

EXERCICE TO DELIVER. Ice Classification in the Groenland

The aim of this study is to evaluate the ability of several learning machines for building a predictive classifier to evaluate, thanks to infrasound records, the (low or high) quantity of ice at a given spot in the Groenland.

This exercice should be conducted by a group of two students (no more). Each file should clearly mentionned the author names.

The requested work includes a report of 8 pages maximum presenting a synthesis of the studied models and the obtained results as well as the Python code file developed for this study.

Both files (report and code) will be submit on the ENSIE Exam web site, directory **MAL2024TP** . The work should be download **before November, Tuesday 17th**.

Context

In this application, there are 4 potential target ouputs corresponding to infrasound signals (Y1, Y3, Y4, Y5) (file Targets.csv). The covariables are defined by 9 variables providing by the European Centre for Medium-Range Weather Forecasts (ECMWF) .

Description of the Output Variable

The displacement of large volumes of air leads to low-frequency acoustic waves in the atmosphere. This so-called infrasound is inaudible to humans as it has frequencies lower than 20 Hz. In the raw data, the 4 output variables are quantitative infra-sound records. For this classification study, **the quantitative valued target signals will be transformed into binary information using a appropriate given threshold**.

Description of the Input Variables

- **climate information:** the European Weather Center provides information on 2 meter below sea temperature (t2m); Sea-surface temperature (SST); and wind speed (u10, v10)
- Sea Ice Concentration information (SIC)
- Groenland liquid water discharge simulated by Region Climate Models for 5 regions (r1_MAR, r2_MAR, r3_MAR, r4_MAR, r5_MAR)

More detailed information can be found in the Geophysical Research Letters, “Long-Term Infrasonic Monitoring of Land and Marine- Terminating Glaciers in Greenland”, Research letter, DOI 10.1029/2021GL097113, AGU advancing earth and space conference.

The data

The following instructions let to read the Input (data_features.csv) and output (data_Targets) data in Python.

```
import pandas as pd

#Load the data
tab = pd.read_csv('data_Features.csv')
tabY = pd.read_csv('data_Targets.csv')
```

Description of the work

- In this work, you have to study only one given target variable (Y1)
- Transform the target variable into a binary variable using an appropriate and motivated threshold.
- Train, test and compare different binary machine learning classifiers to predict the binary target variable.
- Conclusion. Draw conclusions about the performances and the models used.