Dana Naderi

Data Scientist

BREST, FRANCE
R Dana-Naderi-2
in dana-naderi
✉ naderidana@yahoo.com

Summary of qualifications

Data Science graduate with passion about study and research and ability to perform well in a team. Relevant skills include machine learning, statistical learning, problem solving, programming, and creative thinking.

- Hands-on experience in **Machine Learning** and **Deep Learning** algorithms.
- Extensive knowledge in **implementing statistical techniques** in real data analysis.
- Advanced knowledge in **Python** (Pandas, Numpy, SciPy, Matplotlib, Scikit-learn).
- 2+ years experience in data analysis and data management using **R** (data **cleaning**, **wrangling**, **visualization**, **modeling**)
- Expert in using **R**, **RStudio**, and **Rmarkdown** to produce reproducible results.
- Solid background in **statistical methodologies**, **mathematics**, and **problem solving**.
- Extensive knowledge in **High-Dimensional Data analysis**, **Bayesian Inference**, and **Probabilistic Network Analysis**
- Proficient in **Microsoft Office** (Word, Excel, PowerPoint) and LaTeX.
- Familiar with Linux and Shell Scripting
- Motivated, hardworking, creative and detail-oriented with a master degree in Data Science and interested in interdisciplinary research.

Education

2020–2023   **Master degree in Data Science**, *Allameh Tabataba'i University*, Tehran, Iran
GPA: 4/out of 4, 17.88/ out of 20 (University Overall Average= 16.84, Faculty Overall Average= 16.32 , Field Overall Average= 16.82)
Thesis: `Application of Learning Dynamic Bayesian Networks in Brain-Machine Interface`

2013–2017   **Bachelor degree in Mathematics**, *Farhangian University*, Tehran, Iran
GPA: 3.78/out of 4, 17.56/out of 20 (University Overall Average=16.53, Field Overall Average= –)

Research interests

- **Deep learning and Machine learning**
- **Statistical learning**
- **Probabilistic Graphical Modelling**
- **Optimization**

Research experience

2023   **Bayesian Penalized Regression in High-Dimensional Data Analysis: A Case Study on Raman Spectroscopy for Disease Diagnosis**, *(Submitted; Joint work with Dr. Farzad Eskandari, Mohammad Mahdavi)*

In this study, our data is from raman spectroscopy with high-dimensional features, that is extracted from COVID-19 patients. The traditional methods cannot provide accurate analytical results. A good way to deal with this type of data is to use the Bayesian approach, which is more accurate and reliable than conventional methods. We have examined Bayesian logistic regression with several shrinkage functions, model selection criteria, and compare the results with results from ML classifiers. Our analysis show that the results from Bayesian penalized regression models has a remarkably higher accuracy than other methods. In this work, I am contributing in all parts including research design, data analysis and running statistical ML algorithms, writing drafts.

2023 **Decoding Neural Signals with Computational Models: A Systematic Review of Invasive BMI**, *(Submitted; Joint work with Dr. Ayoub Bokani from Queensland University, Dr. Diako Ebrahimi from Texas Biomedical Research Institute, Dr. Jahan Hassan, Rezwan Firouzi, Foad Abdi, Hamed Ahmadiani)*

We took a deep look into Neuralink, brain–machine interfaces (BMIs) data. We provided details regarding possible applications of Neuralink, design challenges, and research opportunities in this area as a new approach of BMIs. My contribution in this work is on computational and statistical part, involve signal acquisition and detection, decoding, feature extraction methods, and ML algorithms where used in BMIs.

2022 **Nonparametric Regression Estimation under Kernel Polynomial Model for Unstructured Data**, *(Published at the Journal of Statistical Research of Iran, Joint work with Dr. Farzad Eskandari, Dr. Sima Naghizadeh Ardabili, Mohammad Mahdavi, and Ali Fakhraei)*

In this study, the non-parametric estimation of kernel polynomial regression (NEKPR) model is introduced. In our proposed model, as a flexible multiple analysis tool in the non-parametric inference, the coefficient of variation is estimated with a polynomial function and each component may have a dependency link to the response variable using the kernel function. In this approach, we described optimal property of estimators. To evaluate the accuracy of our method, we used simulation and real data analysis. Simulation and real-data analysis of COVID-19 data show that the proposed method perform better than the traditional methods. In this work, I contributed in all parts including research design, data analysis and running statistical codes, writing first draft and editing it.

2021 **Predicting the epidemic of febrile infectious diseases of the respiratory system based on pharmacy data**, *(In progress; Joint work with Dr. Abbas Aghaei, Funded by the INEF.)*

In this ongoing project, I am financially supported by the National Elite Foundation of Iran (INEF). We work on forecasting changes in the incidence of systemic febrile infectious diseases (COVID-19) based on data gathered from 19 pharmacies in Sanandaj. This is a joint project with Dr. Abbas Aghaei from department of epidemiology at Kurdistan University of Medical Sciences, January 2021 - September 2022. In this work, I am contributing in all parts including problem definition, research design, data analysis and running statistical codes, writing first drafts and editing them.

COURSE PROJECTS

Jan 2022– **Diagnosis of Covid-19 by Bayesian approach**, *Under supervision of Dr. Farzad*
Apr 2022 *Eskandari, department of statistics at Allameh Tabataba'i University*

This project was a supplemental work in the "Bayesian Statistics in Big Data" course, from which we are currently extracting an article titled "Bayesian Penalized Regression Vs. Machine Learning Models for Studying High-Dimensional Raman Spectroscopy Data".

Oct 2021– **Graphical models and Bayesian networks**, *Under supervision of Dr.Vahid Rezaei-*
Mar 2021 *Tabar, department of statistics at Allameh Tabataba'i University*

This project was done as a supplemental work in the "Graph Theory and Social Networks" course, which mainly focused on graph models, probabilistic networks, and the method of structural and parametric learning in Bayesian networks. This project made me choose the subject of my thesis in this field.

Jan 2021– **A study on of Covid-19 from data-science perspective**, *Under supervision of Dr.*
May 2021 *Farzad Eskandari, department of statistics at Allameh Tabataba'i University*

This project was done as a supplementary work in the "Data Science Basics" course and the "Data Modeling and Visualization" course. we implemented the steps of data modeling and visualization on the COVID-19 data collected by John Hopkins University, from which an article entitled "Nonparametric Regression Estimation under Kernel Polynomial Model for Unstructured Data" was extracted.

Oct 2020– **An essay on Machine Learning algorithms using R**, *Under supervision of Dr. Vahid*
Feb 2021 *Rezaei-Tabar, department of statistics at Allameh Tabataba'i University*

This project was done as a supplemental work in the "Machine Learning" course. The algorithms were tested on several data sets in order to provide practical training.

Conferences

2021 **Nonparametric Regression Estimation under Kernel Polynomial Model for Unstructured Data**, 16th *Iranian Statistics Conference*, 24-26 Aug 2022 in Babolsar, Iran, hosted by University of Mazandaran, Joint work with Dr. Farzad Eskandari, Dr.Sima Naghizadeh Ardebil, Mohammad Mahdavi, Ali Fakhraei.

Teaching experience

2017-2022 **High-school Mathematics Teacher**, *The education department, Marivan, Kurdistan, Iran.*

Graduate courses

Average grade: 17.88 Out of 20, *(University Average=16.12, Field Average= 16.88)*

Thesis: Application of learning dynamic Bayesian networks in brain-machine interface

- **Bayesian Statistics for High Dimensional Data** (3 credits, 18/20)
- **Graph Theory and Social Network Analysis** (3 credits, 19/20)
- **Data Structures and Databases** (3 credits, 17.8/20)
- **Data Modelling and Visualization** (3 credits, 18.5/20)
- **Foundation of Data Science** (3 credits, 17.5/20)
- **Machine Learning** (3 credits, 19/20)
- **Data-Intensive Computing** (3 credits, 19/20)
- **Performance Evaluation of Computer and Systems** (3 credits, 16/20)
- **Optimization and Data Science** (3 credits, 16/20)
- **Seminar** (2 credits, 18/20)

Online courses and certificates

- **Summer School of Data Science and Applications in FinTech & Industry**: Presented by Faculty of Mathematics, Statistics and Computer Science of Allameh Tabataba'i University (ATU) and Universiti Teknologi MARA (UiTM).
- **The R programming for Data Science A-Z Complete Diploma 2022**: Udemy
- **Master in Microsoft Power BI Desktop and Service**: Udemy
- **Data Analysis, Data Science and Visualization: Python and Pandas**: Udemy
- **[2021]Data Science: A-Z Bootcamp + Real Cases**: Udemy
- **Deep Learning Foundation**: Udemy

Analytical skills

- R, Python
- Linux, Shell Scripting
- SPSS, MINITAB, SAS
- Microsoft Office and LaTeX

Language skills

- English (IELTS 6.5)
- Kurdish (Native), Farsi (Native)

References

- Dr. Bahman Rostami-Tabar, Cardiff University, UK, `rostami-tabarb@cardiff.ac.uk`
- Dr. Ayub Bokani, Central Queensland University, Australia, `a.bokani@cqu.edu.au`
- Dr. Farzad Eskandari, Allameh Tabataba'i University, Iran, `askandari@atu.ac.ir`
- Dr. Vahid Rezaei-Tabar, Allameh Tabataba'i University, Iran, `vhrezaei@atu.ac.ir`