Information systems for assessing the environmental impact of industrial processes

Marco Cardia
Department of Computer Science
University of Pisa

Outline

1. Introduction
2. Related Work
3. Project overview
4. Conclusion

Analytical parameters

- Chemical Oxygen Demand (COD)
- Biochemical Oxygen Demand (BOD)
- pH, conductivity, hardness…

Analytical parameters

- **Chemical Oxygen Demand (COD)**
- Biochemical Oxygen Demand (BOD)
- pH, conductivity, hardness…

Objectives

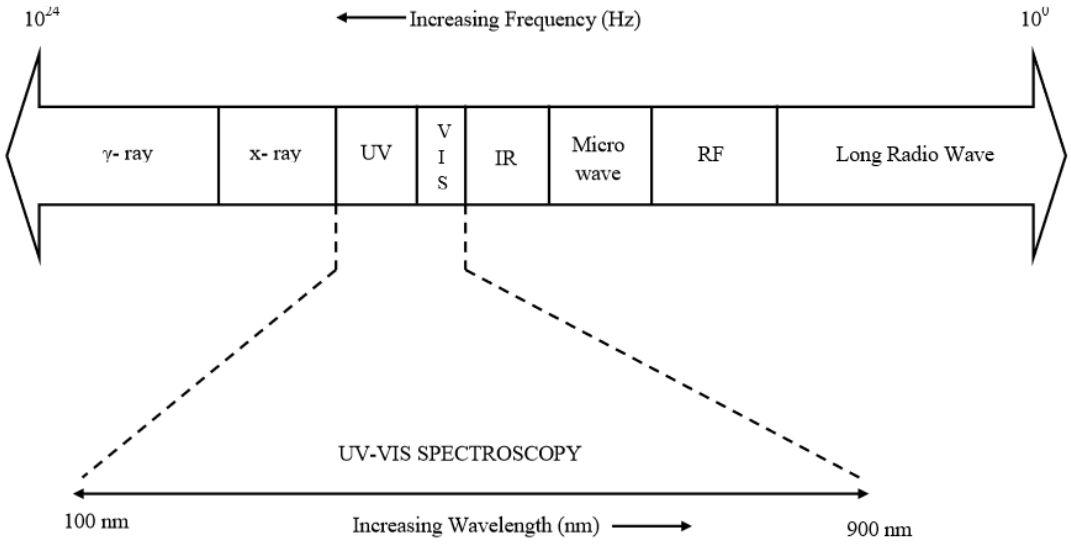
- Build a continuous **monitoring system**
- Reduce laboratory **analysis cost**
- Limit air and water **pollution**

Current method

- Titrimetric analysis
  - Laboratory
- Chemometrics
  - Data-driven

Spectrum

- Ultra-violet and Visible (UV-Vis) spectroscopy

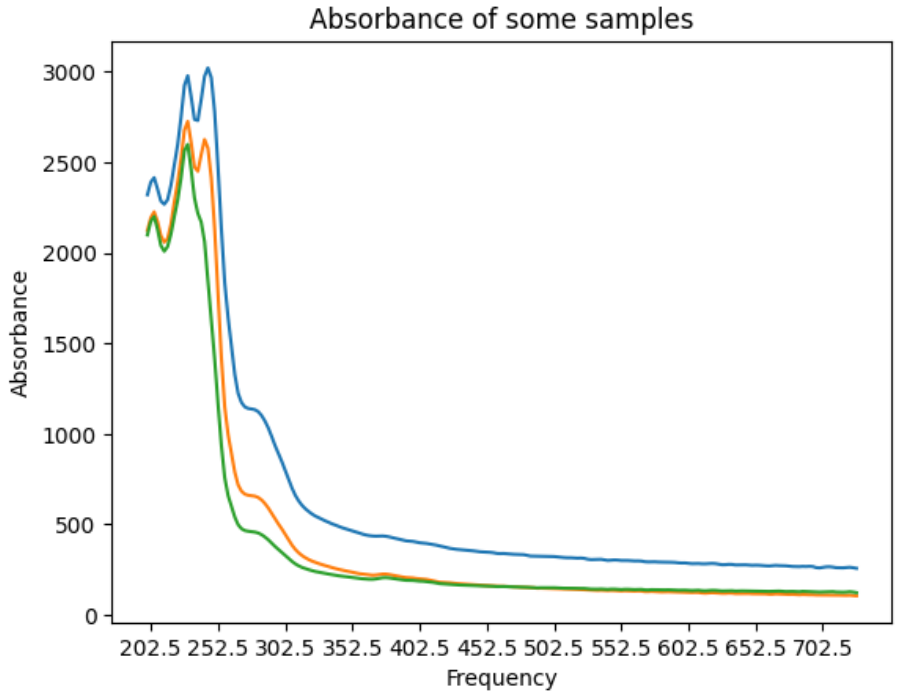


Spectrum

- Ultra-violet and Visible (UV-Vis) spectroscopy
- Beer-Lambert law

Spectrometer

- Absorbance



- Limitations:
  - Optical sensor

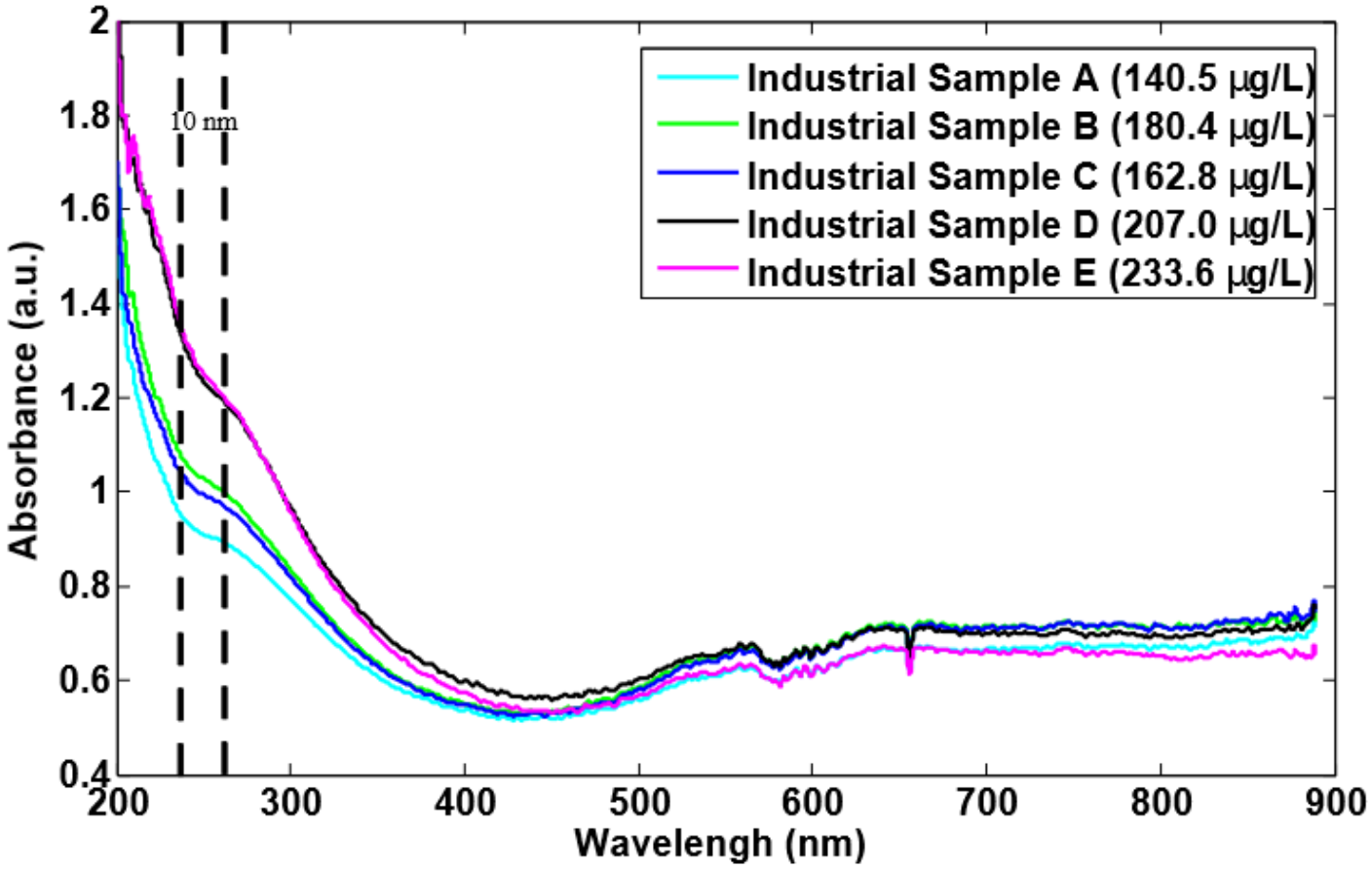
Related work

Collect absonbance

Select a wavelenght

Build linear regressor

Related work



Related work - drawbacks

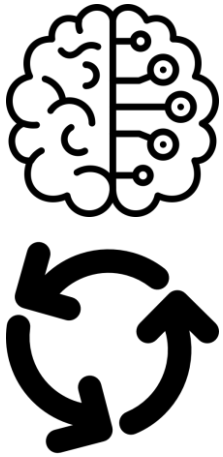
Laboratory prepared samples

One wavelength

Linear regressor stiffness

Our project

- Machine Learning
- Life Cycle Assessment



Project overview

Predict the level of pollutants in wastewaters

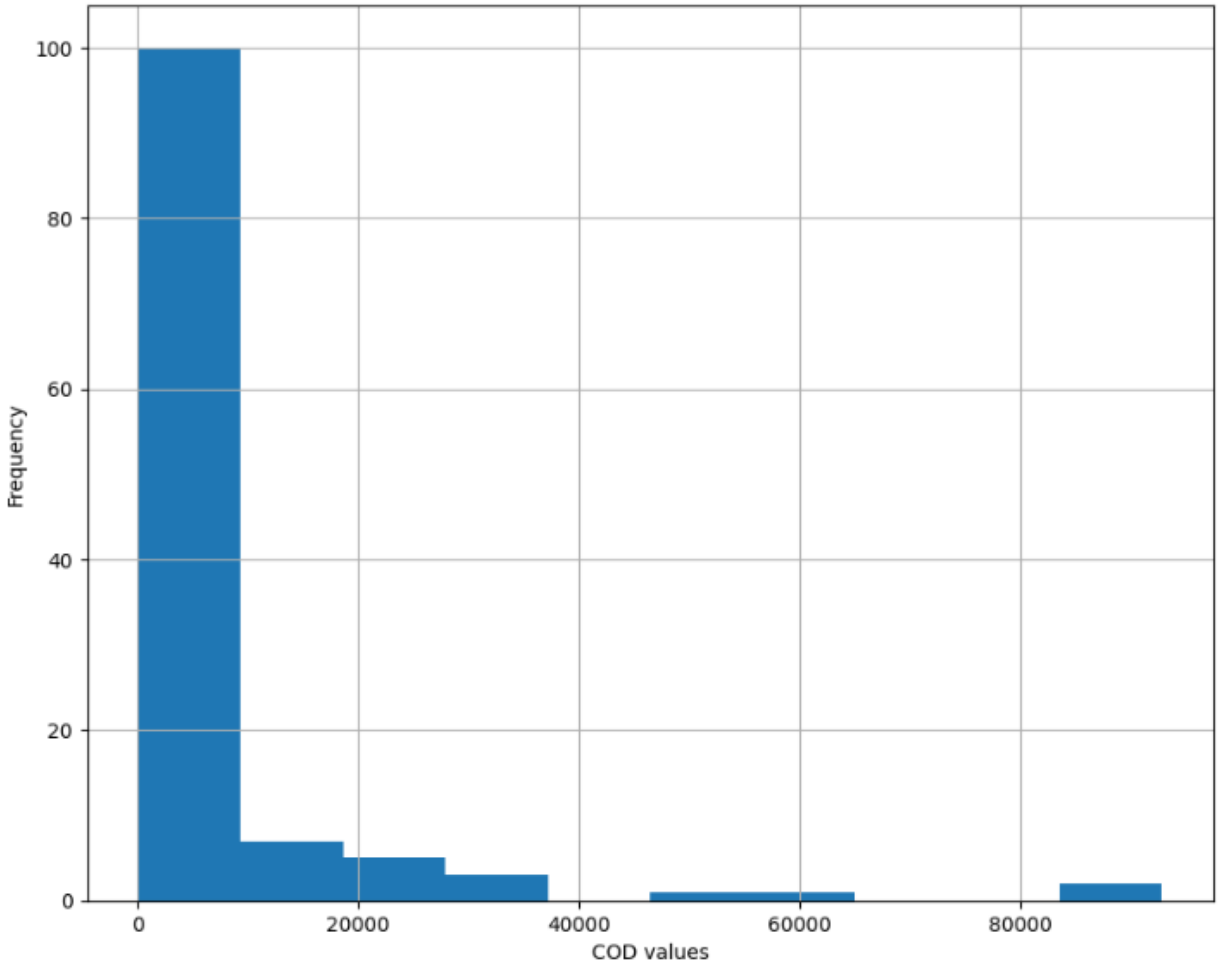
Monitor the pollutants

Simulate the production varying chemical agent

Dataset

- 212 features
- 119 samples

Data distribution



Estimating COD

- Principal Component Analysis (PCA)
- Compare Machine Learning Algorithm

Evaluation measures

- RMSE

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

- R^2

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Results

	RMSE		R2	
Model	Training	Validation	Training	Validation
Null model	14484.14	14165.42	0.00	-0.25
Ridge Regression	275.16	36052.17	0.99	-7.30
Random Forest	4164.05	10638.99	0.86	0.29
SVM	2661.48	4329.57	0.95	0.77
MLP	3522.22	6548.99	0.92	0.55
KNN	7845.81	10672.14	0.68	0.30

Conclusion and next steps

- Objective: Construct a reliable system to constantly monitor pollutants
- Analyse how our work may impact on industries
- Does a chemical agent impact on the pollution?

Thank you!

Marco Cardia
marco.cardia@phd.unipi.it