Lung Cancer Identification by an Electronic Nose based on an Array of MOS Sensors

Rossella Blatt[1], Andrea Bonarini[1], Elisa Calabrò[2]
Matteo Della Torre[3], Matteo Matteucci[1] and Ugo Pastorino[2]

Speaker: Rossella Blatt

blatt@elet.polimi.it

[1] Politecnico di Milano, Department of Electronics and Information, Milan, Italy
[2] Istituto Nazionale Tumori of Milan, Toracic Surgery Department, Milan, Italy
[3] SACMI Imola S.C., Automation & Inspection Systems, Imola (BO), Italy

POLITECNICO MILANO

FONDAZIONE IRCCS ISTITUTO NAZIONALE DEI TUMORI

SACMI

Outline

- Objective: **Lung Cancer** diagnosis classifying the **Olfactory Signal** acquired by an **Electronic Nose**
- Motivation
- Functioning of the Electronic Nose
- **Classification** of volunteers' breath
- Results and comparison with current diagnostic techniques
- Further directions of research

Outline

- Objective: **Lung Cancer** diagnosis classifying the **Olfactory Signal** acquired by an **Electronic Nose**
- Motivation
- Functioning of the Electronic Nose
- **Classification** of volunteers' breath
- Results and comparison with current diagnostic techniques
- Further directions of research

Outline

- Objective: **Lung Cancer** diagnosis classifying the **Olfactory Signal** acquired by an **Electronic Nose**
- Motivation
- Functioning of the Electronic Nose
- **Classification** of volunteers' breath
- Results and comparison with current diagnostic techniques
- Further directions of research

Outline

- Objective: **Lung Cancer** diagnosis classifying the **Olfactory Signal** acquired by an **Electronic Nose**
- Motivation
- Functioning of the Electronic Nose
- **Classification** of volunteers' breath
- Results and comparison with current diagnostic techniques
- Further directions of research

Outline

- Objective: **Lung Cancer** diagnosis classifying the **Olfactory Signal** acquired by an **Electronic Nose**
- Motivation
- Functioning of the Electronic Nose
- **Classification** of volunteers' breath
- Results and comparison with current diagnostic techniques
- Further directions of research

Outline

- Objective: **Lung Cancer** diagnosis classifying the **Olfactory Signal** acquired by an **Electronic Nose**
- Motivation
- Functioning of the Electronic Nose
- **Classification** of volunteers' breath
- Results and comparison with current diagnostic techniques
- Further directions of research

Motivation

- Lung cancer causes more than 160,000 deaths a year in the United States--more than any other cancer
- Once lung cancer is detected the probability of surviving, after 5 years of therapy, is 14%; the survival rate increases to 48% if the cancer is discovered in its earliest stage
- Current diagnostic techniques are invasive, very expensive, have a high risk of complications and a not so good performance
- Efforts at early detection and treatment have been frustrating to date and hence the overall prognosis remains poor

NECESSITY OF A NON INVASIVE, MORE ACCURATE AND CHEAPER DIAGNOSTIC TECHNIQUE, ABLE TO IDENTIFY THE PRESENCE OF LUNG CANCER IN ITS **EARLY STAGE**

Motivation

- Lung cancer causes more than 160,000 deaths a year in the United States--more than any other cancer
- Once lung cancer is detected the probability of surviving, after 5 years of therapy, is 14%; the survival rate increases to 48% if the cancer is discovered in its earliest stage
- Current diagnostic techniques are invasive, very expensive, have a high risk of complications and a not so good performance
- Efforts at early detection and treatment have been frustrating to date and hence the overall prognosis remains poor

NECESSITY OF A NON INVASIVE, MORE ACCURATE AND CHEAPER DIAGNOSTIC TECHNIQUE, ABLE TO IDENTIFY THE PRESENCE OF LUNG CANCER IN ITS **EARLY STAGE**

Motivation

- Lung cancer causes more than 160,000 deaths a year in the United States--more than any other cancer
- Once lung cancer is detected the probability of surviving, after 5 years of therapy, is 14%; the survival rate increases to 48% if the cancer is discovered in its earliest stage
- Current diagnostic techniques are invasive, very expensive, have a high risk of complications and a not so good performance
- Efforts at early detection and treatment have been frustrating to date and hence the overall prognosis remains poor

NECESSITY OF A NON INVASIVE, MORE ACCURATE AND CHEAPER DIAGNOSTIC TECHNIQUE, ABLE TO IDENTIFY THE PRESENCE OF LUNG CANCER IN ITS **EARLY STAGE**

Motivation

- Lung cancer causes more than 160,000 deaths a year in the United States--more than any other cancer
- Once lung cancer is detected the probability of surviving, after 5 years of therapy, is 14%; the survival rate increases to 48% if the cancer is discovered in its earliest stage
- Current diagnostic techniques are invasive, very expensive, have a high risk of complications and a not so good performance
- Efforts at early detection and treatment have been frustrating to date and hence the overall prognosis remains poor

NECESSITY OF A NON INVASIVE, MORE ACCURATE AND CHEAPER DIAGNOSTIC TECHNIQUE, ABLE TO IDENTIFY THE PRESENCE OF LUNG CANCER IN ITS **EARLY STAGE**

Motivation

- **Fundamental Principle of Clinical Chemistry**: "Every pathology changes people chemical composition, modifying the concentration of some chemicals in the human body"
  - In the medical field, clinicians have always considered odor as a fundamental information for the diagnosis of several diseases
- It has been demonstrated (Gordon et al, 1985) that the presence of lung cancer alters the percentage of some **volatile organic compounds (VOCs)** present in **human breath**
  - These VOCs can be considered as **lung cancer markers** and thus used to diagnose it

Motivation

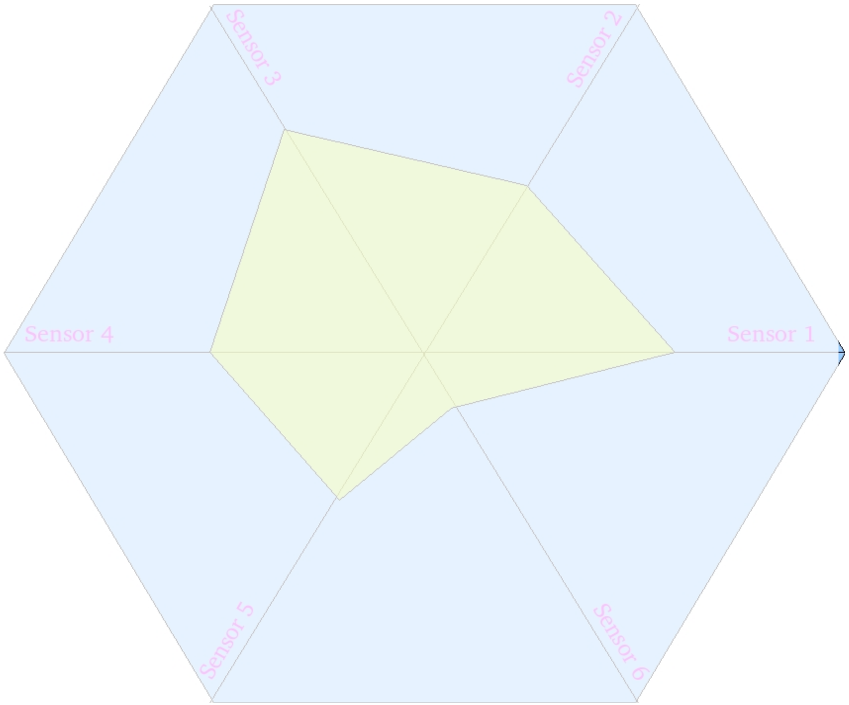
- **Fundamental Principle of Clinical Chemistry**: "Every pathology changes people chemical composition, modifying the concentration of some chemicals in the human body"
  - In the medical field, clinicians have always considered odor as a fundamental information for the diagnosis of several diseases
- It has been demonstrated (Gordon et al, 1985) that the presence of lung cancer alters the percentage of some **volatile organic compounds (VOCs)** present in **human breath**
  - These VOCs can be considered as **lung cancer markers** and thus used to diagnose it

Motivation

- **Fundamental Principle of Clinical Chemistry**: “Every pathology changes people chemical composition, modifying the concentration of some chemicals in the human body”
  - In the medical field, clinicians have always considered odor as a fundamental information for the diagnosis of several diseases
- It has been demonstrated (Gordon et al, 1985) that the presence of lung cancer alters the percentage of some **volatile organic compounds (VOCs)** present in **human breath**
  - These VOCs can be considered as **lung cancer markers** and thus used to diagnose it

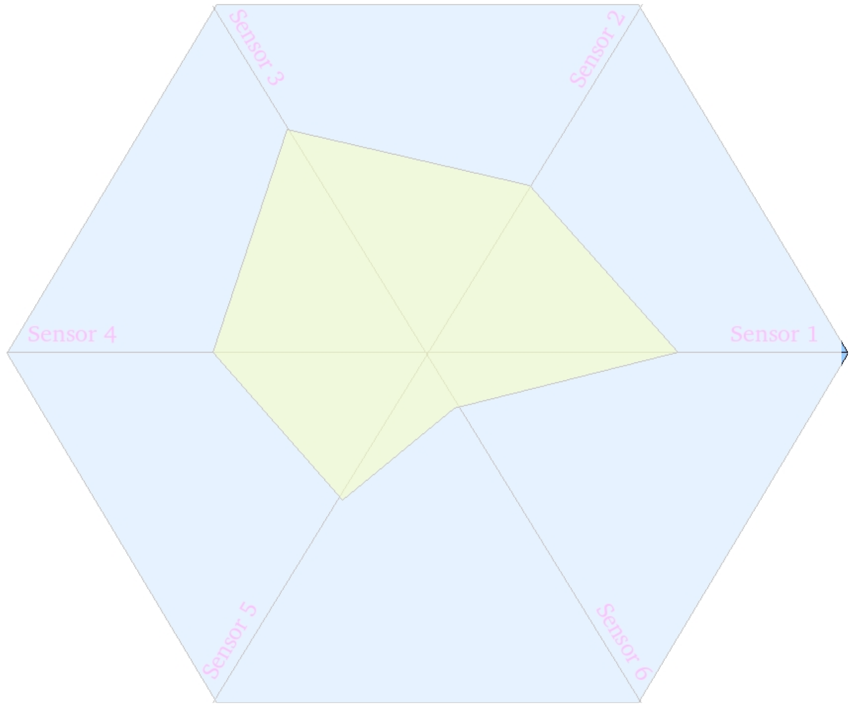
Electronic Nose

- An electronic nose is an instrument able to **acquire, detect and analyse the olfactory signal**
- It is composed of an **array of non specific electronic devices** (sensors) able to **convert a physical or chemical information into an electrical signal**
  - It is non specific because it does not look for particular compounds in the analyzed substance, but for different **patterns**
  - Each sensor reacts in a different way to the analyzed substance providing multidimensional data that can be considered as an **olfactory blueprint** of the substance itself



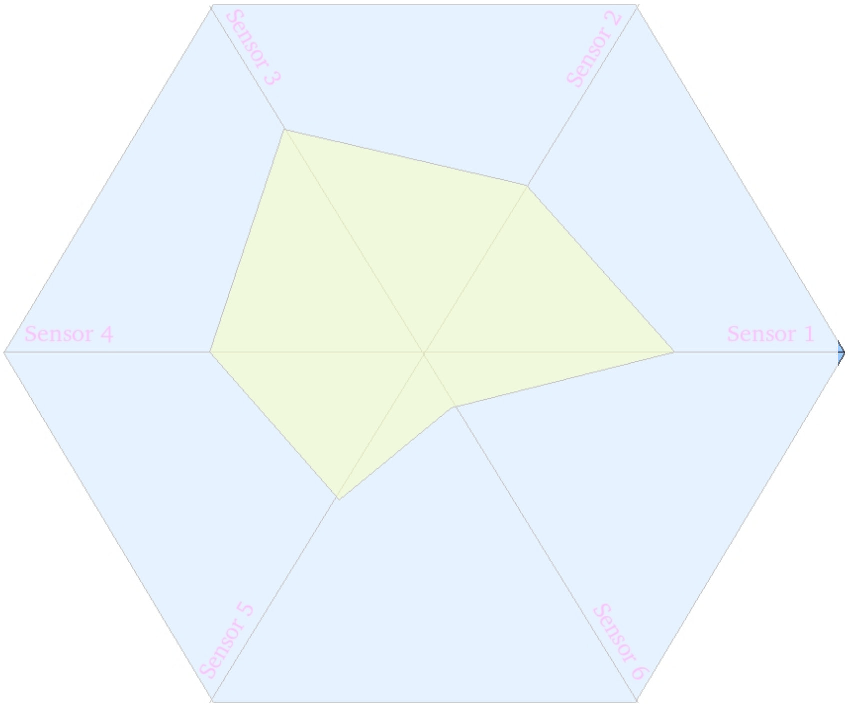
Electronic Nose

- An electronic nose is an instrument able to **acquire, detect and analyse the olfactory signal**
- It is composed of an **array of non specific electronic devices** (sensors) able to **convert a physical or chemical information into an electrical signal**
  - It is non specific because it does not look for particular compounds in the analyzed substance, but for different **patterns**
  - Each sensor reacts in a different way to the analyzed substance providing multidimensional data that can be considered as an **olfactory blueprint** of the substance itself



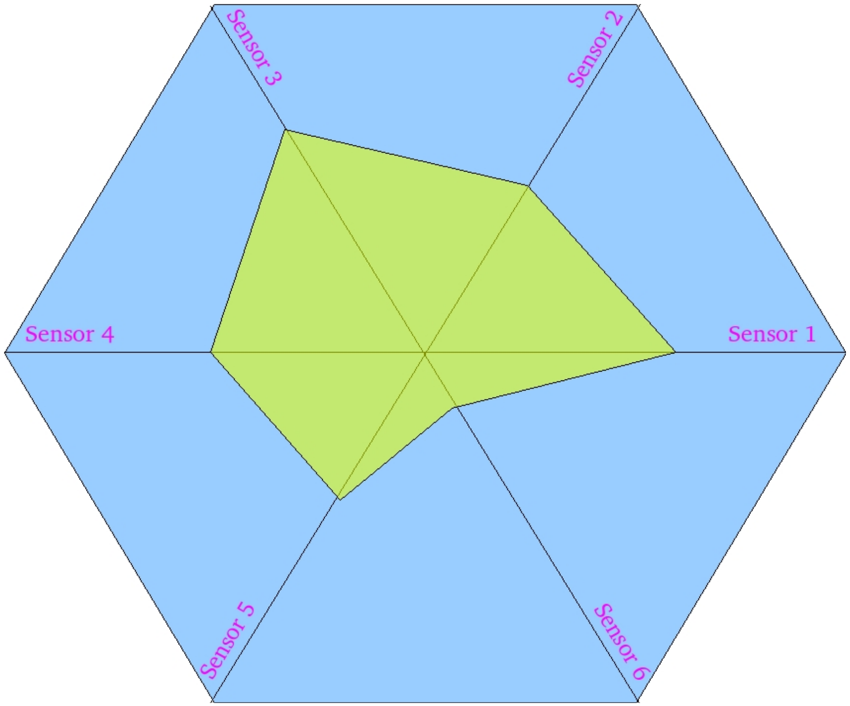
Electronic Nose

- An electronic nose is an instrument able to **acquire, detect and analyse the olfactory signal**
- It is composed of an **array of non specific electronic devices** (sensors) able to **convert a physical or chemical information into an electrical signal**
  - It is non specific because it does not look for particular compounds in the analyzed substance, but for different **patterns**
  - Each sensor reacts in a different way to the analyzed substance providing multidimensional data that can be considered as an **olfactory blueprint** of the substance itself



Electronic Nose

- An electronic nose is an instrument able to **acquire, detect and analyse the olfactory signal**
- It is composed of an **array of non specific electronic devices** (sensors) able to **convert a physical or chemical information into an electrical signal**
  - It is non specific because it does not look for particular compounds in the analyzed substance, but for different **patterns**
  - Each sensor reacts in a different way to the analyzed substance providing multidimensional data that can be considered as an **olfactory blueprint** of the substance itself



Electronic Nose's Output

- According to the used pattern analysis algorithm, the output of an electronic nose can be:
  - the **detection** of a specific substance
  - an estimate of the **concentration** of the odor
  - some particular characteristic of the odor that allows to associate it to a particular **class**

Electronic Nose's Output

- According to the used pattern analysis algorithm, the output of an electronic nose can be:
  - the **detection** of a specific substance
  - an estimate of the **concentration** of the odor
  - some particular characteristic of the odor that allows to associate it to a particular **class**

Electronic Nose's Output

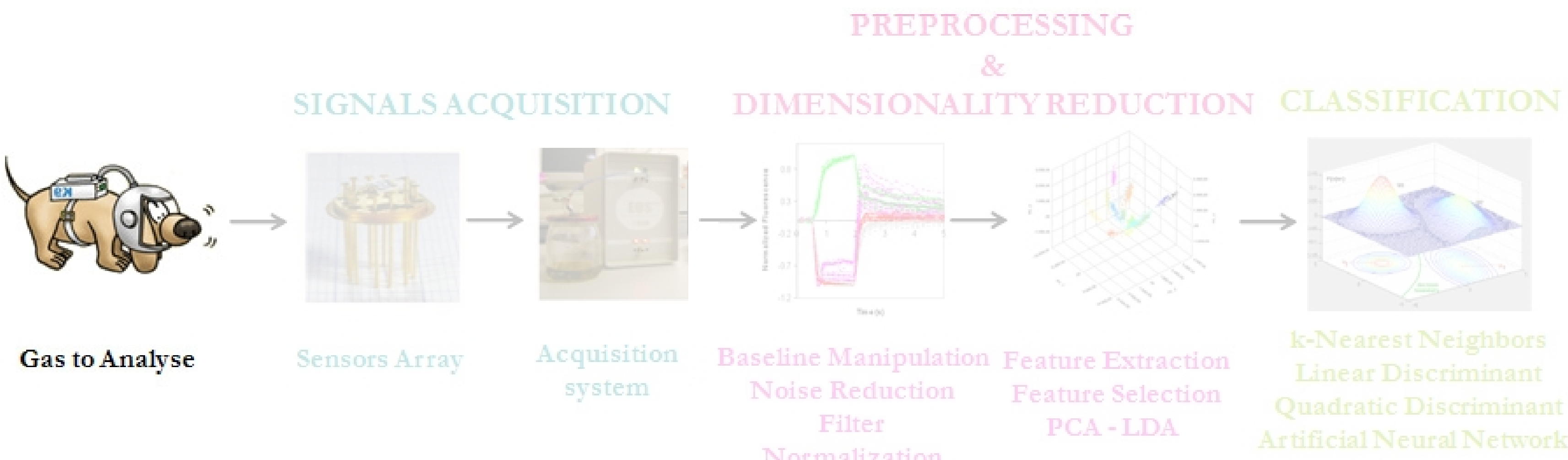
- According to the used pattern analysis algorithm, the output of an electronic nose can be:
  - the **detection** of a specific substance
  - an estimate of the **concentration** of the odor
  - some particular characteristic of the odor that allows to associate it to a particular **class**

Electronic Nose's Output

- According to the used pattern analysis algorithm, the output of an electronic nose can be:
  - the **detection** of a specific substance
  - an estimate of the **concentration** of the odor
  - some particular characteristic of the odor that allows to associate it to a particular **class**

Electronic Nose

1. **Signal Acquisition**
   - Acquisition is done through a sensor array that measures a given physical or chemical quantity and convert it into an electrical signal
2. **Signal Processing**
   - **Preprocessing**: aimed to reduce the impact of noise
   - **Dimensionality Reduction**: reduce the dimensionality of the problem, enhancing classification performance
3. **Classification and Validation**
   - Classification between the two classes “healthy” and “sick”



Electronic Nose

1. Signal Acquisition

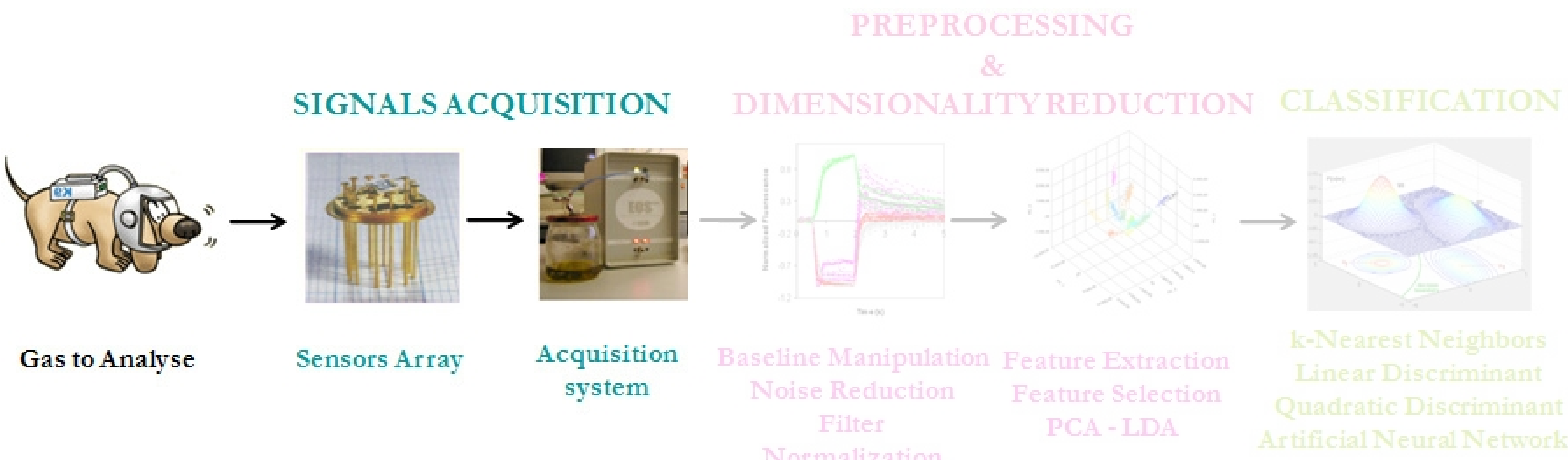
- Acquisition is done through a sensor array that measures a given physical or chemical quantity and convert it into an electrical signal

2. Signal Processing

- **Preprocessing**: aimed to reduce the impact of noise
- **Dimensionality Reduction**: reduce the dimensionality of the problem, enhancing classification performance

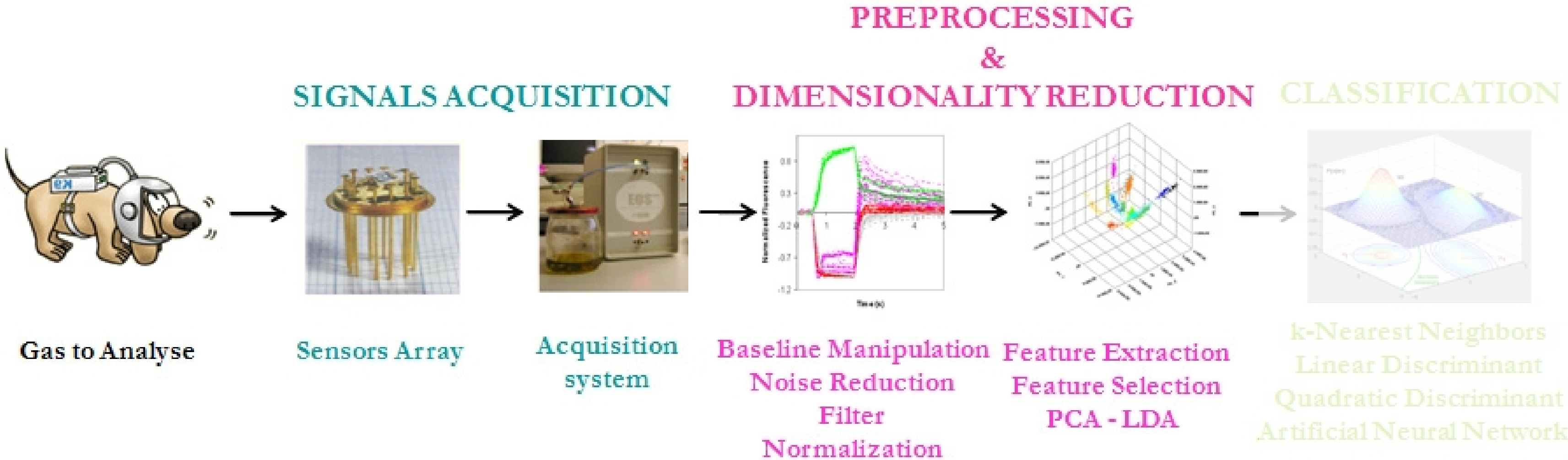
3. Classification and Validation

- Classification between the two classes “healthy” and “sick”



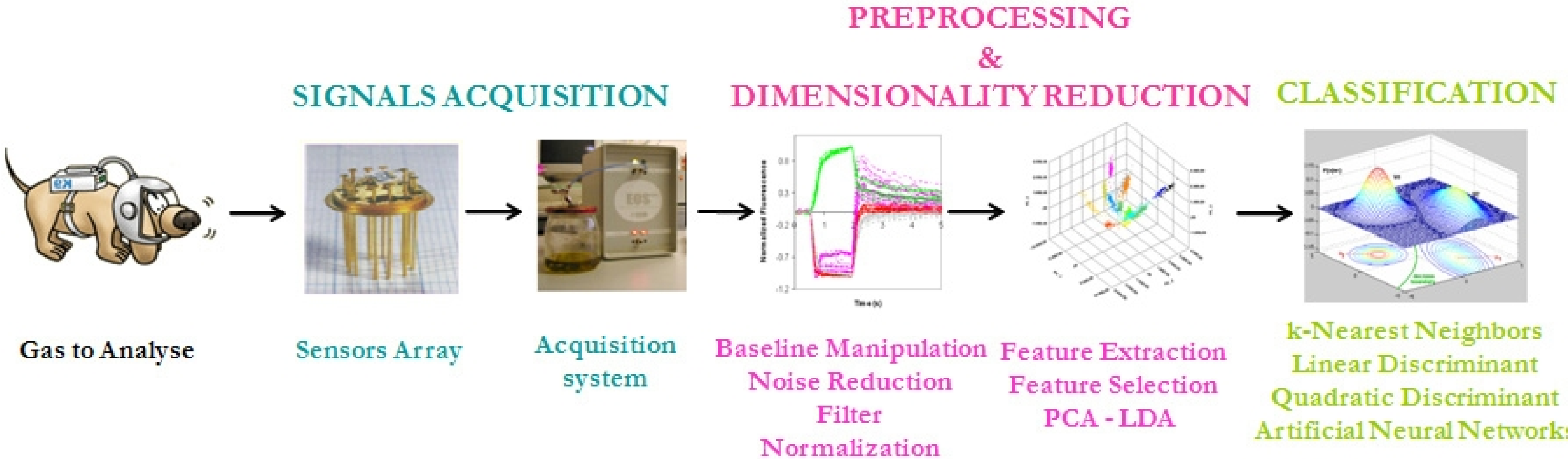
Electronic Nose

1. **Signal Acquisition**
   - Acquisition is done through a sensor array that measures a given physical or chemical quantity and convert it into an electrical signal
2. **Signal Processing**
   - **Preprocessing**: aimed to reduce the impact of noise
   - **Dimensionality Reduction**: reduce the dimensionality of the problem, enhancing classification performance
3. **Classification and Validation**
   - Classification between the two classes "healthy" and "sick"

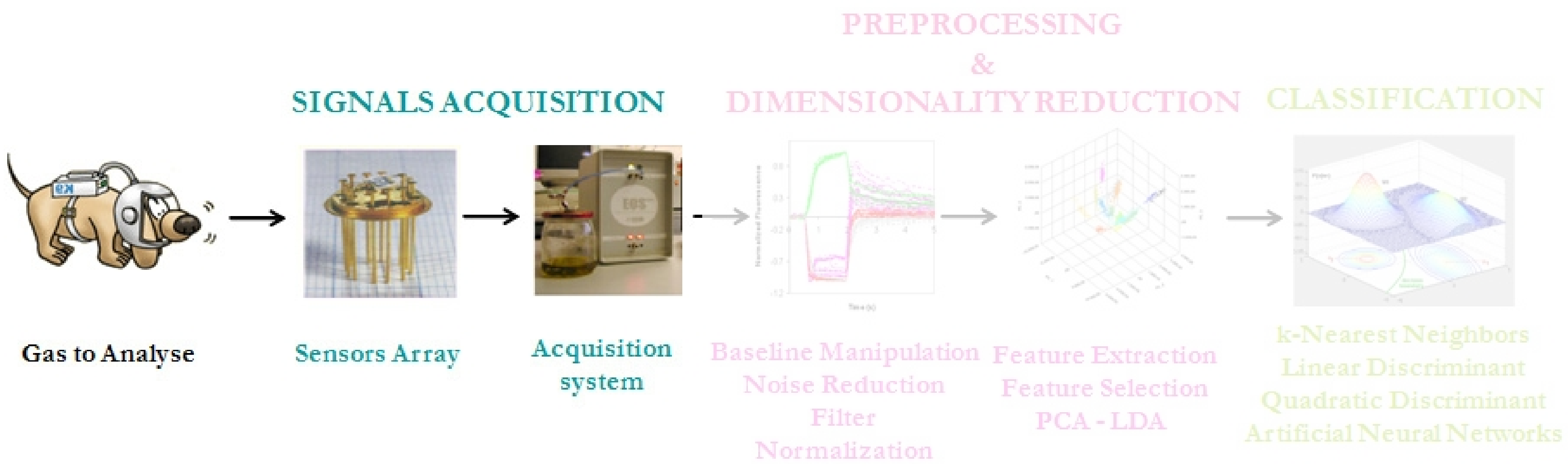


Electronic Nose

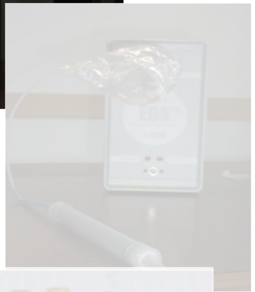
1. **Signal Acquisition**
   - Acquisition is done through a sensor array that measures a given physical or chemical quantity and convert it into an electrical signal
2. **Signal Processing**
   - **Preprocessing**: aimed to reduce the impact of noise
   - **Dimensionality Reduction**: reduce the dimensionality of the problem, enhancing classification performance
3. **Classification and Validation**
   - Classification between the two classes "healthy" and "sick"



Signal Acquisition

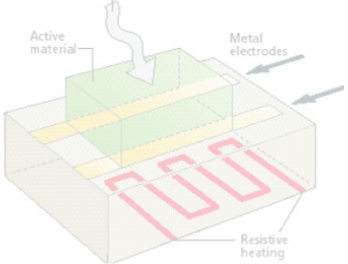


Signal Acquisition

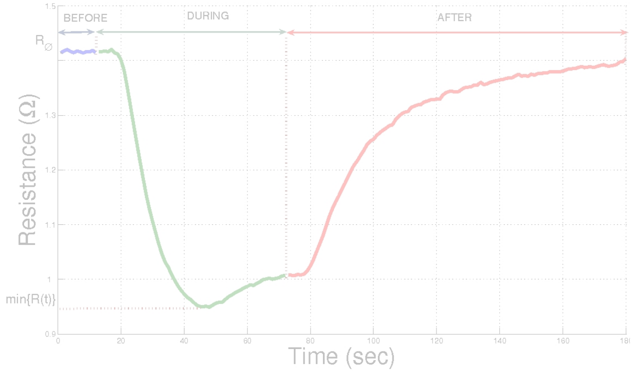


- The breath acquisition has been made inviting all volunteers to blow into a nalophan bag of approximately 400cm^3

- Then, the air contained in the bag was input into the electronic nose



- We used an array of **six MOS sensors** that react to gases with a **variation of resistance**

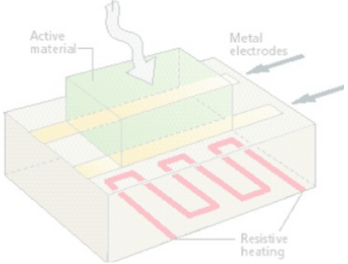


- The registered signal corresponds to the change of resistance through time produced by gas flow

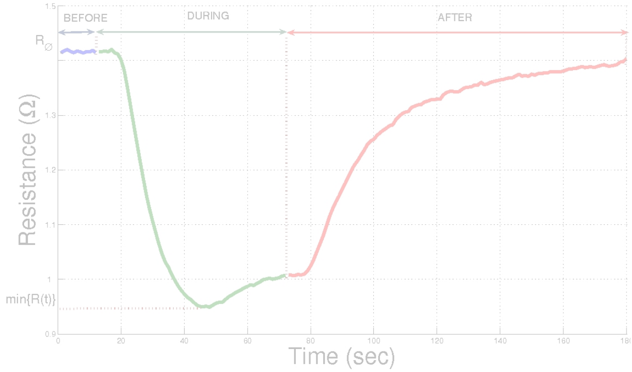
Signal Acquisition

- The breath acquisition has been made inviting all volunteers to blow into a nalophan bag of approximately 400cm^3

- Then, the air contained in the bag was input into the electronic nose



- We used an array of **six MOS sensors** that react to gases with a **variation of resistance**

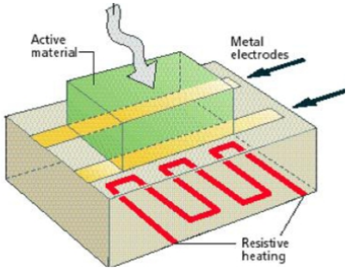


- The registered signal corresponds to the change of resistance through time produced by gas flow

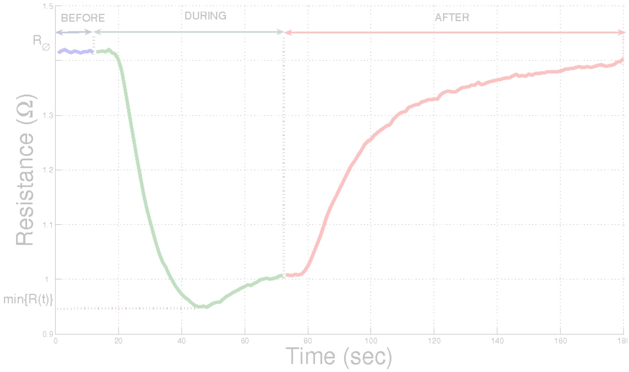
Signal Acquisition

- The breath acquisition has been made inviting all volunteers to blow into a nalophan bag of approximately 400cm^3

- Then, the air contained in the bag was input into the electronic nose



- We used an array of **six MOS sensors** that react to gases with a **variation of resistance**

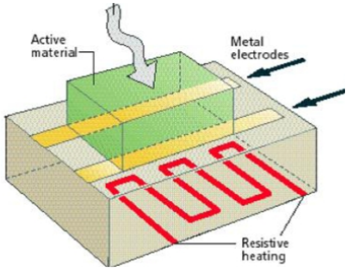


- The registered signal corresponds to the change of resistance through time produced by gas flow

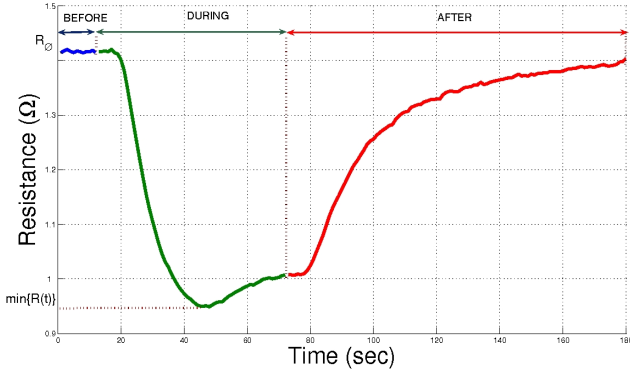
Signal Acquisition

- The breath acquisition has been made inviting all volunteers to blow into a nalophan bag of approximately 400cm^3

- Then, the air contained in the bag was input into the electronic nose



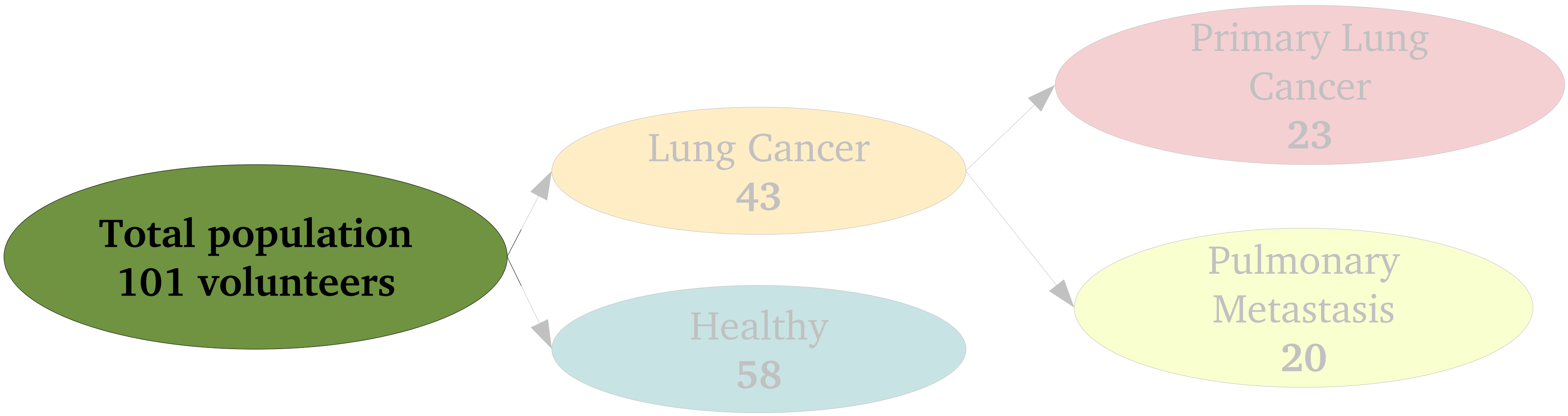
- We used an array of **six MOS sensors** that react to gases with a **variation of resistance**



- The registered signal corresponds to the change of resistance through time produced by gas flow

Involved Population

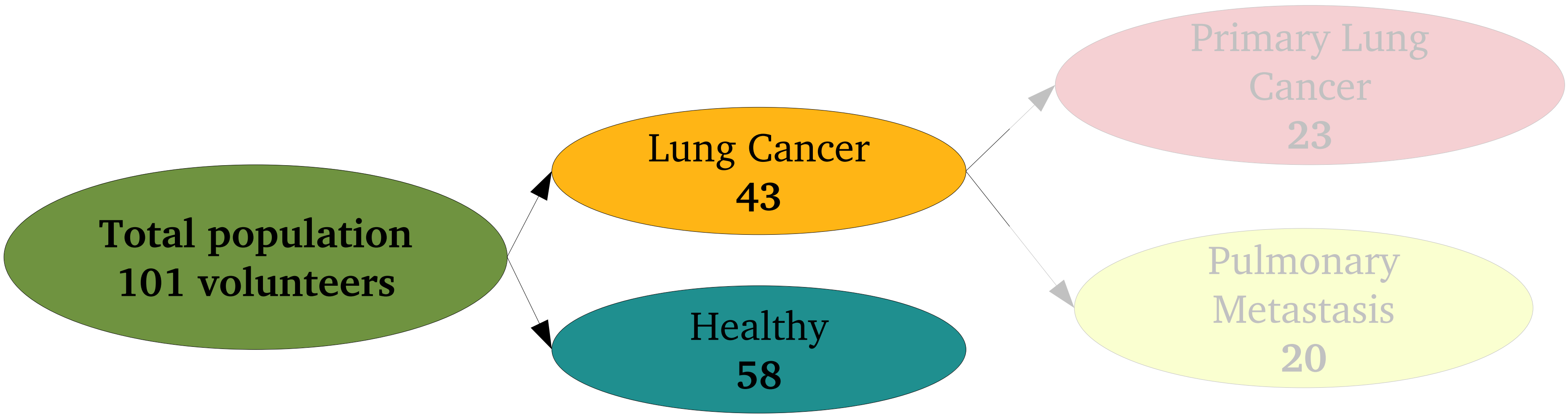
- We analyzed the breath of **101 volunteers**



- For each person we took two measures for a total of **202 measurements** (116 healthy, 86 diseased)

Involved Population

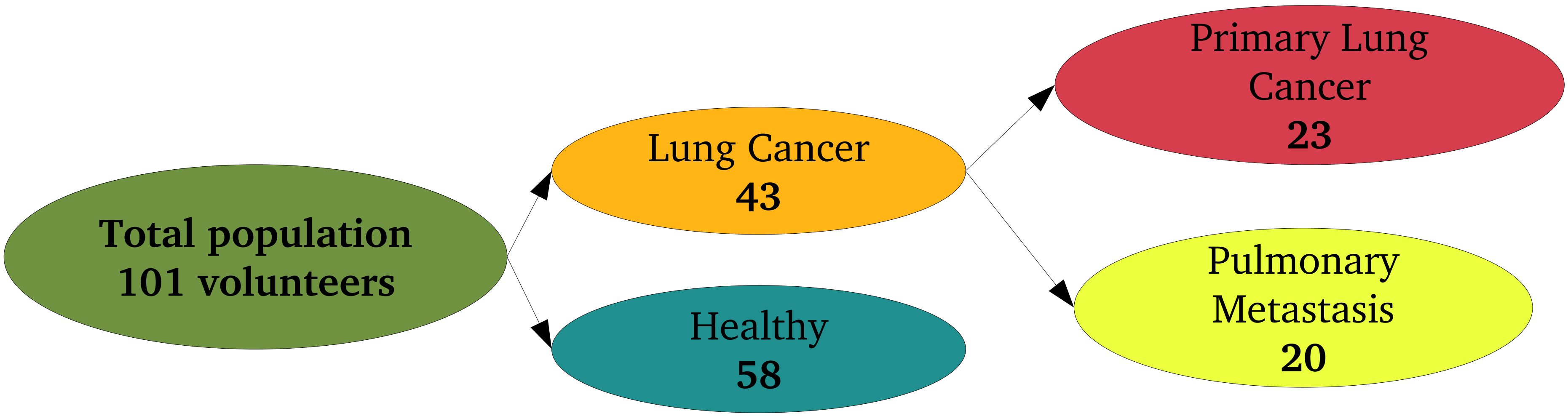
- We analyzed the breath of **101 volunteers**



- For each person we took two measures for a total of **202 measurements** (116 healthy, 86 diseased)

Involved Population

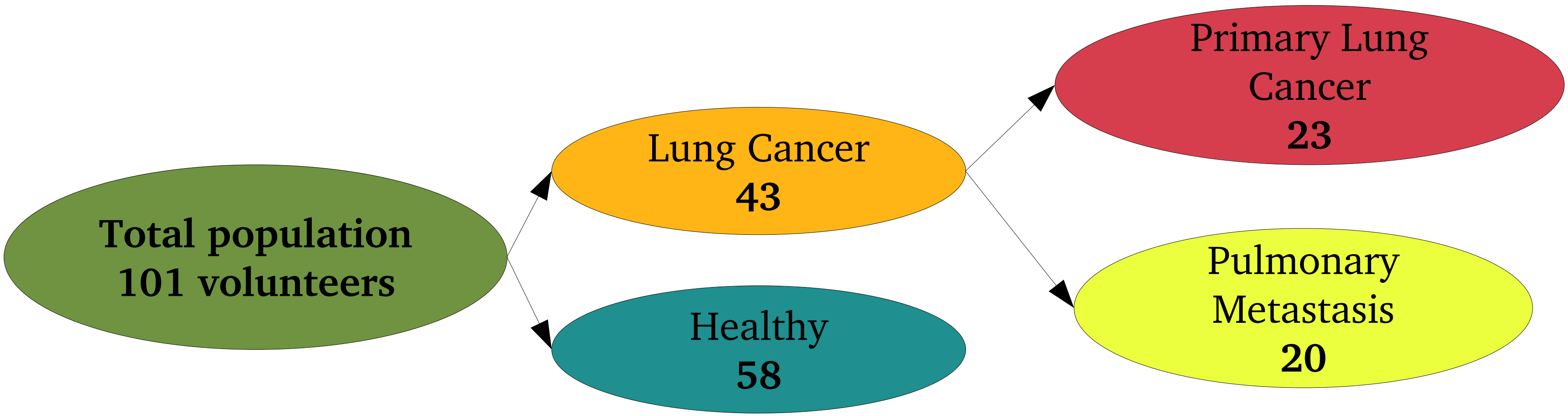
- We analyzed the breath of **101 volunteers**



- For each person we took two measures for a total of **202 measurements** (116 healthy, 86 diseased)

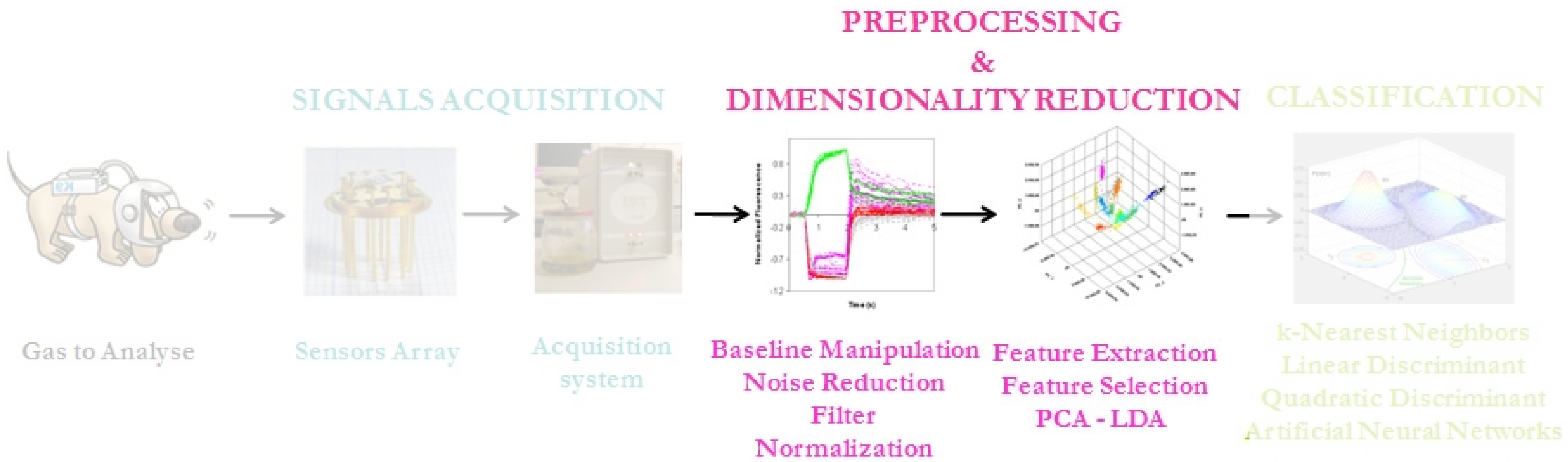
Involved Population

- We analyzed the breath of **101 volunteers**



- For each person we took two measures for a total of **202 measurements** (116 healthy, 86 diseased)

Pre-processing & Dimensionality Reduction



Signal pre-processing

- Manipulation of the **baseline**: transformation of the sensor response w.r.t. its baseline for the purpose of **drift compensation**
- Reduction of **humidity** effects
- **Normalization**: compensation for the scale difference among the six sensors
  - Each sensor has been forced to have zero mean and variance equal to 1

Signal pre-processing

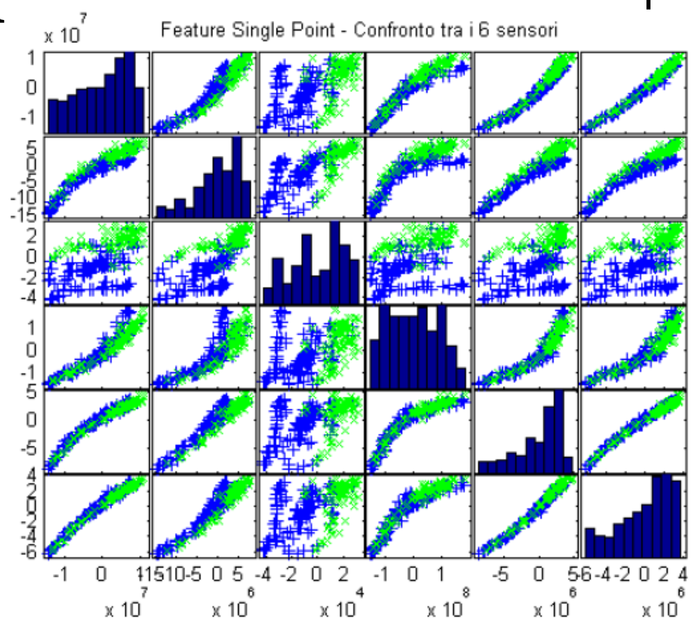
- Manipulation of the **baseline**: transformation of the sensor response w.r.t. its baseline for the purpose of **drift compensation**
- Reduction of **humidity** effects
- **Normalization**: compensation for the scale difference among the six sensors
  - Each sensor has been forced to have zero mean and variance equal to 1

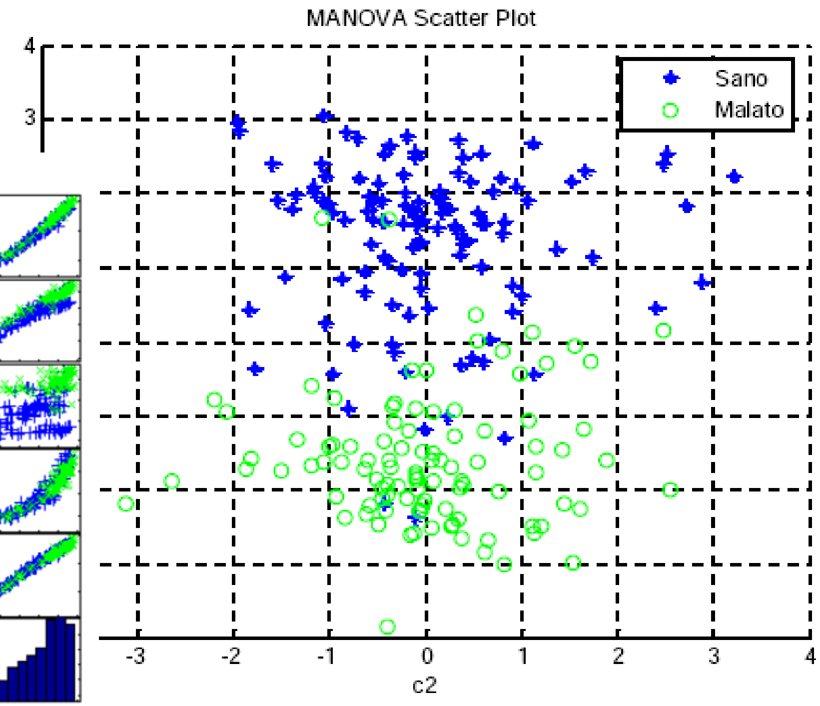
Signal pre-processing

- Manipulation of the **baseline**: transformation of the sensor response w.r.t. its baseline for the purpose of **drift compensation**
- Reduction of **humidity** effects
- **Normalization**: compensation for the scale difference among the six sensors
  - Each sensor has been forced to have zero mean and variance equal to 1

Dimensionality Reduction

- **Feature Selection**
  - Non-parametric test of Mann-Whitney-Wilcoxon
  - Scatter Plot
  - MANOVA





- Feature Extraction

Dimensionality Reduction

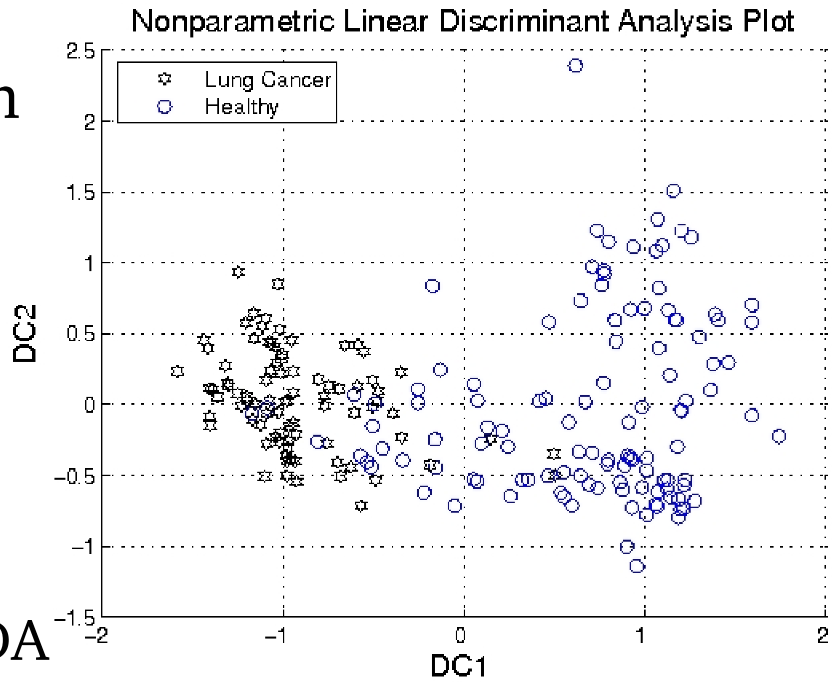
- **Feature Selection**
  - Non-parametric test of Mann-Whitney-Wilcoxon
  - Scatter Plot
  - MANOVA
- **Feature Extraction**
  - **Non Parametric Linear Discriminant Analysis NPLDA** (Fukunaga, 1983)
    - A generalization of Fisher's LDA
    - It **removes the unimodal gaussian assumption** by computing the between scatter matrix S_b using the *k*-NN rule
  - Best projection: 1st NPLDA component

Dimensionality Reduction

- **Feature Selection**
  - Non-parametric test of Mann-Whitney-Wilcoxon
  - Scatter Plot
  - MANOVA
- **Feature Extraction**
  - **Non Parametric Linear Discriminant Analysis NPLDA** (Fukunaga, 1983)
    - A generalization of Fisher's LDA
    - It **removes the unimodal gaussian assumption** by computing the between scatter matrix S_b using the *k*-NN rule
  - Best projection: 1st NPLDA component

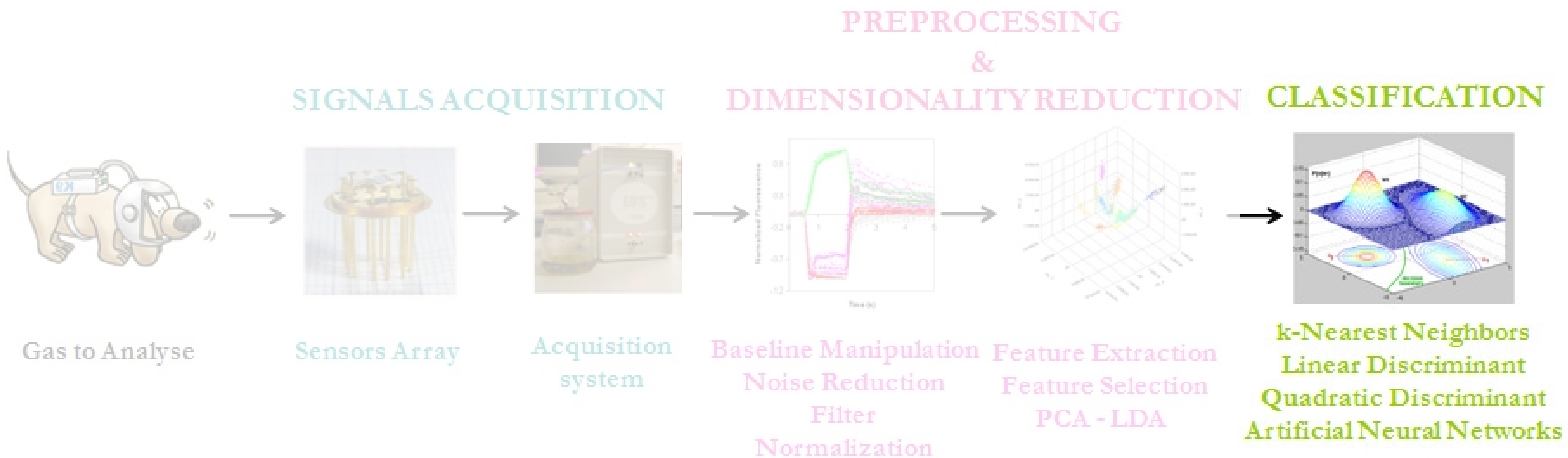
Dimensionality Reduction

- **Feature Selection**
  - Test of Mann-Whitney-Wilcoxon
  - Scatter Plot
  - MANOVA



- **Feature Extraction**
  - **NPLDA** (Fukunaga, 1983)
    - A generalization of Fisher's LDA
    - It **removes the unimodal gaussian assumption** by computing the between scatter matrix S_b using the *k*-NN rule
  - Best projection: 1st NPLDA component

Classification



Classification

- Different families of classifiers:
  - **Nearest Neighbor Classifiers (k-NN)**
    - **Classic k-NN**
    - **Modified k-NN** --> k=number of neighbors belonging all to the same class
    - **Fuzzy *k*-Nearest Neighbors** --> assigns a class membership function to each training and test samples
  - **Discriminant Functions** Classifiers
    - Linear
    - Quadratic
  - **Artificial Neural Network**
    - Feedforward Neural Network with one hidden layer

Classification

- Different families of classifiers:
  - **Nearest Neighbor Classifiers (k-NN)**
    - **Classic k-NN**
    - **Modified k-NN** --> k=number of neighbors belonging all to the same class
    - **Fuzzy *k*-Nearest Neighbors** --> assigns a class membership function to each training and test samples
  - **Discriminant Functions** Classifiers
    - Linear
    - Quadratic
  - **Artificial Neural Network**
    - Feedforward Neural Network with one hidden layer

Classification

- Different families of classifiers:
  - **Nearest Neighbor Classifiers (k-NN)**
    - **Classic k-NN**
    - **Modified k-NN** --> k=number of neighbors belonging all to the same class
    - **Fuzzy *k*-Nearest Neighbors** --> assigns a class membership function to each training and test samples
  - **Discriminant Functions** Classifiers
    - Linear
    - Quadratic
  - **Artificial Neural Network**
    - Feedforward Neural Network with one hidden layer

Classification

- Different families of classifiers:
  - **Nearest Neighbor Classifiers (k-NN)**
    - **Classic k-NN**
    - **Modified k-NN** --> k=number of neighbors belonging all to the same class
    - **Fuzzy *k*-Nearest Neighbors** --> assigns a class membership function to each training and test samples
  - **Discriminant Functions** Classifiers
    - Linear
    - Quadratic
  - Artificial Neural Network
    - Feedforward Neural Network with one hidden layer

Classification

- Different families of classifiers:
  - **Nearest Neighbor Classifiers (k-NN)**
    - **Classic k-NN**
    - **Modified k-NN** --> k=number of neighbors belonging all to the same class
    - **Fuzzy *k*-Nearest Neighbors** --> assigns a class membership function to each training and test samples
  - **Discriminant Functions** Classifiers
    - Linear
    - Quadratic
  - **Artificial Neural Network**
    - Feedforward Neural Network with one hidden layer

Results

- Performance has been evaluated through **confusion matrix** and the corresponding **performance indexes** (CI=95%)
- **Cross-validation:** *modified* leave-one-out
- We considered **different values for *k*** (*k*=1,3,5,9,101)

Classifier	NER	TPR	TNR	$PREC_{POS}$	$PREC_{NEG}$
Classic 9-NN	90.1%	89.5%	90.5%	87.5%	92.1%
Confidence Interval	[85.7-94.5]	[85.3-93.8]	[86.0-95.0]	[81.6-93.4]	[86.8-97.4]
Modified 9-NN	91.1%	91.9%	90.5%	87.8%	93.7%
Confidence Interval	[86.8-95.4]	[87.9-95.9]	[86.0-95.0]	[81.9-93.7]	[89.1-98.4]
Fuzzy k-NN	**92.6%**	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]
LD	89.6%	**96.5%**	84.5%	82.2%	**97.0%**
Confidence Interval	[85.0-94.2]	[93.7-99.3]	[79.1-89.9]	[75.2-89.1]	[93.9-100]
QD	**92.6%**	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]
ANN	91.6%	91.9%	**91.3793%**	**88.8%**	93.8%
Confidence Interval	[87.4-95.8]	[87.9-95.9]	[87.0-95.8]	[84.1-93.4]	[88.2-99.4]

Results

- Performance has been evaluated through **confusion matrix** and the corresponding **performance indexes** (CI=95%)
- **Cross-validation:** *modified* leave-one-out
- We considered **different values for *k*** (*k*=1,3,5,9,101)

Classifier	NER	TPR	TNR	$PREC_{POS}$	$PREC_{NEG}$
Classic 9-NN	90.1%	89.5%	90.5%	87.5%	92.1%
Confidence Interval	[85.7-94.5]	[85.3-93.8]	[86.0-95.0]	[81.6-93.4]	[86.8-97.4]
Modified 9-NN	91.1%	91.9%	90.5%	87.8%	93.7%
Confidence Interval	[86.8-95.4]	[87.9-95.9]	[86.0-95.0]	[81.9-93.7]	[89.1-98.4]
Fuzzy k-NN	**92.6%**	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]
LD	89.6%	**96.5%**	84.5%	82.2%	**97.0%**
Confidence Interval	[85.0-94.2]	[93.7-99.3]	[79.1-89.9]	[75.2-89.1]	[93.9-100]
QD	**92.6%**	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]
ANN	91.6%	91.9%	**91.3793%**	**88.8%**	93.8%
Confidence Interval	[87.4-95.8]	[87.9-95.9]	[87.0-95.8]	[84.1-93.4]	[88.2-99.4]

Results

- Performance has been evaluated through **confusion matrix** and the corresponding **performance indexes** (CI=95%)
- **Cross-validation:** *modified* leave-one-out
- We considered **different values for *k*** (k=*1,3,5,9,101*)

Classifier	NER	TPR	TNR	$PREC_{POS}$	$PREC_{NEG}$
Classic 9-NN	90.1%	89.5%	90.5%	87.5%	92.1%
Confidence Interval	[85.7-94.5]	[85.3-93.8]	[86.0-95.0]	[81.6-93.4]	[86.8-97.4]
Modified 9-NN	91.1%	91.9%	90.5%	87.8%	93.7%
Confidence Interval	[86.8-95.4]	[87.9-95.9]	[86.0-95.0]	[81.9-93.7]	[89.1-98.4]
Fuzzy k-NN	**92.6%**	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]
LD	89.6%	**96.5%**	84.5%	82.2%	**97.0%**
Confidence Interval	[85.0-94.2]	[93.7-99.3]	[79.1-89.9]	[75.2-89.1]	[93.9-100]
QD	**92.6%**	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]
ANN	91.6%	91.9%	**91.3793%**	**88.8%**	93.8%
Confidence Interval	[87.4-95.8]	[87.9-95.9]	[87.0-95.8]	[84.1-93.4]	[88.2-99.4]

Results

- Performance has been evaluated through **confusion matrix** and the corresponding **performance indexes** (CI=95%)
- **Cross-validation:** *modified* leave-one-out
- We considered **different values for *k*** (*k*=*1,3,5,9,101*)

Classifier	NER	TPR	TNR	$PREC_{POS}$	$PREC_{NEG}$
Classic 9-NN	90.1%	89.5%	90.5%	87.5%	92.1%
Confidence Interval	[85.7-94.5]	[85.3-93.8]	[86.0-95.0]	[81.6-93.4]	[86.8-97.4]
Modified 9-NN	91.1%	91.9%	90.5%	87.8%	93.7%
Confidence Interval	[86.8-95.4]	[87.9-95.9]	[86.0-95.0]	[81.9-93.7]	[89.1-98.4]
Fuzzy k-NN	**92.6**%	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]
LD	89.6%	**96.5**%	84.5%	82.2%	**97.0**%
Confidence Interval	[85.0-94.2]	[93.7-99.3]	[79.1-89.9]	[75.2-89.1]	[93.9-100]
QD	**92.6**%	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]
ANN	91.6%	91.9%	**91.3793**%	**88.8**%	93.8%
Confidence Interval	[87.4-95.8]	[87.9-95.9]	[87.0-95.8]	[84.1-93.4]	[88.2-99.4]

Results: Confusion Matrix

- Fuzzy k-NN demonstrated to be **robust** to k changes, keeping its results invariant

CONFUSION MATRIX		TRUE LABELS	
		Positive	Negative
ESTIMATED LABELS	Positive	82	11
	Negative	4	105

Indexes	Average Index	Confidence Interval $(CI = 95\%)$
Accuracy	92.6%	[88.5-96.7]
Sensitivity	95.3%	[91.8-98.9]
Specificity	90.5%	[86.0-95.0]
$PREC_{POS}$	88.2%	[82.3-94.1]
$PREC_{NEG}$	96.3%	[93.2-99.4]

Results: Confusion Matrix

- Fuzzy k-NN demonstrated to be **robust** to k changes, keeping its results invariant

CONFUSION MATRIX		TRUE LABELS	
		Positive	Negative
ESTIMATED LABELS	Positive	82	11
	Negative	4	105

Indexes	Average Index	Confidence Interval ($CI = 95\%$)
Accuracy	92.6%	[88.5-96.7]
Sensitivity	95.3%	[91.8-98.9]
Specificity	90.5%	[86.0-95.0]
$PREC_{POS}$	88.2%	[82.3-94.1]
$PREC_{NEG}$	96.3%	[93.2-99.4]

Results: Confusion Matrix

- Fuzzy k-NN demonstrated to be **robust** to k changes, keeping its results invariant

CONFUSION MATRIX		TRUE LABELS	
		Positive	Negative
ESTIMATED LABELS	Positive	82	11
	Negative	4	105

Indexes	Average Index	Confidence Interval ($CI = 95\%$)
Accuracy	92.6%	[88.5-96.7]
Sensitivity	95.3%	[91.8-98.9]
Specificity	90.5%	[86.0-95.0]
$PREC_{POS}$	88.2%	[82.3-94.1]
$PREC_{NEG}$	96.3%	[93.2-99.4]

Results: Confusion Matrix

- Fuzzy k-NN demonstrated to be **robust** to k changes, keeping its results invariant

CONFUSION MATRIX		TRUE LABELS	
		Positive	Negative
ESTIMATED LABELS	Positive	82	11
	Negative	4	105

Indexes	Average Index	Confidence Interval ($CI = 95\%$)
Accuracy	92.6%	[88.5-96.7]
Sensitivity	95.3%	[91.8-98.9]
Specificity	90.5%	[86.0-95.0]
PREC_{POS}	88.2%	[82.3-94.1]
PREC_{NEG}	96.3%	[93.2-99.4]

Classifiers Comparison

- Performing a **student t-test** between all pairs of classifiers, no relevant differences emerged
  - All implemented classifiers result comparable for the considered problem
- The **robustness** showed by Fuzzy k-NN to *k* changes is not verified in the classic and the modified *k*-NN, that leads to different results according to the value of *k*
- Moreover the output given by Fuzzy *k*-NN can be used to investigate the **relationship between these values and lung cancer stages**

Classifiers Comparison

- Performing a **student t-test** between all pairs of classifiers, no relevant differences emerged
  - All implemented classifiers result comparable for the considered problem
- The **robustness** showed by Fuzzy k-NN to *k* changes is not verified in the classic and the modified *k*-NN, that leads to different results according to the value of *k*
- Moreover the output given by Fuzzy *k*-NN can be used to investigate the **relationship between these values and lung cancer stages**

Classifiers Comparison

- Performing a **student t-test** between all pairs of classifiers, no relevant differences emerged
  - All implemented classifiers result comparable for the considered problem
- The **robustness** showed by Fuzzy k-NN to *k* changes is not verified in the classic and the modified *k*-NN, that leads to different results according to the value of *k*
- Moreover the output given by Fuzzy *k*-NN can be used to investigate the **relationship between these values and lung cancer stages**

Comparison to current diagnostic methods

- The use of an electronic nose as lung cancer diagnostic tool is reasonable if it gives some advantages compared to **current diagnostic techniques**

	Accuracy	Sensitivity	Specificity	$PREC_{POS}$	$PREC_{NEG}$
CAT Confidence Interval	Nd	75% [60-90]	66% [55-77]	Nd	Nd
PET Confidence Interval	Nd	91% [81-100]	86% [78- 94]	Nd	Nd
E-Nose Confidence Interval	92.6% [88.5-96.7]	95.3% [91.8-98.9]]	90.5% [86.0-95.0]	88.2% [82.3-94.1]	96.3% [93.2-99.4]

- Electronic nose results better in terms of **performance**
- Electronic noses are <u>cheaper</u>, <u>smaller</u> (and thus eventually portable), <u>very fast</u> and <u>non invasive</u> instruments

Comparison to current diagnostic methods

- The use of an electronic nose as lung cancer diagnostic tool is reasonable if it gives some advantages compared to **current diagnostic techniques**

	Accuracy	Sensitivity	Specificity	$PREC_{POS}$	$PREC_{NEG}$
CAT Confidence Interval	Nd	75% [60-90]	66% [55-77]	Nd	Nd
PET Confidence Interval	Nd	91% [81-100]	86% [78- 94]	Nd	Nd
E-Nose Confidence Interval	92.6% [88.5-96.7]	95.3% [91.8-98.9]]	90.5% [86.0-95.0]	88.2% [82.3-94.1]	96.3% [93.2-99.4]

- Electronic nose results better in terms of **performance**
- Electronic noses are cheaper, smaller (and thus eventually portable), very fast and non invasive instruments

Comparison to current diagnostic methods

- The use of an electronic nose as lung cancer diagnostic tool is reasonable if it gives some advantages compared to **current diagnostic techniques**

	Accuracy	Sensitivity	Specificity	$PREC_{POS}$	$PREC_{NEG}$
CAT	Nd	75%	66%	Nd	Nd
Confidence Interval		[60-90]	[55-77]		
PET	Nd	91%	86%	Nd	Nd
Confidence Interval		[81-100]	[78- 94]		
E-Nose	92.6%	95.3%	90.5%	88.2%	96.3%
Confidence Interval	[88.5-96.7]	[91.8-98.9]]	[86.0-95.0]	[82.3-94.1]	[93.2-99.4]

- Electronic nose results better in terms of **performance**
- Electronic noses are <u>cheaper</u>, <u>smaller</u> (and thus eventually portable), <u>very fast</u> and <u>non invasive</u> instruments

Extension and further directions of research

1. Improvement of **Sensors Technology:**
   - Development of longer-lyfe and more stable sensors
   - Development of hybrid systems, able to provide both selective and sensitive abilities
2. Improvement of **Olfactory Signal Analysis techniques and Classification Algorithms**
3. **Exploration of Informations hidden in the Olfactory Signal**
   - Analysis of the **olfactory patterns' changes due to surgery**
     - **Variation of VOCs** in the breath before and after the surgery
     - It could turn out to be useful for **therapy**
   - Individuation of **risk factors** connected to lung cancer
   - Involving a larger population, partitioning it according to different stages and using the Fuzzy output information, it would be possible to study the possibility of **early diagnose**

Extension and further directions of research

1. Improvement of **Sensors Technology**:
   - Development of longer-lyfe and more stable sensors
   - Development of hybrid systems, able to provide both selective and sensitive abilities
2. Improvement of **Olfactory Signal Analysis techniques and Classification Algorithms**
3. **Exploration of Informations hidden in the Olfactory Signal**
   - Analysis of the **olfactory patterns' changes due to surgery**
     - **Variation of VOCs** in the breath before and after the surgery
     - It could turn out to be useful for **therapy**
   - Individuation of **risk factors** connected to lung cancer
   - Involving a larger population, partitioning it according to different stages and using the Fuzzy output information, it would be possible to study the possibility of **early diagnose**

Extension and further directions of research

1. Improvement of **Sensors Technology**:
    - Development of longer-lyfe and more stable sensors
    - Development of hybrid systems, able to provide both selective and sensitive abilities
2. Improvement of **Olfactory Signal Analysis techniques and Classification Algorithms**
3. **Exploration of Informations hidden in the Olfactory Signal**
    - Analysis of the **olfactory patterns' changes due to surgery**
        - **Variation of VOCs** in the breath before and after the surgery
        - It could turn out to be useful for **therapy**
    - Individuation of **risk factors** connected to lung cancer
    - Involving a larger population, partitioning it according to different stages and using the Fuzzy output information, it would be possible to study the possibility of **early diagnose**

Extension and further directions of research

1. Improvement of **Sensors Technology**:
   - Development of longer-lyfe and more stable sensors
   - Development of hybrid systems, able to provide both selective and sensitive abilities
2. Improvement of **Olfactory Signal Analysis techniques and Classification Algorithms**
3. Exploration of Informations hidden in the Olfactory Signal
   - Analysis of the **olfactory patterns' changes due to surgery**
     - **Variation of VOCs** in the breath before and after the surgery
     - It could turn out to be useful for **therapy**
   - Individuation of **risk factors** connected to lung cancer
   - Involving a larger population, partitioning it according to different stages and using the Fuzzy output information, it would be possible to study the possibility of **early diagnose**

Extension and further directions of research

1. Improvement of **Sensors Technology**:
   - Development of longer-lyfe and more stable sensors
   - Development of hybrid systems, able to provide both selective and sensitive abilities
2. Improvement of **Olfactory Signal Analysis techniques and Classification Algorithms**
3. **Exploration of Informations hidden in the Olfactory Signal**
   - Analysis of the **olfactory patterns' changes due to surgery**
     - **Variation of VOCs** in the breath before and after the surgery
     - It could turn out to be useful for **therapy**
   - Individuation of **risk factors** connected to lung cancer
   - Involving a larger population, partitioning it according to different stages and using the Fuzzy output information, it would be possible to study the possibility of **early diagnose**

Extension and further directions of research

1. Improvement of **Sensors Technology**:
    - Development of <u>longer-lyfe</u> and <u>more stable</u> sensors
    - Development of <u>hybrid systems</u>, able to provide both selective and sensitive abilities
2. Improvement of **Olfactory Signal Analysis techniques and Classification Algorithms**
3. **Exploration of Informations hidden in the Olfactory Signal**
    - Analysis of the **olfactory patterns' changes due to surgery**
        - **Variation of VOCs** in the breath before and after the surgery
        - It could turn out to be useful for **therapy**
    - Individuation of **risk factors** connected to lung cancer
    - Involving a larger population, partitioning it according to different stages and using the Fuzzy output information, it would be possible to study the possibility of **early diagnose**

Extension and further directions of research

1. Improvement of **Sensors Technology**:
   - Development of longer-lyfe and more stable sensors
   - Development of hybrid systems, able to provide both selective and sensitive abilities
2. Improvement of **Olfactory Signal Analysis techniques and Classification Algorithms**
3. **Exploration of Informations hidden in the Olfactory Signal**
   - Analysis of the **olfactory patterns' changes due to surgery**
     - **Variation of VOCs** in the breath before and after the surgery
     - It could turn out to be useful for **therapy**
   - Individuation of **risk factors** connected to lung cancer
   - Involving a larger population, partitioning it according to different stages and using the Fuzzy output information, it would be possible to study the possibility of **early diagnose**

Extension and further directions of research

1. Improvement of **Sensors Technology**:
   - Development of longer-lyfe and more stable sensors
   - Development of hybrid systems, able to provide both selective and sensitive abilities
2. Improvement of **Olfactory Signal Analysis techniques and Classification Algorithms**
3. **Exploration of Informations hidden in the Olfactory Signal**
   - Analysis of the **olfactory patterns' changes due to surgery**
     - **Variation of VOCs** in the breath before and after the surgery
     - It could turn out to be useful for **therapy**
   - Individuation of **risk factors** connected to lung cancer
   - Involving a larger population, partitioning it according to different stages and using the Fuzzy output information, it would be possible to study the possibility of **early diagnose**

The end

Thanks!