AI for the industry 4.0

recruitment task

—

,, A company producing fiber optic sensors decided to implement the concept of Industry 4.0 in one of their production processes. They are going to implement the AI techniques to automate verification of produced sensors.
Currently, after the production of the sensor, its characteristics should be verified each time in three reference substances in order to configure:

- Air
- Water
- Isopropanol

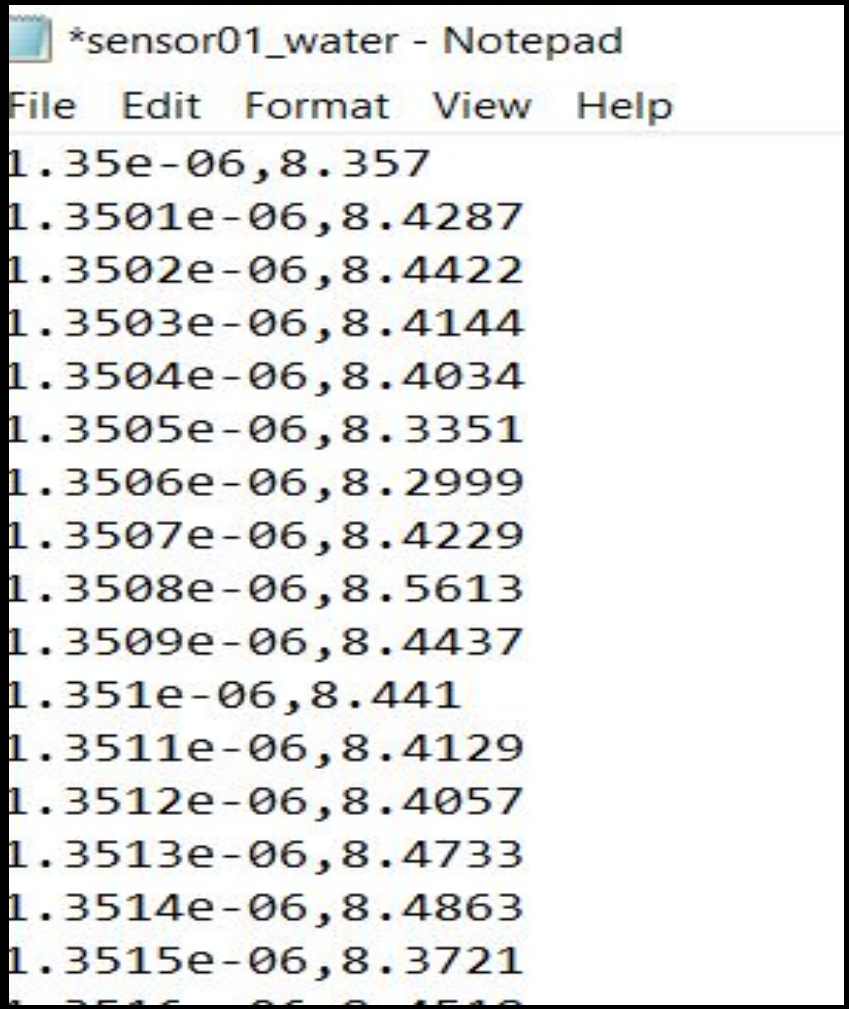
After the procedure in the laboratory they collected three spectroscopic signals. Unfortunately, such procedures are extremely time consuming and require large human resources.
Therefore they are planning to implement AI algorithms in order to predict characteristics (or any other parameters – you can propose something) of a working sensor in water and isopropanol based on measured characteristics in air.
It will enable to reduce validation process only to measurements in the air, but other characteristics will be collected based on the proposed model.

Dataset [...] is a collection of measurements based on 10 sensors . Each sensor was measured three times [...]. Each file contains two-dimensional signal (signal wavelength, signal amplitude) [...].

What I was working on?

Collection of measurements of 10 sensors. Each of them carried out the measurement three times - in air, water and isopropanol.

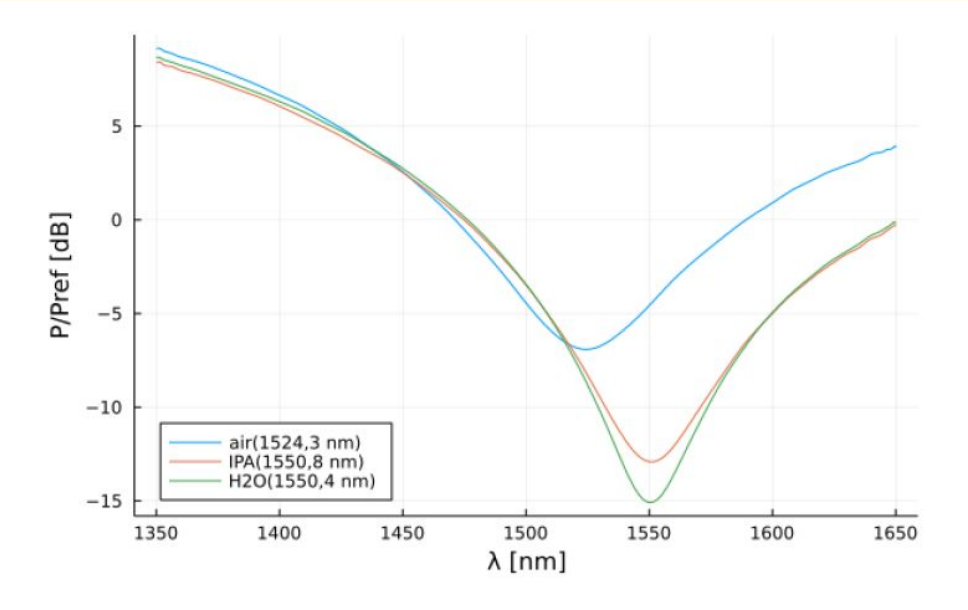
I had 30 files.txt with two-dimensional signal (signal wavelength, signal amplitude)



Assumptions

Main goal:

The use of artificial intelligence is to create algorithms that predict the parameters of optical sensors based on the results of measurements of sensors in the air.

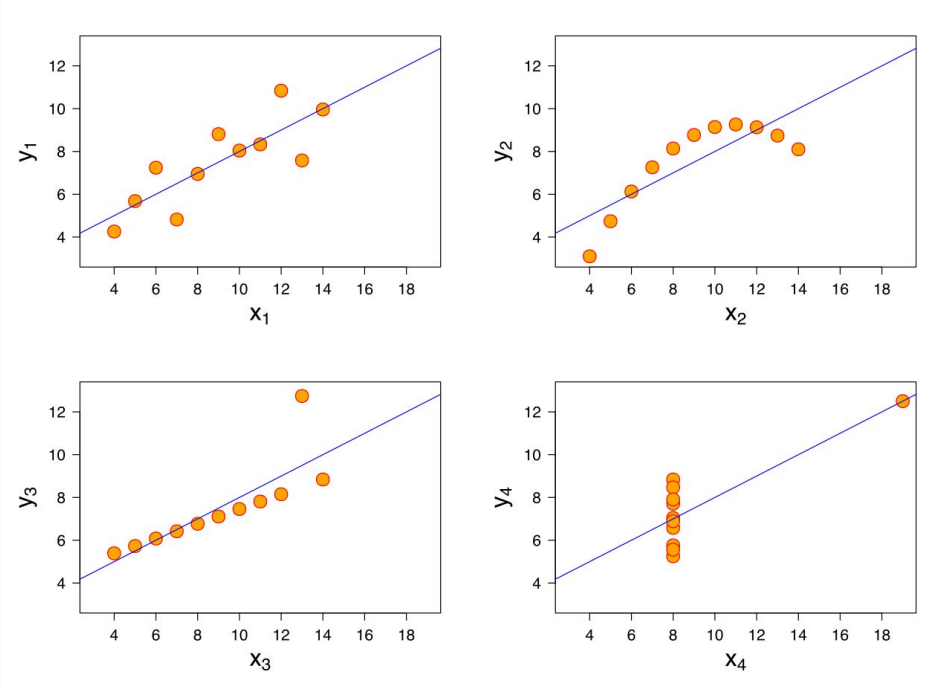


Extra goal:

The use of available data and their dependency for new observations.

My working method

Classic regression problem - predicted continuous values.
I used machine learning models for this and determine their effectiveness with metrics.



What did I want to achieve?

Replacing manual
sensor measurements
in isopropanol and
water.

The result.

Simple regression models - Linear Regression model, Decision Tree and Random Forest Regressor gave very promising results.
Linear Regression gave the best results, although RandomForestRegressor also performed well.

	water			izopropanol		
	Linear Regression	Decision Tree	Random Forest Regressor	Linear Regression	Decision Tree	Random Forest Regressor
R2	0,79	0,63	0,73	0,78	0,60	0,76
MAE	2,56	2,41	2,25	2,72	2,65	2,5
MAPE	1,67	2,46	2,38	2,08	2,58	2,56
MSE	9,02	15,71	11,45	10,73	19,06	14,23
RMSE	3,00	3,96	3,38	3,27	3,28	3,77

Conclusion

Such results are very promising for the creation of more advanced models. The data is highly correlated with each other, which makes it possible to achieve satisfactory results.

Additional comments

Due to the fact that the topic of optical sensors and the measurements made with them is foreign to me, and the time in which I had to complete the task, I could not use extra mechanisms e.g. data enrichment, feature extraction or signal analysis.

Thank you!

Katarzyna Michalak