HAS SOTHEA

PHD IN APPLIED MATHEMATICS

CONTACT

📞 06 18 23 22 02

✉️ sothea.has@lpsm.paris

📍 Paris, France

🌐 https://hassothea.github.io/

EDUCATION

2018 - 2022
SORBONNE UNIVERSITÉ
- PhD in Applied Mathematics

2018
UNIVERSITÉ PARIS DIDEROT
- Mater's degree of Data Science

ENSIIE
- Engineer of Applied Mathematics

2015
ROYAL UNIVERSITY OF PHNOM PENH, CAMBODIA
- Bachelor's degree in mathematics

SKILLS

- Applied machine learning (ML)
- Theoretical ML and Statistics
- Data analysis and modeling
- ML for climate science
- Python, PyTorch, R, C++, ...

LANGUAGES

- Khmer (mother tongue)
- English (fluent)
- French (conversational)

PROFILE

I am currently a postdoctoral researcher in applied Machine Learning (ML) in climate science at Laboratoire de Météorologie Dynamique (LMD) of École Normale Supérieure (ENS), France. My area of expertise is theoretical/applied ML, Statistics, and Data Science. I have been very interested in data and tasks involving using ML tools.

WORK EXPERIENCE

CNRS/LPSM - ENS/LMD - École polytechnique   **2022 - PRESENT**
Postdoctoral researcher

- Applied ML in reconstructing a time series of Gravity Wave Momentum Fluxes (GWMF)
- Extracting import features for reconstruction and insights
- Building a Python library for various aggregation methods

Teaching at Université Paris Cité   **2019 - 2024**
Teaching assistant for 1st and 2nd year Master students

- Data Analysis with R and Rstudio
- Data Mining with R and Rstudio
- Exploratory Data Analysis with R and Rstudio
- Algorithm and Programming with Python
- Big Data Technologies with Python and Spark
- Statistical Inference and Data Modeling

LPSM - Université Paris Cité   **2018 - 2022**
PhD research and projects

- Theoretical studies of consensual aggregation method in ML
- Combining supervised and unsupervised ML for data modeling
- Energy modeling (wind turbine & air compression machine)
- Applied ML in physical quantity modeling (pitch angle diffusion)

PUBLICATION

- **2024**, Estimating balloon-observed gravity wave momentum fluxes using ML & input from ERA5. *Published in JGR - Atmosphere.*

- **2023**, Gradient COBRA: A kernel-based consensual aggregation for regression. *Published in Journal of Data Science, Statistics and Visualization.*

- **2022**, A consensual aggregation of randomly projected high-dimensional features of predictions. *Available in HAL.*

- **2022**, Machine learning methods applied to the global modeling of event-driven pitch angle diffusion coefficients during high-speed streams. *Published in Frontiers Physics.*

- **2021**, KFC: A clusterwise supervised learning procedure based on aggregation of distances. *Published in Journal of Statistical Computation and Simulation.*