



# Zahra Zabihinpour

Data Scientist, PhD in Physical Chemistry & Statistics

Highly motivated data scientist with a background in healthcare, biotech, research, and statistics. Eager to learn in a fast-paced environment, while thriving in an agile team. I really enjoy coding and having an impact on people. I also value teamwork and building strong relationships.

✉ zabihin@gmail.com

📍 Paris, France

🐙 zabihin.github.io

☎ +33666201578

🌐 linkedin.com/in/zahra-zabihinpour

## TECHNICAL SKILLS

### Programming languages

Python, C, MATLAB, SQL

### Data Analysis & Management

Data visualization, Tableau, GCP, AWS (S3), MySQL, ETL - Web Scraping, Spark.

### Machine Learning / Deep Learning

Keras, TensorFlow, spaCy, Transformers, Dash, Flask, FastAPI, LLM.

### Web Scraping

Data Extraction, HTML Parsing, Web Data Collection, BeautifulSoup, Selenium.

### Data Engineering

Docker, MLflow, Streamlit, Heroku,

### Project Management

Agile methodologies, Scrum, Jira, Git, GitHub, Bitbucket.

## PERSONAL PROJECTS

### Symptom checker, diagnosis app

- ◇ Tech stack: Tensorflow, BERT, Hugging face (Training model), Streamlit, GPT-3.
- ◇ <https://huggingface.co/spaces/Zabihin/Symptom-Diagnoses>

### Lung cancer diagnosis, CT scans

- ◇ Tech stack: Image classification, CNN, transfer learning, Inception model.
- ◇ <https://lung-cancer-ct.streamlit.app>

### NLP with disaster tweets (Kaggle)

- ◇ Tech stack: Tensorflow, Spacy, RNN, LSTM, GRU.
- ◇ [https://github.com/Zabihin/NLP\\_disaster\\_tweet](https://github.com/Zabihin/NLP_disaster_tweet)

## LANGUAGES

English / French  
Professional Working Proficiency

Persian  
Native or Bilingual Proficiency

## INTERESTS

AI

Healthcare

Impact

## WORK EXPERIENCE

### Data Scientist

Luna

01/2023 - 12/2023

Paris, France

Luna is an emerging impact & AI startup dedicated to shape the future of women health, starting with optimizing endometriosis care path journey.

#### Achievements/Tasks

- ◇ Led a core end-to-end NLP project, from data collection to deployment - on drug discovery.
- ◇ Implemented state-of-the-art machine learning models and OCR tools to read drug prescriptions and analyse hormonal treatments for endometriosis care.
- ◇ Close collaboration within a team of 3 data scientists in an agile environment.
- ◇ Technologies used : PyTesseract, EasyOCR, PaddleOCR, Regex, FastAPI, Docker, Streamlit, OpenCV.

### Researcher

Institute for Research in Fundamental Sciences

12/2018 - 01/2021

Teheran, Iran

#1 Research center in Iran, known for its fundamental sciences publications.

#### Achievements/Tasks

- ◇ Led a project on stochastic simulation of Gene Expression with delay-differential equations.
- ◇ Stochastic simulation of chemical kinetics using Gillespie algorithm • Mathematical modelling of biological systems.

### Lecturer

Kharazmi University

09/2011 - 01/2021

Teheran, Iran

Historical prestigious university in Iran specialised in engineering and fundamental sciences.

#### Achievements/Tasks

- ◇ Courses taught: Statistical thermodynamics • MATLAB programming • Physical chemistry • Molecular spectroscopy • Quantum chemistry.

## EDUCATION

### Fullstack Data Science Bootcamp

Jedha Bootcamp - #1 in Europe

05/2022 - 08/2022

Paris, France

#### Courses

- ◇ **Image classification, Convolutional Neural Network** • Designed and trained a deep learning model to analyse chest CT scan images for the early detection of lung cancer.
- ◇ **Transfer Learning** • Implemented neural network and deployed a web app on Streamlit to classify chest CT scan images, achieving 96.3% specificity, and 80% accuracy.

### PhD Physical Chemistry, Statistics

Kharazmi University

09/2007 - 06/2012

Teheran, Iran

#### Courses

- ◇ Numerical simulations of super-diffusion using power law probability distribution.
- ◇ Proof of non-ergodicity of Lévy flight through the extraction of analytical formulas.