ARTIFICIAL CANCER BREATH DETECTING E-NOSE : K-NN VS LSTM

Justin D.M. Martin, Anne-Claude Romain*

Sensing of Atmospheres and Monitoring (SAM), University of Liège (ULiège), Arlon Campus Environnement, Arlon, Belgium

Sensing of Atmospheres and Monitoring

OBJECTIVES

- Evaluate the obtainable classification performance from a breath-spiking experiment involving an electronic nose
- Discover how a trusted non-parametric method such as k-Nearest Neighbours (k-NN) performs in a classification task against Long Short Term Memory (LSTM) neural network.

METHODS

The electronic nose : SAMBre

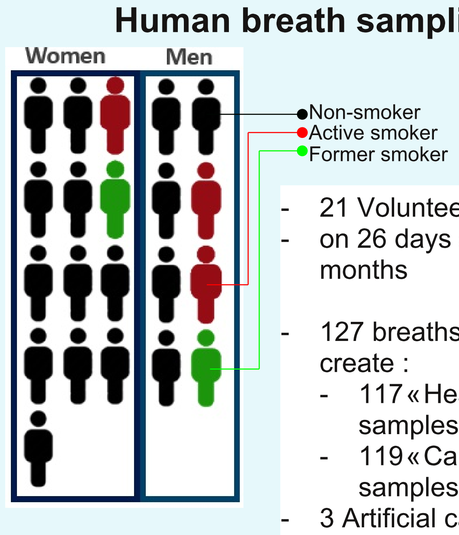
Sensors

- T2603 (Figaro®)
- 3530T, 1430T, 2530T, 8530T (Umwelt Sensor Technik®)
- MP901 (Winsen®)
- BME680 (Bosh®)

Specifications

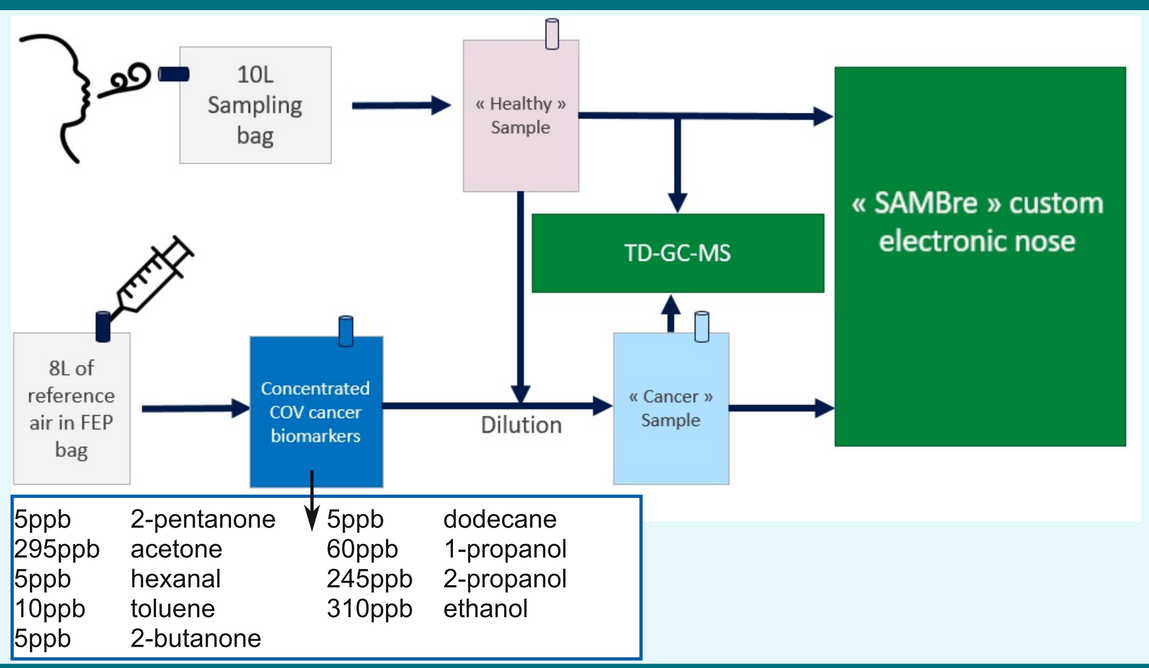
- Temperature maintained at 45°C within chamber
- Flow constant at 200mL/min
- Moisture and temperature monitored

Human breath sampling



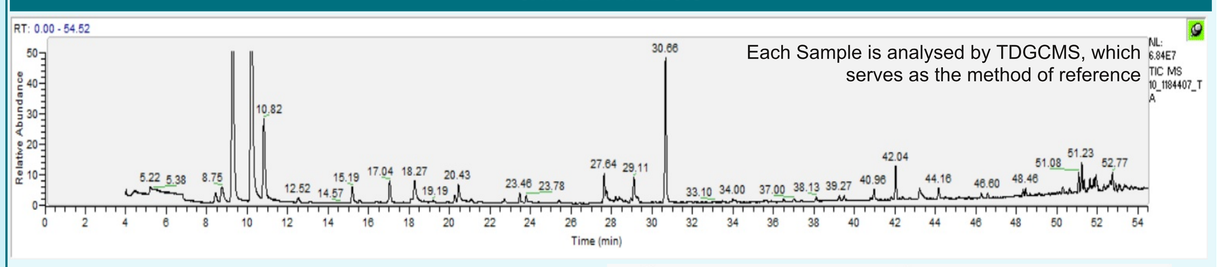
- 21 Volunteers
- on 26 days across 3 months
- 127 breaths used to create :
  - 117 « Healthy » samples
  - 119 « Cancer » samples by spiking
- 3 Artificial cancer Concentrations : 1.5x, 4x, 8x (comp. with lit.)

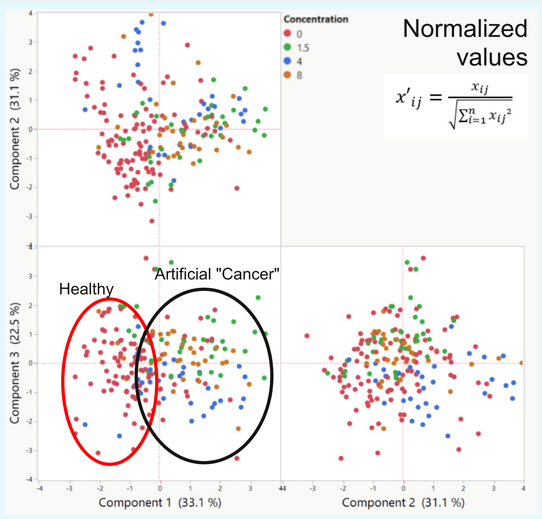
TESTING PROCEDURE



5ppb	2-pentanone	5ppb	dodecane
295ppb	acetone	60ppb	1-propanol
5ppb	hexanal	245ppb	2-propanol
10ppb	toluene	310ppb	ethanol
5ppb	2-butanone		

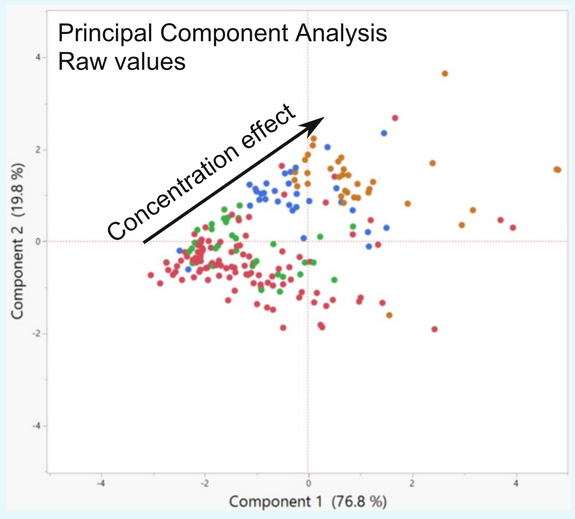
SAMPLE ANALYSIS



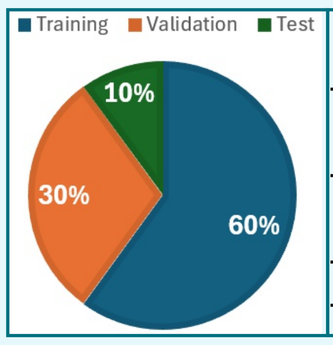


Normalized values

$$x'_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} {x_{ij}}^2}}$$



MACHINE LEARNING



For both models

- Same share of dataset for training, validation and external testing
- External testing created with data from a separate day
- Same initial experiment
- 5-fold cross-validation

LSTM

- Time series of 315 seconds per sample
- 100 memory blocks
- Adam optimizer
- Rescaling to 0-1 range
- Done using Python

k-NN

- 1 data per sample (maximum baseline-corrected conductance increase)
- Best model out of the first 30 values of k is selected for validation and testing
- Done using JMP Software

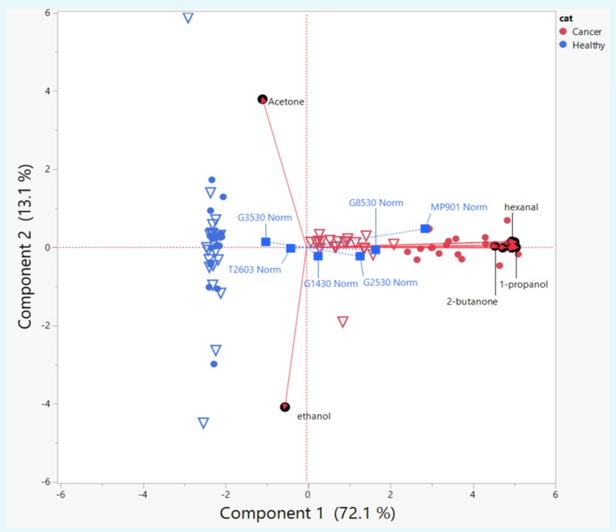
RESULTS

LSTM

- 88±2% accuracy (training)
- 80±5% accuracy (validation)
- 81±4% accuracy (test)
- Using the "time series" dataset on kNN did not produce any performance improvement.
- Underperforming
- No additionnal information within time series ?

k-NN

- 90±3% accuracy (training)
- 91±2% accuracy (validation)
- 91±2% accuracy (test)
- Best k between 1-16
- Best contributing sensors (to discriminating PC)
  - G2530 (62%)
  - MP901 (23%)
  - G3530 (11%)
  - T2603 (4%)



TD-GC-MS ANALYSIS

- Perfect separation between groups
- Acetone influence from fasting (12h overnight nothing by mouth policy)
- Ethanol influence from disinfection
- Overall imprecision on artificial sample creation : below 5%
- Atmosphere stability in FEP bag is above 2h if heated above body temperature

CONCLUSIONS

- Data from 246 individual measurements on breath and spiked breath was collected and successfully used to train a k-NN model and LSTM network.
- LSTM did not perform significantly better than k-NN, but this conclusion should be confirmed by training both on future clinical trial data.
- Overall, classification error is low but cancer biomarkers concentrations had to be raised, which underlines the need for a preconcentration step.

MARTIN Justin
PhD Student for S.A.M. team, ULiège Arlon

CONTACT
- jdm.martin@uliege.be
- +32(0)63230947
- www.labo-sam.uliege.be

This research was funded by Interreg France-Wallonie-Vlaanderen, under the Pathacov project