Importance of features using Artificial Neural Networks: my journey within two paradigms

F Velásquez-Rojas[1,2*], JE Fajardo, D Zacharías and MF Laguna[2]
F Velásquez Rojas[1,2*] and JJ Ramasco[1]

[1]IFISC (CSIC - UIB) Palma de Mallorca – Spain.
[2]CONICET - UNLP, Argentina.

*fatima@ifisc.uib-csic.es

Introduction

Artificial Neural networks (ANNs) have proven their usefulness in contexts as diverse as one can imagine. They've been applied from social networks to particle physics research and a virtually endless variety of applications such as nuclear fusion research, medical imaging improvement and recognition, stock market predictions, voice recognition, translation, among many others. In addition to these multiple applications, ANNs can give us an idea of how important the characteristics or parameters defining a given model are and how much they contribute to the results obtained. In this poster, I will try to summarize what has been my journey within two different paradigms using this method: my last work which is related to complex systems and higher education and my current work at the IFISC which is related to human mobility and the most important features involved in this phenomenon.

Effects of the COVID-19 pandemic in higher education †

Classroom: Individuals who influence each other → Dynamics of collective behaviour → Complex Systems

... the COVID-19 pandemic abruptly changed the classroom context ...

OUR GOAL

To analyze the Knowledge Acquisition (KA) process in two different contexts: face-to-face (before the onset of the pandemic) and virtual (during confinement.)

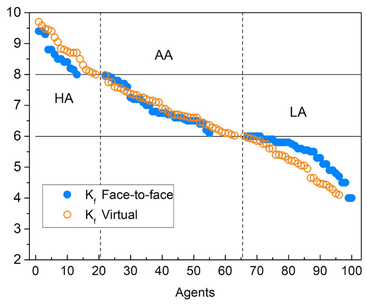
EDUCATIONAL CONTEXT

- *National University of La Plata (UNLP), Argentina, School of Engineering*
- *Physics II course*
- *4 semesters, 8 sections (2019 - 2020)*
- *N = 173 students participating in the whole process*

GROUP CLASSIFICATION*

- *High-achieving (**HA**) students ($8 \leq K_f \leq 10$)*
- *Average-achieving (**AA**) students ($6 < K_f < 8$)*
- *Low-achieving (**LA**) students ($K_f \leq 6$)*

Context	HA	AA	LA	Total
Face-to-face	13	34	34	81
Virtual	19	42	31	92
Total	32	76	65	173



KNOWLEDGE ACQUISITION MODEL**

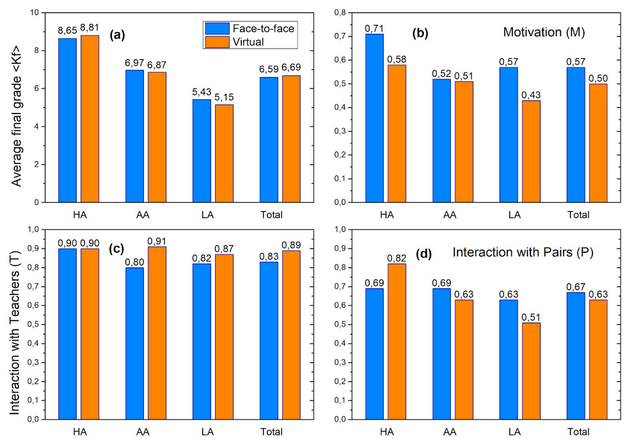
$$K_f^i = \beta_M^X M^i + \beta_T^X T^i + \beta_P^X P^i$$

Final knowledge acquired by individual *i* in the course

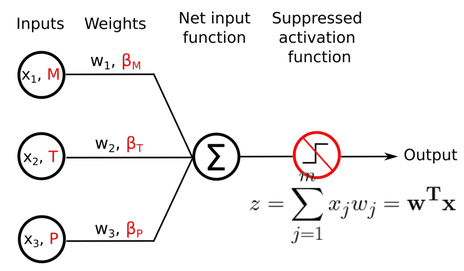
$X = HA, AA, LA$

M: Motivation
T: Interaction with teachers
P: Interaction with pairs
□: Weights for each term

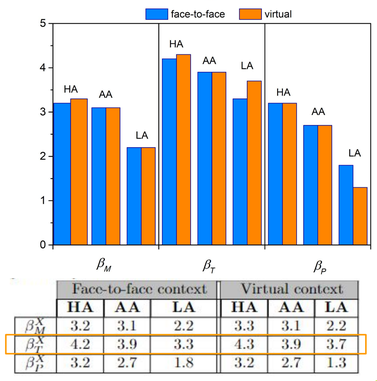
Quantities	relative change (%) HA	AA	LA	Total
K_f	1.8	-1.5	-5.1	1.5
M	-17.5	-2.2	-23.7	-12.3
T	-0,1	12.8	5.9	7.1
P	17.8	-8.5	-19.7	-5.8



SINGLE LAYER PERCEPTRON (SLP) NETWORK



$$K_f^i = \beta_M^X M^i + \beta_T^X T^i + \beta_P^X P^i$$



	Face-to-face context HA	AA	LA	Virtual context HA	AA	LA
β_M^X	3.2	3.1	2.2	3.3	3.1	2.2
β_T^X	4.2	3.9	3.3	4.3	3.9	3.7
β_P^X	3.2	2.7	1.8	3.2	2.7	1.3

MULTIPLE LINEAR REGRESSION (MLR)

$$K_f = \beta_M M + \beta_T T + \beta_P P + \gamma_{HA} G_{HA} + \gamma_{LA} G_{LA} + \alpha_C C_P + \varepsilon$$

	β	SE	p-value
M	2.2776	0.3768	1.00e-08
T	4.6893	0.3063	< 2e-16
P	1.4660	0.2308	12.09e-09
Student group according to their final achievements K_f (Reference → AA)			
HA	1.9131	0.2664	2.43e-11
LA	-1.0585	0.2136	1.81e-06
Context (Reference → Virtual context)			
Face-to-face context	0.7424	0.1902	0.000139
		Adjusted R-squared	0.9664
		p-value	< 2.2e-16

References:
† *F Velásquez-Rojas, JE Fajardo, D Zacharías and MF Laguna. Effects of the COVID-19 pandemic in higher education: a particular case from the perspective of complex systems. arxiv: https://arxiv.org/abs/2203.05819 (2022).*
** Bordogna & Albano EV. Theoretical description of teaching-learning processes: A multidisciplinary approach. Physical Review Letters. 2001; 87(11):118701.*
*** Velásquez-Rojas F & Laguna MF. The knowledge acquisition process from a complex system perspective: observations and models. Nonlinear Dynamics, Psychology, and Life Sciences. 2021; 25(1):41–67.*

Finding the most determining factor in human mobility

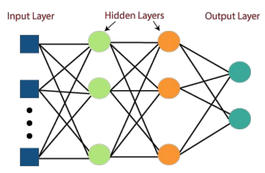
Human mobility: Movement of human beings (individuals, groups) in space and time → Different patterns → Applications: migratory flows, traffic forecasting, urban planning and epidemic modeling [1].

... extensive data to explain human mobility: geographic variables, distances, populations...

OUR GOAL

To understand what is the most important characteristic (or feature) that intervenes in human mobility using Artificial Neural Networks.

GRAVITY MODEL $w_{ij} \propto \frac{P_i P_j}{r_{ij}}$ + DEEP ANN = DEEP GRAVITY MODEL

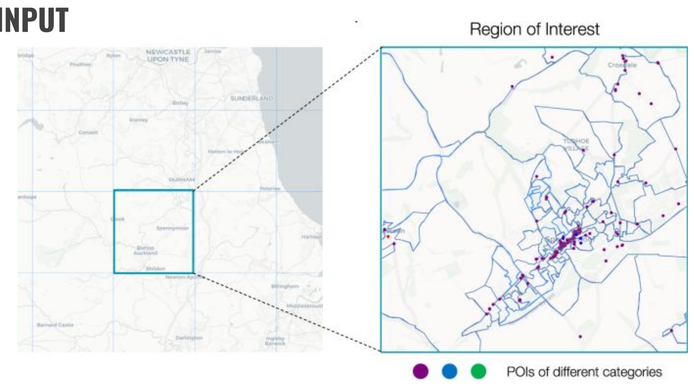


More realistic data → More realistic mobility flows

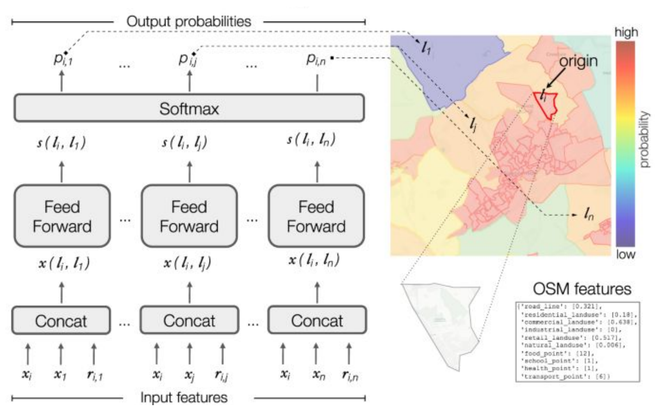
PREPARATION OF GEOGRAPHIC DATA INPUT

- *The geographic space is divided into regions of interest (squared tiles).*
- *Each region of interest is split into locations.*
- *Each location may have several Points Of Interest (POI's) obtained by OpenStreetMap.*

Image credit: Reference [2]



MULTI-LAYER PERCEPTRON (MLP) NETWORK



- *The input features x_i (feature vector of the origin location l_i), x_j (feature vector of the destination location l_j), and $r_{i,j}$ (distance between origin and destination) are concatenated to obtain the input vectors $x(l_i, l_j)$.*
- *The output of the last hidden layer is a score $s(l_i, l_j) \in [-\infty, +\infty]$. The higher this score for a pair of locations (l_i, l_j), the higher the probability to observe a trip from l_i to l_j (estimating flows).*
- *To estimate how the input geographic features contribute to determine the output, we propose different techniques as SHapley Additive exPlanations (SHAP) [3] or Random forest (ensembles of decision trees).*

Image credit: Reference [2]

References:
[1] *Barbosa, H. et al. Human mobility: models and applications. Phys. Rep. 734, 1–74 (2018).*
[2] *F. Simini, G. Barlacchi, M. Luca, L. Pappalardo, A Deep Gravity model for mobility flows generation, Nature Communications 12, 6576 (2021).*
[3] *Štrumbelj, E. & Kononenko, I. Explaining prediction models and individual predictions with feature contributions. Knowl. Inf. Syst. 41, 647–665 (2014)*

Conclusions and Perspectives

■**Effects of the COVID-19 pandemic in higher education**:
- We analyze various quantities that participate in the knowledge acquisition process in face-to-face and virtual contexts for a specific case of study. The shift to virtuality reflected a lack of motivation to learn and a change in the way students interact with pairs and teachers.
- We use ANN and MLR to know the weight of the different factors considered: in all cases, interaction with teachers is of utmost importance in the process of acquiring knowledge. In both contexts the weights are similar.

■**Finding the most determining factor in human mobility**: Ongoing work + Add different mobility types: week and weekend days.