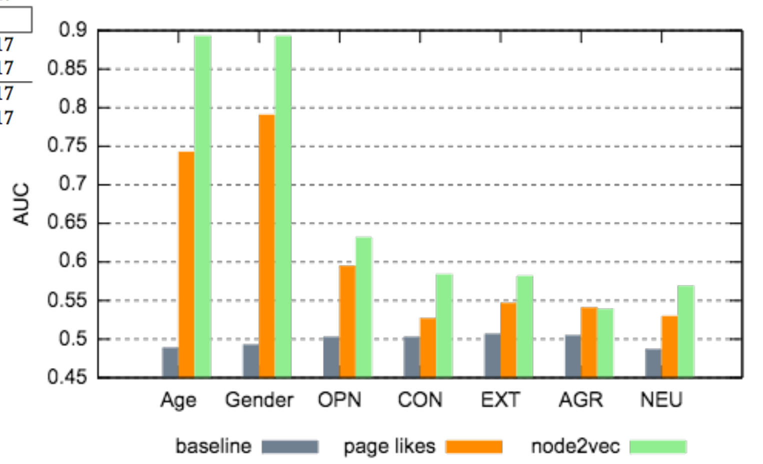
Model	Stack	Age	Gender	Opn	Con	Ext	Agr	Neu
Baseline		0.488	0.492	0.502	0.502	0.506	0.506	0.486
One source								
Text	✗	0.741±0.022	0.668±0.020	0.550±0.016	0.575±0.017	0.536±0.016	0.547±0.016	0.523±0.016
	✓	0.748±0.022	0.668±0.020	0.553±0.017	0.574±0.017	0.545±0.016	0.550±0.016	0.524±0.016
Image	✗	0.552±0.016	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	✗	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
TI	✓	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	✗	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
TR	✓	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	✗	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
IR	✓	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	✗	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
(Late approach)	✓	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	✗	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
TIR	✓	**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

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



- Has shared representation between modalities to integrate three sources of data at the **feature level**,

- Combines the decision of separate networks that operate on each combination of data sources at the **decision level**.

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±0.0161	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**

For Further Information

Please contact gfarnadi@ucsc.edu.

CENTER FOR DATA SCIENCE
UNIVERSITY of WASHINGTON | TACOMA
Institute of Technology
KATHOLIEKE UNIVERSITEIT LEUVEN
UNIVERSITEIT GENT
UNIVERSITY OF CALIFORNIA SANTA CRUZ