



# Sara Malvar, PhD

## PERSONAL PROFILE

Electrical engineer with 7+ years of experience in data analysis and machine learning. Certified professional with university teaching and distance learning experience, predictive analysis for business and data mining. I have more than 13 publications in international conferences and journals. Worked as a consultant for large companies such as FastShop, Shell, Symrise, IBM and acts as a reviewer and mentor at Udacity, in addition to producing Machine Learning courses for Alura, Data Science Dojo, and mentoring Awari Data Science students.

## WORK EXPERIENCE

### Machine Learning Instructor

DATA SCIENCE DOJO | 2020 - Present

- Teaches unsupervised algorithm and time series classes

### Data Science Instructor/Mentor

AWARI | 2019 - Present

- Mentor students from the Data Science course
- Instructor from the whole 16 weeks Data Science program.

### Machine Learning Consultant

ALURA | 2019 - Present

- Create Machine Learning content and courses for Alura

### Data Science and Machine Learning Mentor

UDACITY | 2018 - 2021

- Mentor students from 7 different Nanodegrees
- Review projects and codes

### Assistant Professor

University of Brasilia | 2014 - 2015

- Teaches Robotics, Machine Learning, Control and Fluid Mechanics

### Electrical Engineer

VERT TIC - 2014 - 2014

- Responsible for the electrical part of the first Brazilian modular datacenter


### Technical Specialist


IBM | 2011 - 2014


- Help customers with pre and postsales technical issues regarding data storage

## CONTACT ME AT

 São Paulo, Brazil (UTC-3)

 malvar.sara@gmail.com  
malvar.sara@usp.br

 <https://smalvar.github.io/>

 @saramalvar

 <https://github.com/smalvar>

## SKILLS SUMMARY

●●●● Python

●●●● Machine Learning


●●●● Deep Learning (Keras and Tensorflow)


●●●● SQL


●●●● Tableau / Dashboards

●●●● Matlab, C and FORTRAN

## AWARDS RECEIVED

 Best undergraduation project (ABCM/EMBRAER) - 2014

 Best master dissertation (ABCM/EMBRAER) - 2015

 FAMELAB finalist - Scientific Communication Prize - 2017 and 2020

## EDUCATIONAL HISTORY

### Research Centre for Gas Innovation

Postdoctoral Fellow | 2019 - Present

- NLP development for SDG
- Deep Learning for FWI Inversion
- Machine Learning for Rheological Tests
- Wavelet, Lapped Transforms and RPCA for signal processing.

### University of São Paulo

PhD | 2015 - 2019

- GPA: 5
- "Hydrodynamics of living fluids in microflows"

### University of Pennsylvania

PhD | 2019

- Research Internship under Prof. Paulo Arratia laboratory

### University of Tokyo

PhD | 2018

- Research Internship under Prof. Shu Takagi laboratory

### University of Brasilia

Master degree | 2014 - 2015

- GPA: 5
- "Bubble dynamics in ferrofluids: theory and applications"

### University of Brasilia

Bachelor in Electrical Engineering | 2009 - 2014

- GPA: 4.89

# Portfolio

---

You can find some of my projects on these links. Check [Google Scholar](#) and [Lattes](#) for publications

## MACHINE LEARNING

- [ACTIVITY AWARE PULSE RATE ALGORITHM](#): I've estimated the pulse rate from the PPG signal and a 3-axis accelerometer.
- [PREDICT WINE QUALITY](#): In this project, I've predicted the quality of wine based on their features (there is also a discussion about scalers, EDA and metrics)
- [TIME SERIES USING COVID-19](#): The notebook explains the most important characteristics of Time Series based on COVID-19 data. This project became a [course](#).
- [FRAUD DETECTION](#): Comparison of supervised and unsupervised algorithms to detect fraud in e-commerce data. This project became a [course](#).
- [RENT CALCULATOR](#): This project computes all the Data Science Pipeline: web scraping, cleaning and wrangling, EDA, prediction and conclusions using Brazilian data.
- [CLIENT SEGMENTATION](#): Using clustering algorithm to segment clients from a big e-commerce. This project became a [course](#).
- [NLP ON PAPERS FOR ONU'S SUSTAINABLE DEVELOPMENT GOALS](#): In this project, I've used Natural Language Processing on papers produced by the Research Centre for Gas Innovation to quantify their actions regarding ONU's Sustainable Development Goals. Segments were also created.

## DEEP LEARNING

- [PNEUMONIA DETECTION](#): In this project, I've analyzed data from the NIH Chest X-ray Dataset and trained a CNN to classify a given chest x-ray for the presence or absence of pneumonia.
- [BUBBLE DYNAMICS VIBRATION](#): Published paper on identifying the vibration parameters on bubble dynamics

## DEEP REINFORCEMENT LEARNING

- [NAVIGATION](#): For this project, I've trained an agent to navigate and collect bananas in a square world using Unit environment. This project implements a Deep Q-Networks, which combines two approaches: Q Learning (reinforcement learning method) and Q-table approximation learnt by a deep neural network.
- [CONTINUOUS CONTROL](#): In this project, I've controlled a double-jointed arm that can move to target locations. A reward of +0.1 was provided for each step that the agent's hand is in the goal location. Thus, the goal of your agent is to maintain its position at the target location for as many time steps as possible. The off-policy DDPG algorithm was used.
- [COLLABORATION AND COMPETITION](#): In this environment, two agents control rackets to bounce a ball over a net. If an agent hits the ball over the net, it receives a reward of +0.1. If an agent lets a ball hit the ground or hits the ball out of bounds, it receives a reward of -0.01. Thus, the goal of each agent is to keep the ball in play. I've used the deep deterministic policy gradient (DDPG) adjusted to a multi-agent case

## SELECTED MENTEES PROJECTS

- [LUMI APP](#): Movie recommendation system using NLP (using Scikit's CountVectorizer and Cosine Similarity)
- [HEART FAILURE PREDICTION](#): Model for predicting mortality caused by Heart Failure
- [PCA / CLUSTERING / TIME SERIES](#): Complete project using PCA, clustering analysis and time series forecast with robust e-commerce database.