



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

itee^{PhD}
information technology
electrical engineering



DIE
TI

UNI
NA

Patrizia Quaranta

Early Markers detection in Occupational Health

Tutor: Roberto Pietrantuono

Cycle: XXXVIII

Year: First

My background

- MSc degree in Computer Engineering
- Research group: DESSERT
- PhD start date: 01/11/2022
- Scholarship type: PNRR- DM 352
- Partner company: CML VESUVIO s.r.l Nola

Research field of interest

Occupational Health is an area of work to promote and maintain the highest degree of workers' physical, mental, and social well-being in all occupations.

Data Analysis in Occupational Health:

- Understand and improve the Health Surveillance via **early markers identification**

Summary of study activities

	Course	Seminars	Research	Tutorship
Total	35	6	24	0
Expected	20-40	5-10	10-35	0-1.6

- Ad hoc PhD courses:
 - ✓ Scienza moderna e disciplina giuridica dell'Intelligenza Artificiale
 - ✓ I pilastri della Trasformazione Digitale
- PhD Schools:
 - ✓ SIESTA2023: 3rd International Software Engineering Summer School

Research activity: Overview(1/4)

Problem:

What is an early marker?

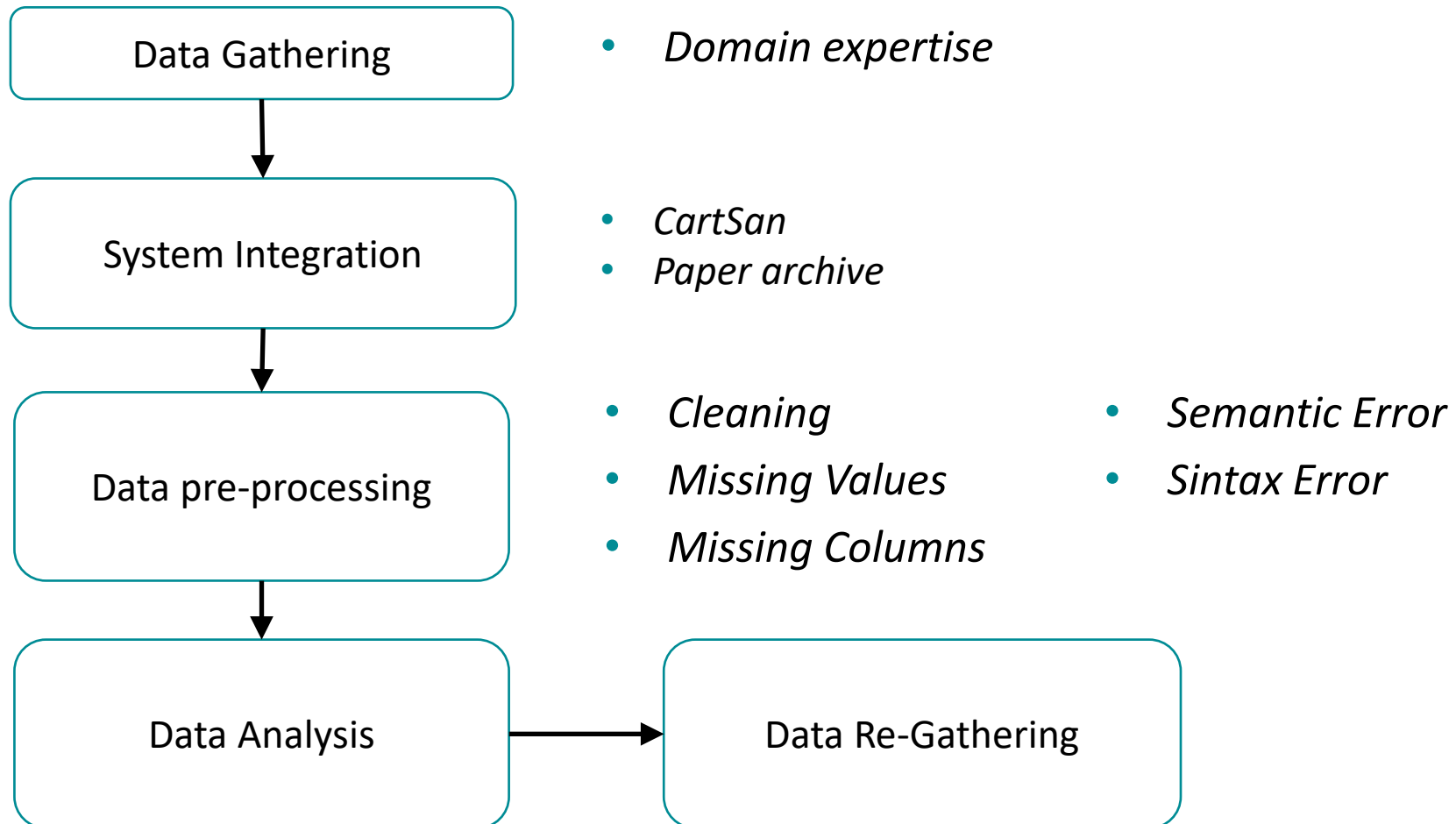
- It's a reversible nonsymptomatic biological alteration
- It can predict symptoms of health impairment that may cause aggravation of conditions that reduce the quality of life or result in a disabling disease

Methodology:

- First method: Use of machine learning techniques:
 - identifying correlations between features that characterize work-related diseases
 - performing Features Ranking and Selection (LASSO Regression, Random Forest) on-going

Research activity: Overview(2/4)

Activity workflow



Research activity: Overview(3/4)

Id	Visit Date	Weight (kg)	Height (cm)	Cig num	Type	Outcome
1	14/07/21	85	177	10	Visita periodica	Idoneo
1	14/07/21	85	177	10	spirometria	Normale&&
1	14/07/21	85	177	10	Alcol test	Negativo
1	14/07/21	85	177	10	Emocromo	Nei limiti di riferimento
1	14/07/21	85	177	10	audiometria	Percentuale di danno biologico pari a: 0 % (D.Lgs. 38/2000)&&Classe 0&&Classe 1a
2	26/07/21	108	180	Sigari 10	Visita periodica	Idoneo
2	26/07/21	108	180	Sigari 10	Emocromo	Nei limiti di riferimento
2	26/07/21	108	180	Sigari 10	Esame obiettivo del rachide	
3	26/07/21	80	170		Visita periodica	Idoneo
3	26/07/21	80	170		spirometria	Normale&&Classe9
4	26/07/21	120	140		Visita periodica	Idoneo

Research activity: Overview (4/4)

Using KNIME Analytics Platform a preprocessing workflow has been implemented.

- *“Type” and “Outcome” Pivot operation (17->41 features)*
- *Synthetic columns creation*

Id	Visit date	...	Audiometry Score	Audiometry Class	Blood count	Visit Type	Outcome
1	14/07/21	...	0	Classe0 Classe 1a	Nei limiti di riferimento	Visita periodica	idoneo
2	26/07/21	...	Not done	Not done	Nei limiti di riferimento	Visita periodica	idoneo
3	26/07/21	...	Not done	Not done	Not done	Visita periodica	idoneo

Next Year

- Second method: Causal inference to identify the causes of work-related diseases
- *Use the Output of Feature Selection as prior knowledge in a Causal Model Implementation*
- *Extract blood examination values from pdf to improve Machine Learning models -> Causal Model*

Thanks for your attention