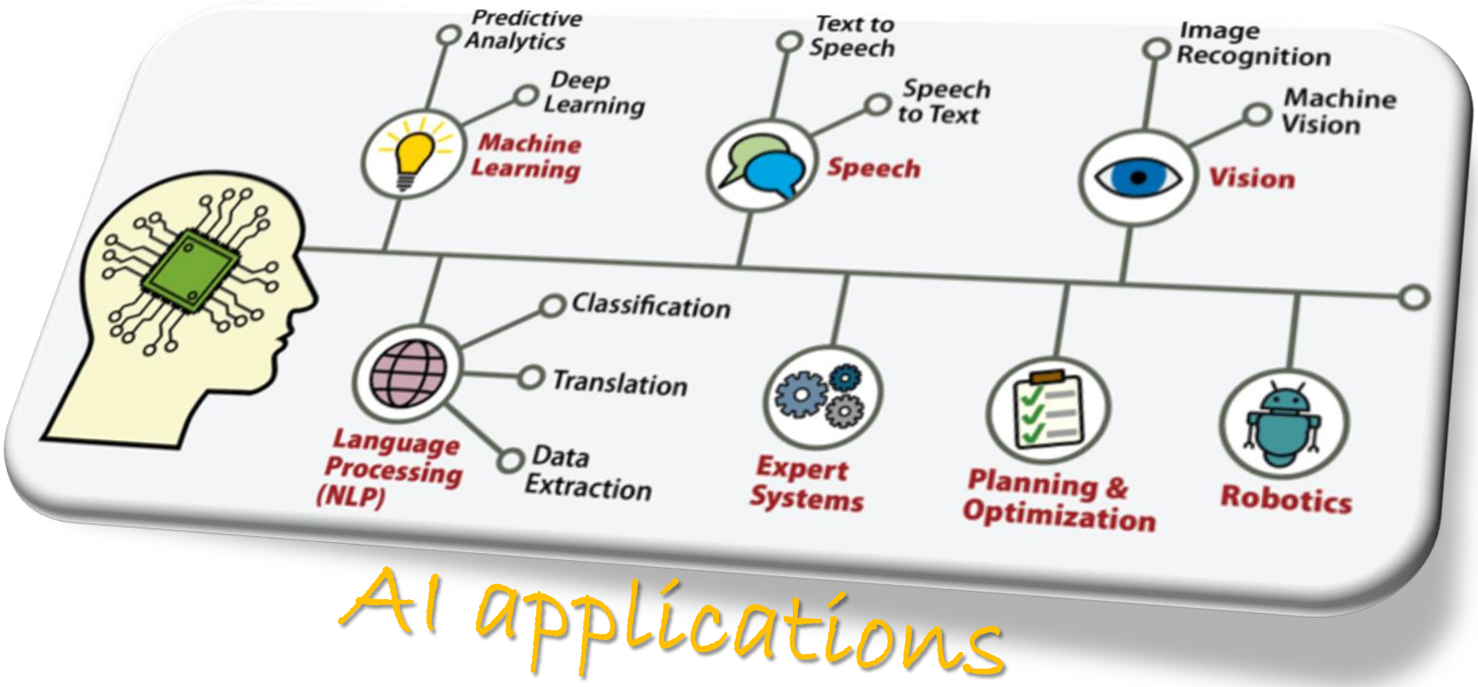
Artificial intelligence in the early diagnosis of diseases through breath analysis

Dr. **José S. Torrecilla,** MSc, MBA
Department of Chemical Engineering
Complutense University of Madrid
Madrid, Spain
jstorre@ucm.es

Table of contents

What does Artificial Intelligence mean?

What does Artificial Intelligence mean?



AI applications

Google has managed to create an artificial brain that plays a **video game better than a human player**, without having received any instructions on what the game is about.

Games

IBM's Watson computer defeated the best participants in the history of the American question and answer contest Jeopardy.

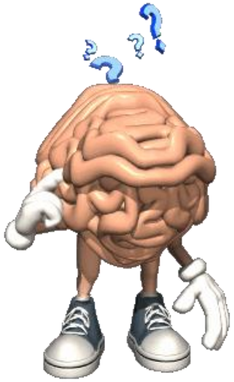
The University of Science and Technology of China have developed a machine capable of obtaining **better results than humans in an intelligence test**.

UNIVERSITY OF CAMBRIDGE

The University of Cambridge has an algorithm capable of **discovering your age, gender, profession, level of intelligence, political opinion and sexual orientation** through your facebook activity.

What does Artificial Intelligence mean?

<u>Artificial Intelligence</u> (AI) makes it possible for machines to **learn from experience**, adjust to new inputs, and perform tasks as humans do. Most of the examples of AI you hear about today are largely based on **predictive analysis and deep learning**. Through the use of these technologies, computers can be **trained** to perform **specific tasks by processing large amounts of data** and **recognizing patterns within.**



What does Artificial Intelligence mean?

Learn from experience

Predictive analysis, and deep learning

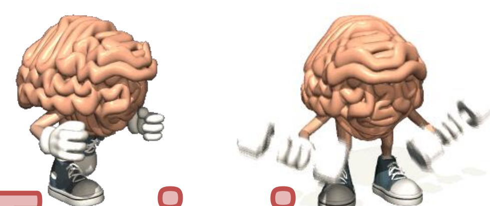
Recognizing patterns by processing data

What does Artificial Intelligence mean?

Learn from experience

Predictive analysis, and deep learning

Recognizing patterns by processing data



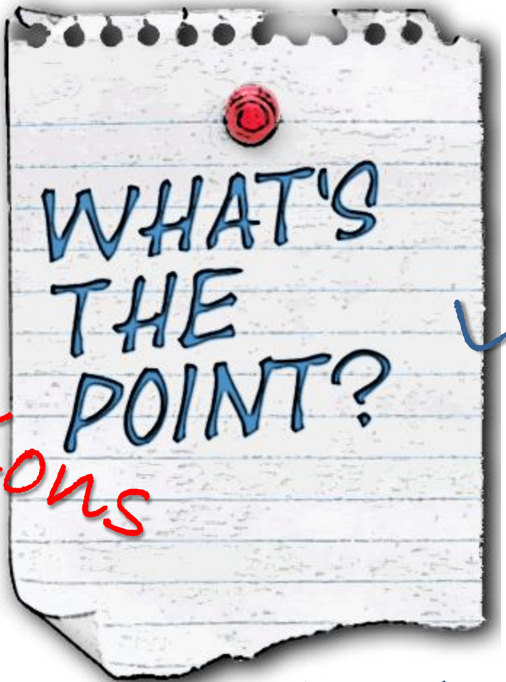
Training

Where is Artificial Intelligence needed?

What does Artificial Intelligence mean?

What does Artificial Intelligence mean?

Boring calculations



Large databases

Many potential benefits

The point is to develop intelligent applications

Where is Artificial Intelligence needed?

Our priorites!!!

Where is Artificial Intelligence needed?

What are the main applications which **make our lives better?**

What are the main applications which make our lives better?

M. Izquierdo, M. Lastra, E. González, J.C. Cancilla, M. Pérez, J.S. Torrecilla. Convolutional decoding of thermographic images to locate and **quantify honey adulterations**. Talanta (Accepted).

M. Izquierdo, M. Lastra, E. González, S. Padrana, J.C. Cancilla, J.S. Torrecilla. Visible imaging to convolutionally discern and **authenticate varieties of rice**. Food Control 110 **2020** 106971.

M. Lastra, E. González, M. Izquierdo, J. C. Cancilla, J.S. Torrecilla. Cognitive chaos on spectrofluorometric data to quantitatively unmask adulterations of a **PDO vinegar**. Food Control 108 **2020** 106860.

M. Lastra, M. Izquierdo, A. Torreblanca, R. Aroca, J.C. Cancilla, J.E. Sepulveda, J.S. Torrecilla. Cognitive fluorescence sensing to monitor the storage conditions and locate **adulterations of extra virgin olive oil**. Food Control 103 **2019** 48-58.

M. Izquierdo, A. Villa, M. Lastra, R. Aroca, J.C. Cancilla, J.S. Torrecilla. Intelligent real-time quantification of cheese whey in **rivers and reservoirs** in Madrid (Spain). Sensors & Actuators: B. Chemical 298 **2019** 126895.

What are the main applications which make our lives better?

What are the main applications which make our lives better?

How can we diagnose diseases?

What would you prefer as a diagnostic method?

What would you prefer as a diagnostic method?

Non-invasive

Painless

Easy

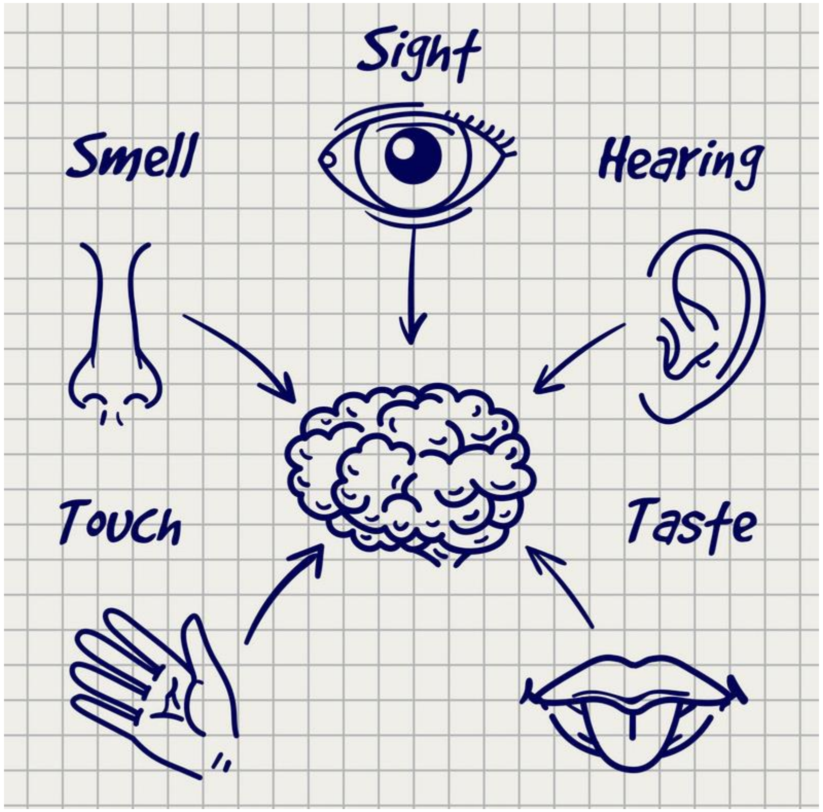
Broad

Fast

Safe

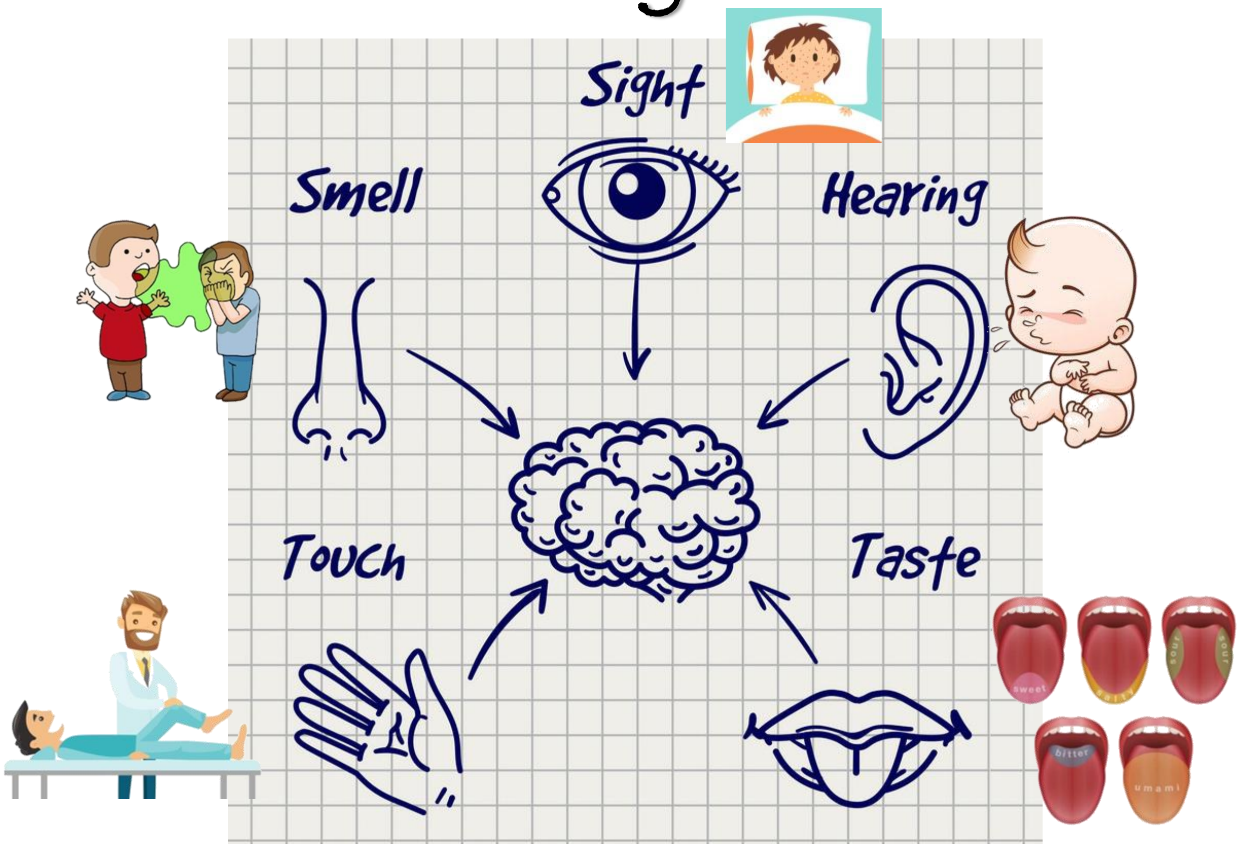
Reliable

How can we diagnose like this?



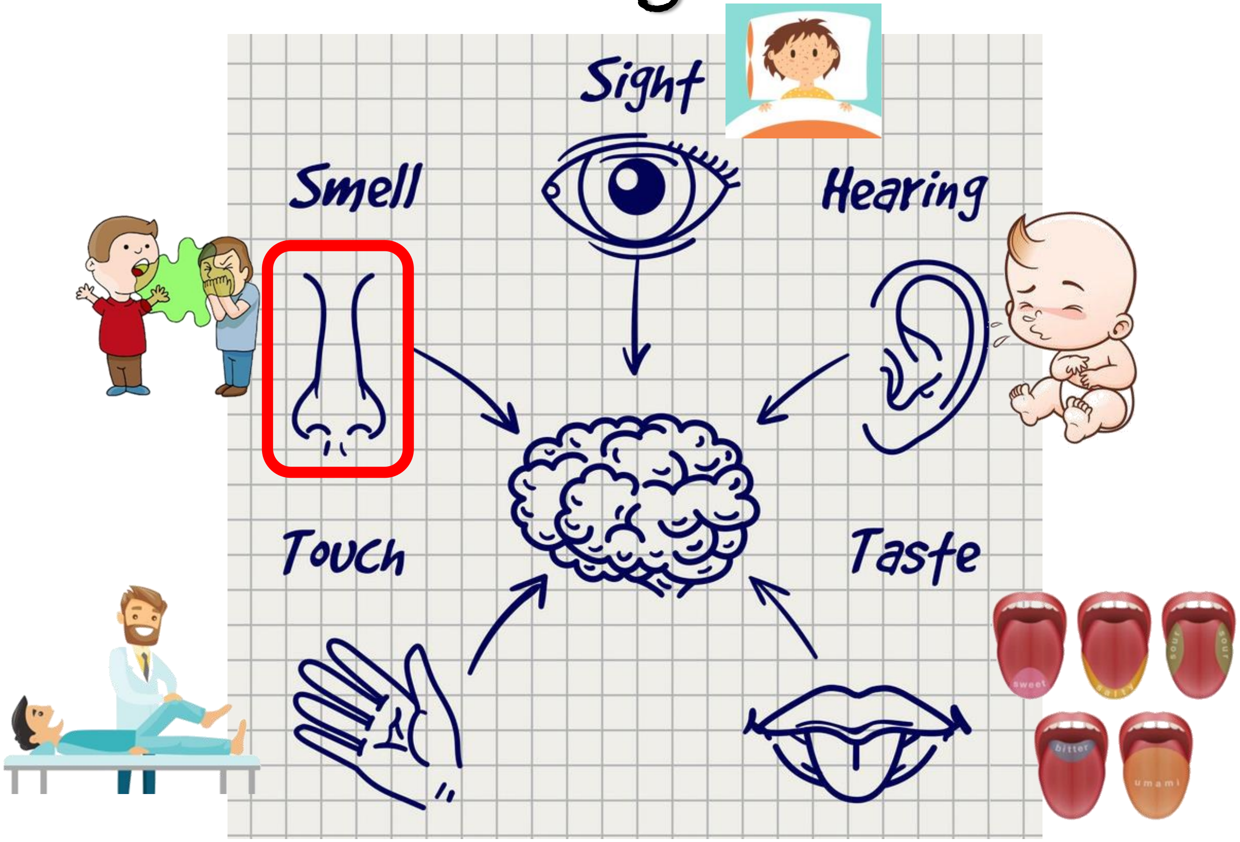
Please think how you would diagnose a disease using just one of your senses, then two of them, and then three, and…

How can we diagnose like this?



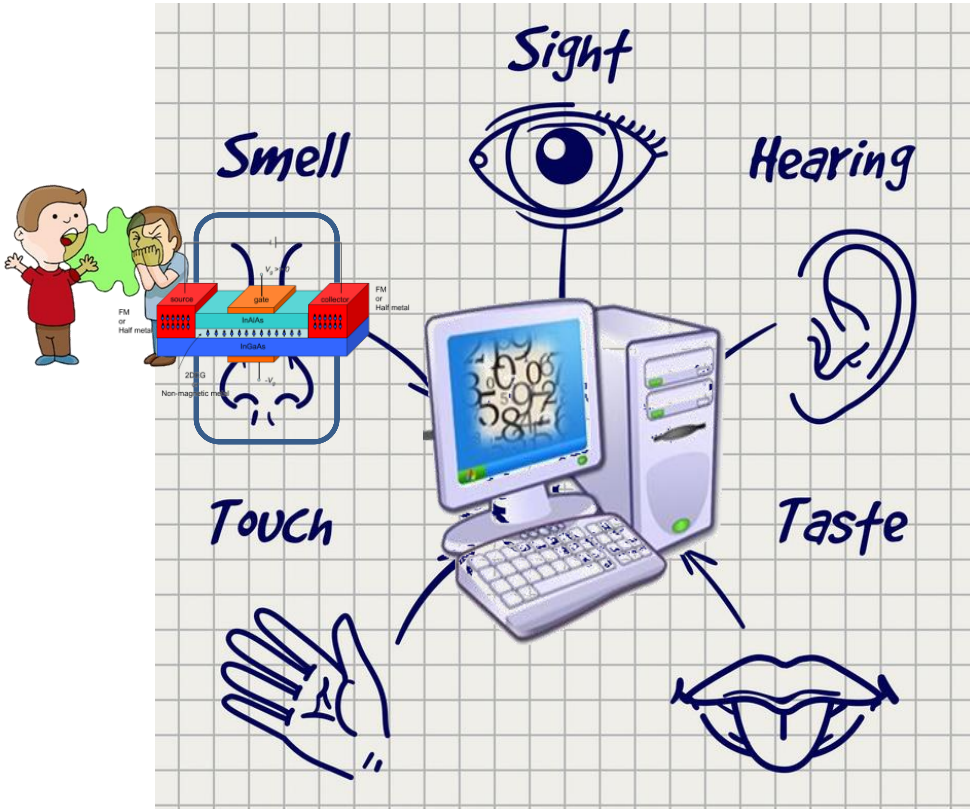
Please think how you would diagnose a disease using just one of your senses, then two of them, and then three, and…

How can we diagnose like this?

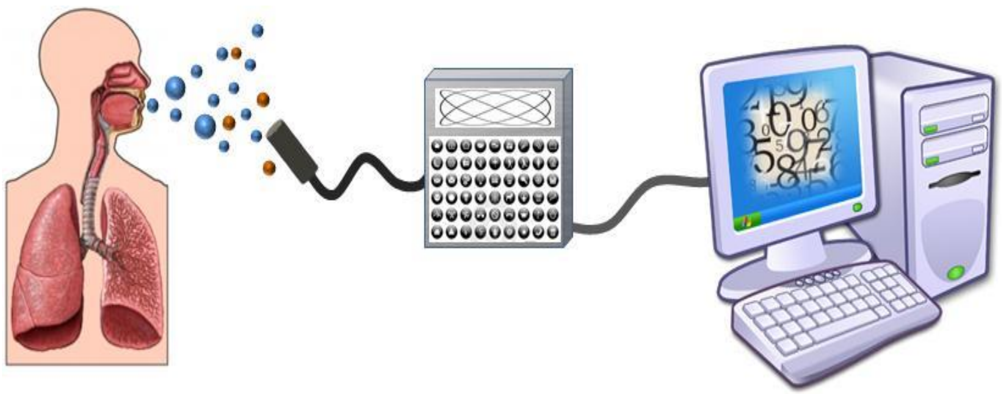


Please think how you would diagnose a disease using just one of your senses, then two of them, and then three, and...

How can we diagnose like this?



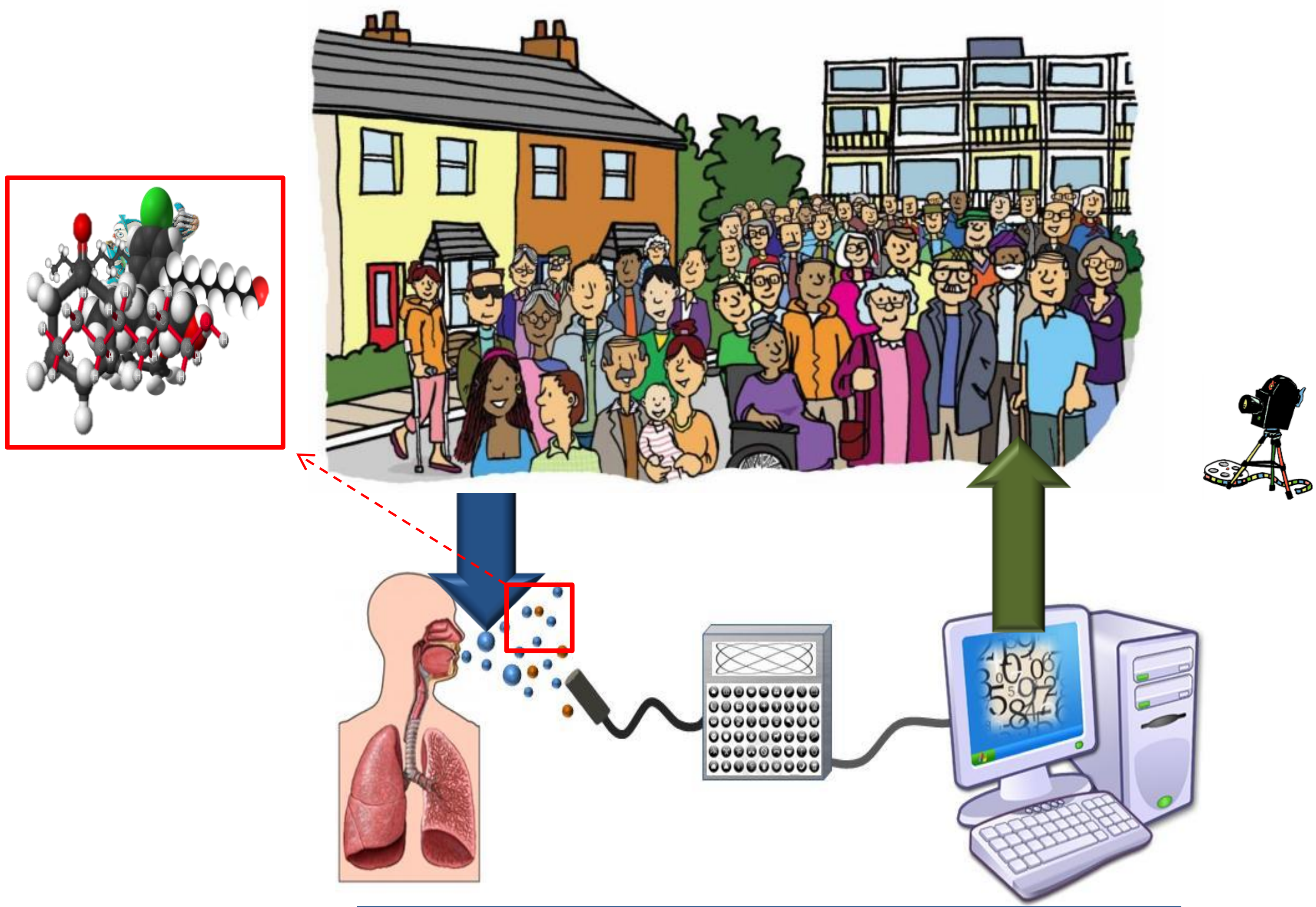
How can lung cancer be diagnosed in this way?



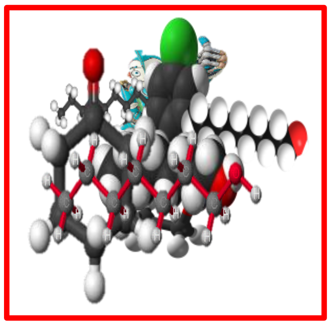
LCaos A Nanoscale Artificial Nose To Detect Lung Cancer

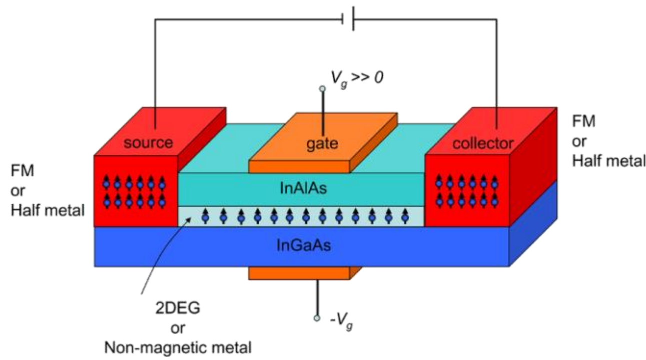
UNIVERSIDAD COMPLUTENSE MADRID

AlgoReach
Analyzing Complex Systems

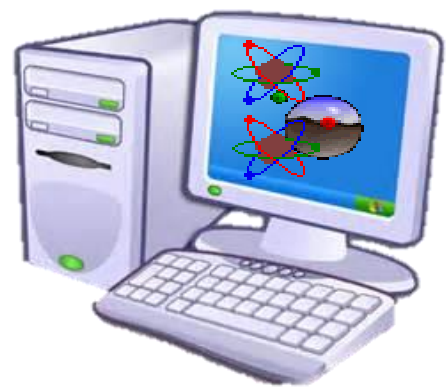


LCaos A Nanoscale Artificial Nose To Detect Lung Cancer





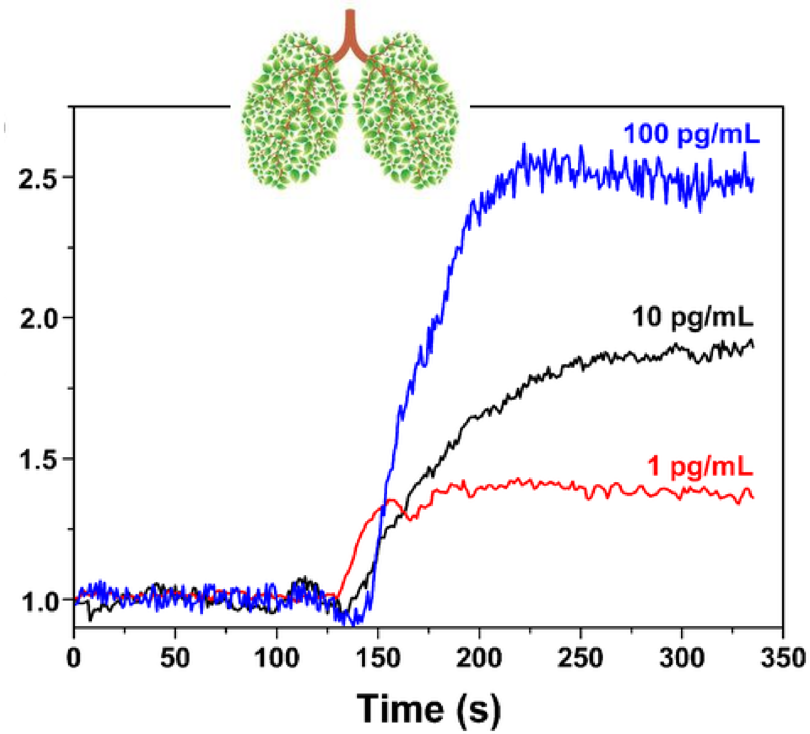
Cross-reactive and functionalized FET sensor

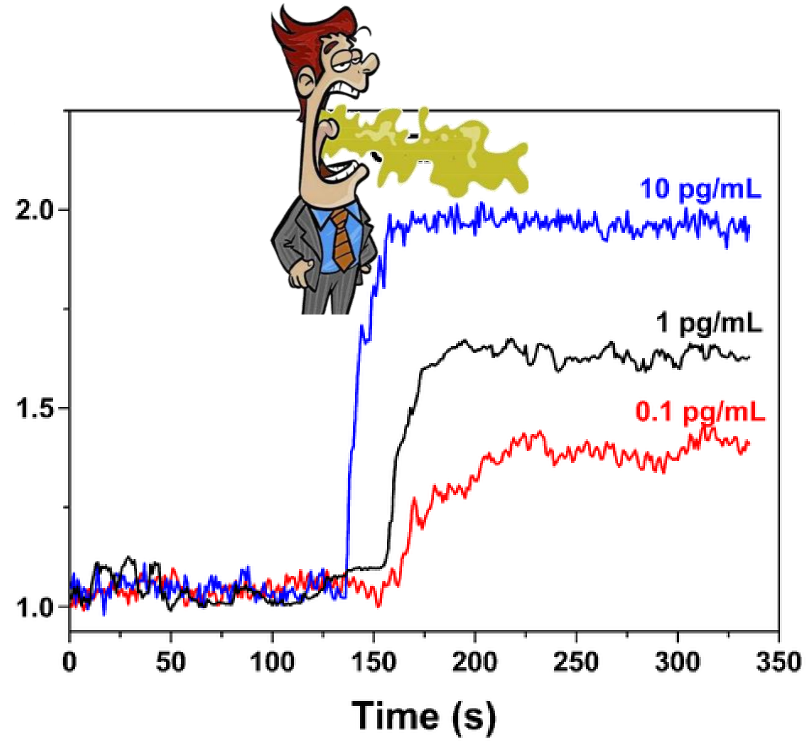


How much is there?

What is this?

Comparing Patterns





What are the mathematics within?

What are the mathematics within?

Advanced artificial intelligence based feature selection software. It enables the determination of:

1. **Most important features** or variables that characterize a system,
2. The optimal **estimative model.**

Discriminating features

What are the mathematics within?

1. Eliminating noise, anomalous data or redundant features

Statistical tools

1.1. ANN based calculations

The algorithm and its application is being patented

1.1.1. Statistical analysis

Comparison of estimation errors and/or success rates

Most important features

The best estimative model

BEST OF THE BEST

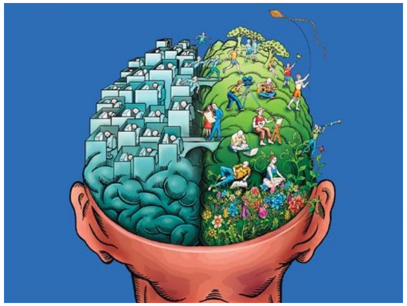
What are the mathematics within?

Statistical analysis of the database

Discriminating features

Estimative model

The nature of both algorithms are similar



Under patent request

What are the mathematics within?

- Parameters of Algorithms
- Activation function
- Transfer function

Statistical analysis of the database

Discriminating features

Estimative model

The nature of both algorithms are similar

- Training function
- Optimization processes
- Validation methods

Under patent request

What are the mathematics within?

Statistical analysis of the database

Discriminating features

Estimative model

The nature of both algorithms are similar

Powerful workstation

Under patent request

Advantage over other feature selection methods

Algo Pack

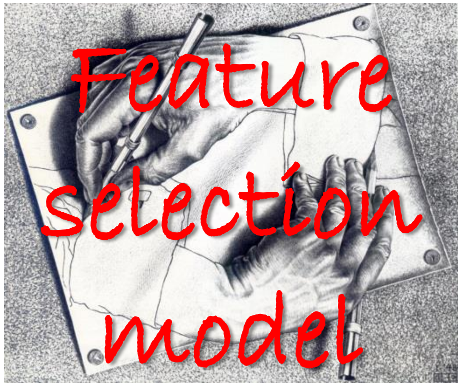
Feature selection algorithms

ReliefF, Kruskal-Wallis, Chi^2…

Independent variables selected

Estimative Model

AlgoTron



In contrast with other feature selection algorithms, this approach ensures that the optimal variables are found for the system as the model itself is used for the selecting task.

Main Results

NANO LETTERS
Letter
pubs.acs.org/NanoLett

Artificial Sensing Intelligence with Silicon Nanowires for Ultraselective Detection in the Gas Phase

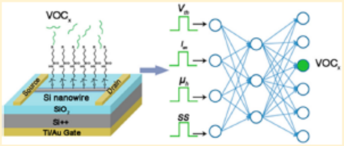
Bin Wang,[†] John C. Cancilla,[‡] Jose S. Torrecilla,[‡] and Hossam Haick*,[†]

[†]Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion-Israel Institute of Technology, Haifa 3200003, Israel

[‡]Department of Chemical Engineering, Complutense University of Madrid, Madrid 28040, Spain

Supporting Information

ABSTRACT: The use of molecularly modified Si nanowire field effect transistors (SiNW FETs) for selective detection in the liquid phase has been successfully demonstrated. In contrast, selective detection of chemical species in the gas phase has been rather limited. In this paper, we show that the application of artificial intelligence on deliberately controlled SiNW FET device parameters can provide high selectivity toward specific volatile organic compounds (VOCs). The obtained selectivity allows identifying VOCs in both single-component and multicomponent environments as well as estimating the constituent VOC concentrations. The effect of the structural properties (functional group and/or chain length) of the molecular modifications on the accuracy of VOC detection is presented and discussed. The reported results have the potential to serve as a launching pad for the use of SiNW FET sensors in real-world counteracting conditions and/or applications.



KEYWORDS: *Field effect transistor, sensor, artificial neural network, selective, volatile organic compound, vapor*

Silicon nanowire field effect transistors (SiNW FETs) are emerging as promising candidates for signal transduction and recognition of (bio)chemical species.[1−10] SiNW FETs provide several advantages over other sensing strategies, attributed to the methodical controllability over the sensing signals by means of gate voltages,[11,12] the ability to provide multiple device features to evaluate sensing signals,[13] the low power consumption of the SiNW FET, and to the extreme miniaturization features of the device dimensions.[1,5]

Implementation of the SiNW FETs for selective detection in liquid phase has been demonstrated. Functionalization of the Si NWs with amine/oxide,[5] biotin,[6] antigen,[7] or calmodulin[5] has allowed real-time selective (or specific) detection of pH, streptavidin, antibody, and calcium ions, respectively, in a liquid phase. Also, functionalization of SiNW FETs with tyrosine kinase,[8] antibody receptor,[9] or peptide nucleic acids (PNAs)[10] allowed selective detection of Ab1 tyrosine kinase, influenza-A virus, or DNA, respectively, in biological samples. In these applications, the receptors attached to the SiNW's surface bind specifically with the targeted (bio)molecule of interest, mostly through the "lock-and-key" approach,[14,15] and transduced via the FET platform as a change in the electrical signal(s).

In contrast to the liquid phase, it has been a challenge to obtain selective SiNW FET sensors for gaseous chemical species, such as, volatile organic compounds (VOCs) that are associated with environmental pollution, quality control, explosive materials, or various diseases.[4,16−22] As an intermediate way to increase the selectivity toward a specific VOC of interest, arrays of cross-reactive (that is, semiselective) sensors in conjugation with pattern recognition methods have been utilized.[4,17,23−30] This approach mimics the human olfactory system and is often referred to as "electronic nose".[4,17,23,25] Excellent progress in the electronic nose field has been achieved during the past two decades but increased selectivity remains elusive, especially when the targeted VOC exists in an environment that includes counteracting contaminants or VOCs. Theoretically, increasing the number of the cross-reactive sensors in the array might lead to some improvements in the selectivity.[4] Nevertheless, increasing the number of sensors increases the power consumption and complicates the device circuitry and the related computation parts. Additionally, the higher the number of the sensing elements, the higher the risk of overfitting toward the analyzed data (i.e., cases where the applied algorithm describes random errors or noise instead of the underlying relationship),[31] particularly in cases where the sample size is limited.

Here, we present an approach that allows highly accurate selective detection as well as estimation of VOC(s) concentrations in both single-component and multicomponent mixtures. The approach relies on the use of multiple independent parameters of a specific molecularly modified SiNW FET (e.g., voltage threshold, V_{th}, hole mobility, μ_h,

Received: November 21, 2013
Revised: January 9, 2014
Published: January 17, 2014

   dx.doi.org/10.1021/nl404335p | Nano Lett. 2014, 14, 933−938

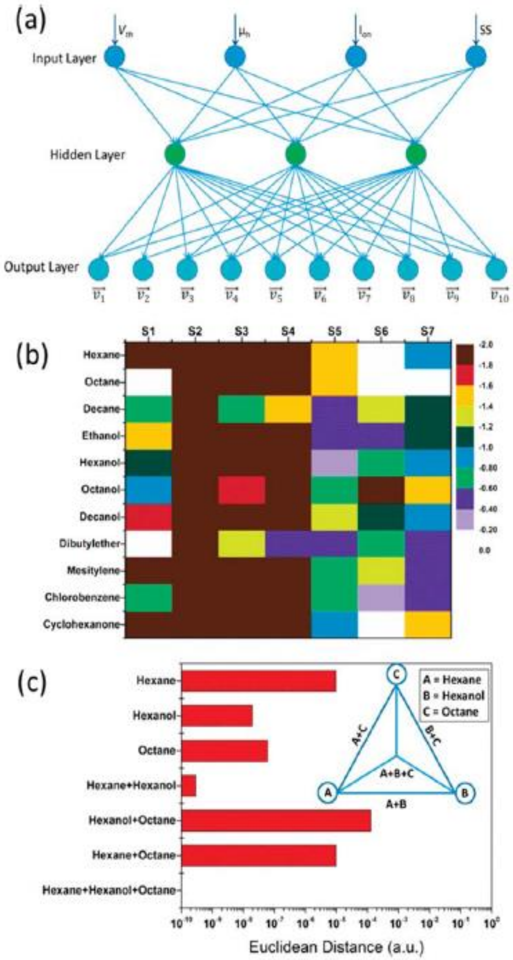


Figure 3. (a) Schematic illustration of an ANN model for VOC recognition. (b) Hot plots of logarithm Euclidean distance at VOC concentration of p_a/p_o = 0.08. (c) Euclidean distance of ANN outputs using sensor S3 to identify hexane, hexanol, octane, and their binary and ternary mixtures. All the tested VOC concentrations are at p_a/p_o = 0.08. Inset: Schematics of the relationship among single VOCs and their binary and ternary mixtures in ANN outputs.

ADVANCED MATERIALS
www.advmat.de
Materials Views
www.MaterialsViews.com

COMMUNICATION

A Highly Sensitive Diketopyrrolopyrrole-Based Ambipolar Transistor for Selective Detection and Discrimination of Xylene Isomers

Bin Wang, Tan-Phat Huynh, Weiwei Wu, Naseem Hayek, Thu Trang Do, John C. Cancilla, Jose S. Torrecilla, Masrur Morshed Nahid, John M. Colwell, Oz M. Gazit, Sreenivasa Reddy Puniredd, Christopher R. McNeill, Prashant Sonar, and Hossam Haick**

Xylenes, which include three isomers, i.e., *o*-xylene, *m*-xylene, and *p*-xylene, are important petrochemicals.[1] Individual and mixed xylene isomers are widely used as chemical intermediates to make polyesters and as solvents in the printing, rubber, and leather industries, as well as in household products (aerosol paints and lacquers).[1] Volatilization of xylenes in air can impair the human respiratory system, the central nervous system, liver, kidneys, eyes, and skin.[2] The U.S. National Institute for Occupational Safety and Health (NIOSH) has recommended long- and short-term exposure limits for xylenes of 100 and 150 ppm, respectively.[2a] Though xylene isomers have similar structural and physical properties, they have different industrial applications.[1] Moreover, the three xylene isomers are metabolized by different pathways in the body.[3] Hence, the detection and discrimination of xylene isomers is essential in environmental monitoring and healthcare.[2,3]

In principle, high-resolution gas chromatography (HRGC) can be used to detect and discriminate xylene isomers.[4] However, it requires expensive equipment and well-trained operators, making HRGC unsuitable for environmental monitoring. A number of methods, including mass sensor arrays,[5] chemiresistor sensor arrays,[6] sensors based on metal oxide semiconductors,[7] and cataluminescence sensor arrays,[8] have been developed for cost-effective detection of xylene isomers. Although selective detection has been achieved with arrays of sensors, their detection limits are always very high (e.g., 10^4, 10^3, and 400 ppm as reported in refs. [7c], [5], and [6] respectively). 5 ppm *p*-xylene has been detected using cobalt-doped zinc oxide nanowire-based sensors.[7b] However, such metal-oxide-semiconductor-based sensors work only at a relatively high temperature (>400 °C) and no selectivity regarding the other isomers of xylenes has been reported for these sensors. Therefore, detection of xylene isomers with high selectivity and sensitivity below the health-risk threshold level of 100 ppm remains a significant challenge. A reliable selective sensor for the detection of mixed xylene isomers, widely used as solvents in the chemical industry, is lacking as well.

We report here a sensor based on an ambipolar diketopyrrolopyrrole (DPP) copolymer field-effect transistor (FET) to selectively and sensitively detect xylene isomers at 40 ppm, which is below the NIOSH long-term exposure limit for xylene isomers. The rationale behind this approach is that an ambipolar FET that can operate alternatively in n-channel and p-channel modes has been adopted for the generation of double-sensing parameters in hole and electron changes, rather than using a unipolar FET.[9] Therefore, it is reasonable that an ambipolar FET sensor would provide more degrees of freedom to enhance the selective detection of structurally highly similar chemicals, e.g., between isomers. Indeed, by combining multiple sensing parameters derived from hole and electron effects obtained from the one ambipolar FET with a pattern-recognition algorithm, chemical fingerprints for each of the xylene isomers can be taken to clearly discriminate between structurally similar isomers. The ambipolar FET sensor can also selectively detect solvents of binary and ternary *o*-, *m*-, and *p*-xylene mixtures at different ratios via vapor sensing.

An alternating copolymer, poly(diketopyrrolopyrrole-ter-thiophene) referred to as **PDPPHD-T3** (**Figure 1**a), was synthesized by Stille coupling, using hexyldecyl-substituted DPP flanked with thiophene as previously reported[10] (see also "Experimental Section" below and Supporting Information, Section S1). Polymer characterization data can be found in Section S1 and Figure S1–S5 of the Supporting Information.

Dr. B. Wang, Dr. T.-P. Huynh, Dr. W. Wu, N. Hayek, Prof. O. M. Gazit, Prof. H. Haick
Department of Chemical Engineering and Russell Berrie Nanotechnology Institute
Technion-Israel Institute of Technology
Haifa 3200003, Israel
E-mail: hhossam@technion.ac.il

T. T. Do, Dr. J. M. Colwell, Prof. P. Sonar
School of Chemistry
Physics and Mechanical Engineering
Queensland University of Technology (QUT)
4001 Brisbane, Australia
E-mail: sonar.prashant@qut.edu.au

J. C. Cancilla, Prof. J. S. Torrecilla
Department of Chemical Engineering
Complutense University of Madrid
Madrid 28040, Spain

M. M. Nahid, Prof. C. R. McNeill
Department of Materials Science and Engineering
Monash University
Clayton, Victoria 3800, Australia

Dr. S. R. Puniredd
Institute of Materials Research and Engineering (IMRE)
Agency for Science, Technology and Research (A*STAR)
2 Fusionopolis Way, Innovis, Singapore 138634, Singapore

DOI: 10.1002/adma.201505641

ADVANCED MATERIALS
www.advmat.de
Materials Views
www.MaterialsViews.com

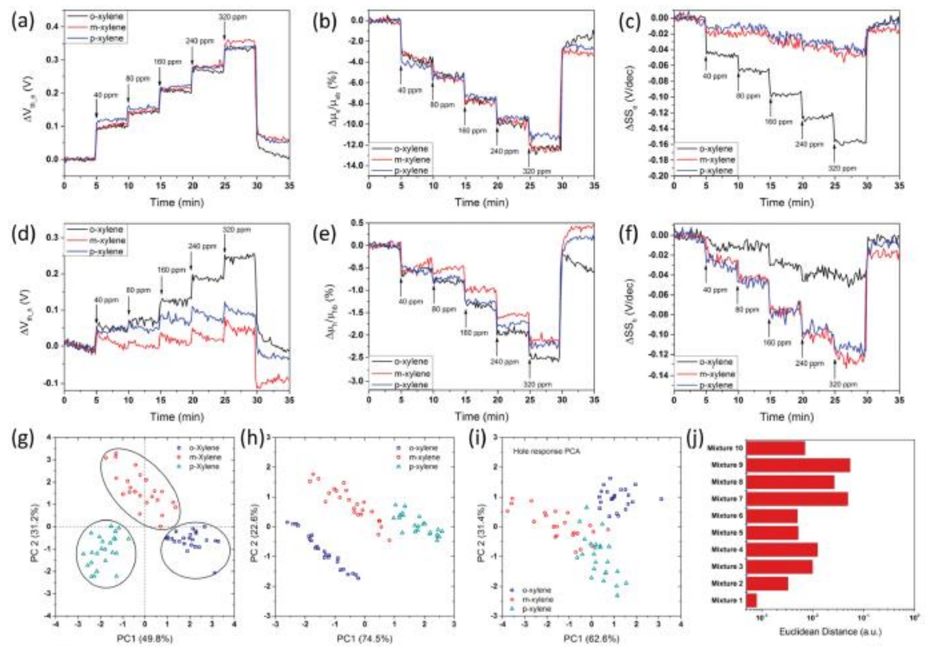


Figure 2. Time-dependent responses as expressed by: a) $\Delta V_{th\text{-}e}$, b) $\Delta\mu_e/\mu_{eb}$, c) ΔSS_e, d) $\Delta V_{th\text{-}h}$, e) $\Delta\mu_h/\mu_{hb}$, and f) ΔSS_h of a **PDPPHD-T3** FET after exposure to different xylene isomer vapors at varying concentrations. PCA score plots for the three xylene isomers at 40 ppm. g) Six sensing parameters ($\Delta V_{th\text{-}h}$, $\Delta V_{th\text{-}e}$, $\Delta\mu_h/\mu_{hb}$, $\Delta\mu_e/\mu_{eb}$, ΔSS_h, and ΔSS_e) obtained by one **PDPPHD-T3** sensor were used. h) Only electron-response parameters ($\Delta V_{th\text{-}e}$, $\Delta\mu_e/\mu_{eb}$, and ΔSS_e) were used. i) Only hole-response parameters ($\Delta V_{th\text{-}h}$, $\Delta\mu_h/\mu_{hb}$, and ΔSS_h) were used. j) Euclidean distance of ANN outputs using a sole **PDPPHD-T3** ambipolar FET sensor to identify *o*-xylene, *m*-xylene, *p*-xylene, and their binary and ternary mixture solvents.

UNIVERSIDAD COMPLUTENSE MADRID

AlgoReach
Analyzing Complex Systems

ACS Chemical Neuroscience
Research Article
Cite This: ACS Chem. Neurosci. 2017, 8, 2402-2413
pubs.acs.org/chemneuro

Exhaled Breath Markers for Nonimaging and Noninvasive Measures for Detection of Multiple Sclerosis

Yoav Y. Broza,[†] Lior Har-Shai,[‡,§] Raneen Jeries,[†] John C. Cancilla,[∥] Lea Glass-Marmor,[‡,§] Izabella Lejbkowicz,[‡,§] José S. Torrecilla,[∥] Xuelin Yao,[⊥] Xinliang Feng,[⊥] Akimitsu Narita,[⊥] Klaus Müllen,[⊥] Ariel Miller,[‡,§] and Hossam Haick*[,†]

[†]Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion−Israel Institute of Technology, Haifa 32000003, Israel
[‡]Division of Neuroimmunology and Multiple Sclerosis Center, Carmel Medical Center, Haifa 34362, Israel
[§]Rappaport Faculty of Medicine & Research Institute, Technion−Israel Institute of Technology, Haifa 31096, Israel
[∥]Departamento de Ingeniería Química, Facultad de Ciencias Químicas, Universidad Complutense de Madrid, 28040 Madrid, Spain
[⊥]Max Planck Institute for Polymer Research, Ackermannweg 10, D-55128 Mainz, Germany

Supporting Information

ABSTRACT: Multiple sclerosis (MS) is the most common chronic neurological disease affecting young adults. MS diagnosis is based on clinical characteristics and confirmed by examination of the cerebrospinal fluids (CSF) or by magnetic resonance imaging (MRI) of the brain or spinal cord or both. However, neither of the current diagnostic procedures are adequate as a routine tool to determine disease state. Thus, diagnostic biomarkers are needed. In the current study, a novel approach that could meet these expectations is presented. The approach is based on noninvasive analysis of volatile organic compounds (VOCs) in breath. Exhaled breath was collected from 204 participants, 146 MS and 58 healthy control individuals. Analysis was performed by gas-chromatography mass-spectrometry (GC-MS) and nanomaterial-based sensor array. Predictive models were derived from the sensors, using artificial neural networks (ANNs). GC-MS analysis revealed significant differences in VOC abundance between MS patients and controls. Sensor data analysis on training sets was able to discriminate in binary comparisons between MS patients and controls with accuracies up to 90%. Blinded sets showed 95% positive predictive value (PPV) between MS-remission and control, 100% sensitivity with 100% negative predictive value (NPV) between MS not-treated (NT) and control, and 86% NPV between relapse and control. Possible links between VOC biomarkers and the MS pathogenesis were established. Preliminary results suggest the applicability of a new nanotechnology-based method for MS diagnostics.

KEYWORDS: *Biomarkers, breath, diagnosis, multiple sclerosis, sensors, spectrometry, volatile organic compound*

INTRODUCTION

Multiple sclerosis (MS) is the most common chronic neurological disease affecting young adults, with usual onset at the age of 20−50 years old, and is more common in women (about 3 times more than in men).[1,2] The diagnosis of MS is based on clinical characteristics. History and physical examination are the most significant factors in determining the diagnosis of the disease, but there are several ancillary laboratory tests that assist the clinician in establishing the diagnosis.[1,2]

The most commonly employed tool for confirmation of MS diagnosis is magnetic resonance imaging (MRI) of brain or spinal cord.[1,3] MRI is considered the test-of-choice to support the clinical diagnosis of MS, applying the McDonald's diagnostic criteria, with demonstration of dissemination in time and space of characteristic lesions is necessary to establish the diagnosis.[1] Characteristic lesions on MRI are found in 70−95% of patients diagnosed with MS.[4] Most lesions detected on MRI correlate with MS pathologic lesions. However, in some cases, detected lesions may be extensive on MRI show only small plaques on pathological examination.[5] Furthermore, central nervous system (CNS) lesions that are related to other etiologies (e.g., ischemia, systemic lupus erythematosus (SLE), Behcet's disease, vasculitis, etc.) and can appear very similar to MS lesions on MRI, making them often difficult to be distinguished from one another.[6] Moreover, MRI is an expensive procedure that is not available for routine medical care and follow-up.

Another ancillary test is analysis of the presence of IgG oligoclonal bands (OCBs), extracted from cerebrospinal fluids (CSF), by means of electrophoresis. This test involves performing a lumbar puncture, which is an invasive procedure, accompanied by pain and discomfort to the patient and potentially infection, and can lead to neurological complica-

Received: May 16, 2017
Accepted: August 2, 2017
Published: August 2, 2017

DOI: 10.1021/acschemneuro.7b00181
ACS Chem. Neurosci. 2017, 8, 2402−2413

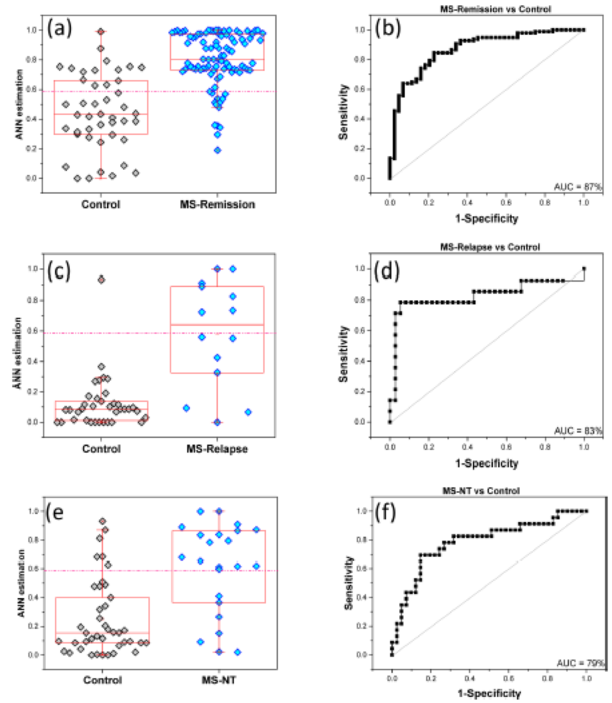


Figure 3. Results of the ANN model training set in distinguishing between samples. The following comparisons are presented: (a, d) MS remission (n = 95) vs control (n = 44); (b, e) MS relapse (n = 14) vs control (n = 37); (c, f) MS NT (n = 23) vs control (n = 41), represented in a cluster formation (left) and a ROC curve (right). The red dashed line is the set threshold per training.

AI Diagnosis Results Conclusion

ACS NANO
www.acsnano.org
ARTICLE

Silicon Nanowire Sensors Enable Diagnosis of Patients *via* Exhaled Breath

Nisreen Shehada,[†] John C. Cancilla,[‡] Jose S. Torrecilla,[‡] Enrique S. Pariente,[‡] Gerald Brönstrup,[§] Silke Christiansen,[§,•] Douglas W. Johnson,[∥] Marcis Leja,[⊥,#,∇] Michael P. A. Davies,[¶] Ori Liran,[○] Nir Peled,[○] and Hossam Haick*[,†]

[†]Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion–Israel Institute of Technology, Haifa 3200003, Israel

[‡]Department of Chemical Engineering, Complutense University of Madrid, Madrid 28040, Spain

[§]Max-Planck-Institute for the Science of Light, Günther-Scharowsky-Strasse 1, Erlangen 91058, Germany

[•]Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, Hahn-Meitner-Platz 1, 14109 Berlin, Germany

[∥]Florida Radiation Oncology Group, Department of Radiation Oncology, Baptist Cancer Institute (BCI), 1235 San Marco Boulevard, Suite 100, Jacksonville, Florida 32207, United States

[⊥]Faculty of Medicine, University of Latvia, 19 Raiņa boulv., LV1586 Riga, Latvia

[#]Department of Research, Riga East University Hospital, 6 Linezera iela, LV1006 Riga, Latvia

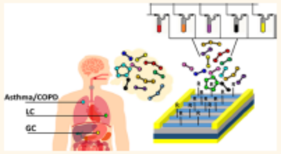
[∇]Digestive Diseases Centre GASTRO, Riga, Latvia. 6 Linezera iela, LV1006 Riga, Latvia

[¶]Molecular & Clinical Cancer Medicine, University of Liverpool, William Duncan Building, 6 West Derby Street, Liverpool L7 8TX, United Kingdom

[○]Thoracic Cancer Unit, Davidoff cancer center, Petah Tiqwa and the Tel Aviv University, Tel Aviv, Israel

Supporting Information

ABSTRACT: Two of the biggest challenges in medicine today are the need to detect diseases in a noninvasive manner and to differentiate between patients using a single diagnostic tool. The current study targets these two challenges by developing a molecularly modified silicon nanowire field effect transistor (SiNW FET) and showing its use in the detection and classification of many disease breathprints (lung cancer, gastric cancer, asthma, and chronic obstructive pulmonary disease). The fabricated SiNW FETs are characterized and optimized based on a training set that correlate their sensitivity and selectivity toward volatile organic compounds (VOCs) linked with the various disease breathprints. The best sensors obtained in the training set are then examined under real-world clinical conditions, using breath samples from 374 subjects. Analysis of the clinical samples show that the optimized SiNW FETs can detect and discriminate between almost all binary comparisons of the diseases under examination with >80% accuracy. Overall, this approach has the potential to support detection of many diseases in a direct harmless way, which can reassure patients and prevent numerous unpleasant investigations.



KEYWORDS: *nanowire, sensor, disease, cancer, diagnosis, breath, volatile organic compound*

Physicians are always challenged by the need to give the correct diagnosis as early in the onset of a disease as possible, whether the disease-related symptoms are absent or not evident.[1] Symptoms are not always characteristic of one particular disease.[2] For example, patients with different respiratory diseases, such as malignant or benign tumors or substantially less severe diseases, may have similar symptoms, for example, cough, chest pain, difficulty breathing, *etc.* These symptoms may be characteristic of lung cancer (LC), pneumonia, asthma, and chronic obstructive pulmonary disease (COPD).[1,2] Therefore, it is of particular clinical importance to find a diagnostic tool capable of distinguishing between these diseases. A diagnostic tool that does not involve any needle, surgery, active material(s), or radioactive exposure would have a significant benefit.

A highly promising approach that could meet the aforementioned need is based on the detection and classification of the disease breathprint, namely the chemical profiles of highly

Received: May 11, 2016
Accepted: July 6, 2016
Published: July 6, 2016

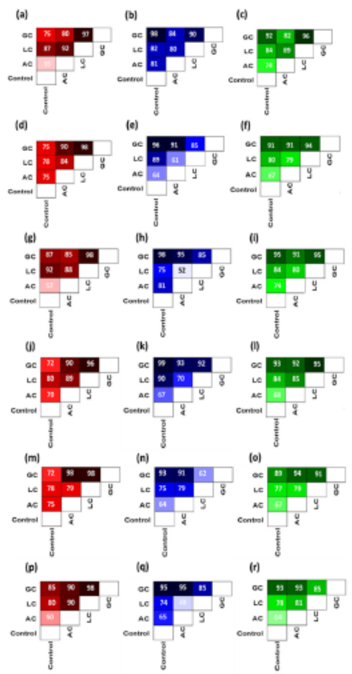


Figure 5. Sensitivity (red; a, d, g, j, m, p), specificity (blue; b, e, h, k, n, q), and accuracy (green; c, f, i, l, o, r) for binary comparisons between the response of S1–S6 to breath samples of volunteers from all groups: LC, GC, AC, and healthy control. Classification was obtained by ANN analysis (each row corresponds to a different sensor: top row (a, b, c) corresponds to S1, the second (d, e, f) to S2, *etc.*). In this classification, all stages of a disease were considered as one group (without staging). Numbers in white are the values, for example, the value in the red box relating GC and control is the sensitivity of their comparison, the value in the blue box is the specificity, and that in the green box is the accuracy. NOTE: In comparing any of the groups with the "control" group, the latter was clearly considered the "negative" group, while the others were considered the "positive" groups. In comparing one of the cancer groups (LC or GC) to AC, the former was considered the "positive" and AC the "negative" group. Finally, in comparing the 2 types of cancer, LC was taken as the "positive" group, and GC the "negative" group. The values of the comparisons are reviewed in Supporting Information, Table S3.

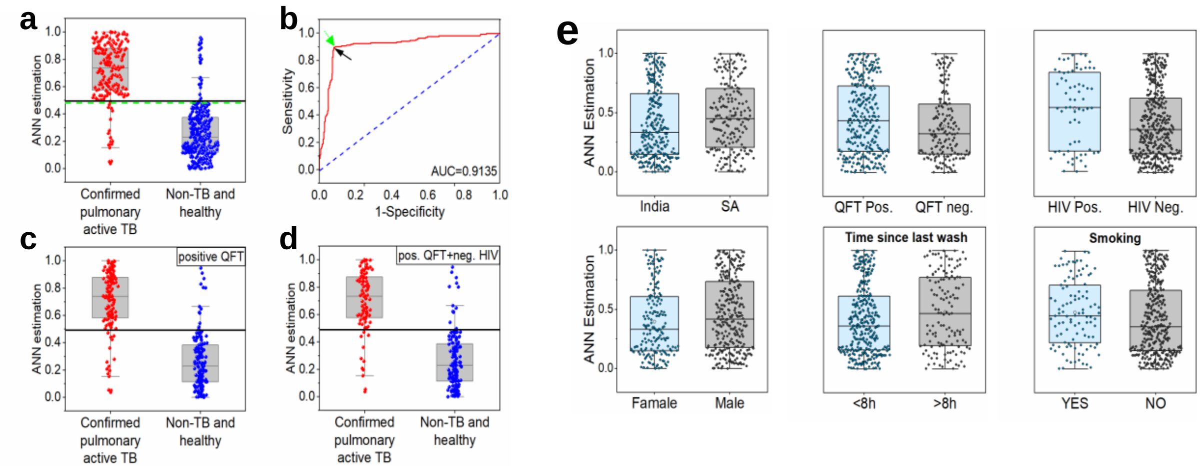
AI Diagnosis Results Conclusion

Unique profile of skin volatile biomarkers for non-invasive and point-of-care detection of tuberculosis

Vishinkin R.[1], Mansour E.[1], Falk Fish[1], Esmail A.[2], Kumar P.[3], Gharaa A.[1], Cancilla J. C.[4], Torrecilla J. S.[5], Dheda K.[2], Singh S.[3] and Haick H.[1]

1. *Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion-Israel Institute pf Technology, Haifa 3200003, Israel.*
2. *Lung Infection and Immunology Unit, Division of Pulmonology, Department of Medicine, University of Cape Town, Cape Town, South Africa.*
3. *Division of Clinical Microbiology and Molecular Medicine, Department of Laboratory Medicine, All India Institute of Medical Sciences, New Delhi 110029, India.*
4. *Scintillon Institute, San Diego, CA, USA*
5. *Department of Chemical Engineering, Complutense University of Madrid, Madrid 28040, Spain.*

Tuberculosis (TB) is an infectious disease that threatens >10 million people annually. Despite dramatic advances in TB diagnostics, millions of patients continue to receive incomplete or delayed diagnosis, as the physical signs and symptoms of TB are not specific. Many existing bio-diagnostic tests are slow, have low sensitivity and/or specificity, and can be too expensive or complex for resource-limited settings. Here we report on a new diagnostics pathway and related sensing tools enabling a non-invasive, affordable, fast and highly accurate way of detecting TB. The approach relies on TB-specific volatile organic compounds (VOCs) that are detected, identified and quantified from air trapped above the skin (the "skin headspace"). A specifically-designed nanoarray could translate these findings into a point-of-care diagnosis and discriminating between confirmed active pulmonary TB status and controls with 92.3% sensitivity, 80.4% specificity and 84.9% overall accuracy, fulfilling the requirements of a triage TB test according to the World Health Organization (WHO) guidelines.



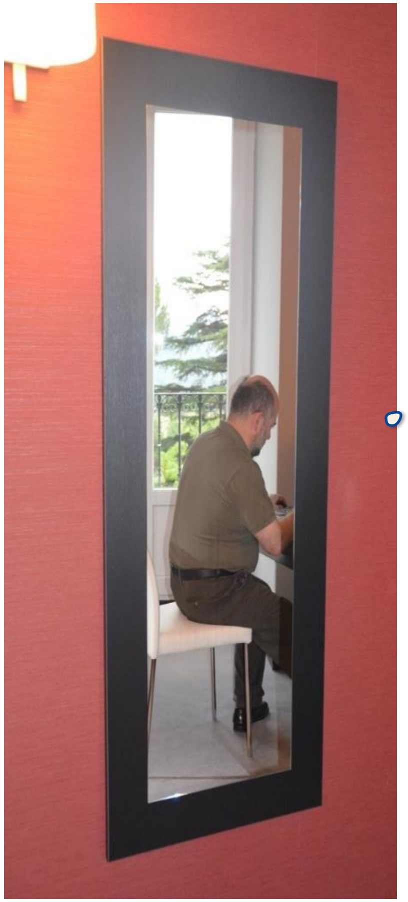
Artificial neural network results of the global classifier. **a**, boxplot of ANN estimation score of the model. Each point represents one sample. The central solid line represents Youden's cut-point. The dashed line represents modified cut-point to meet the WHO criterion of sensitivity above 90%. Samples above the cut-point are classified as Non-TB and healthy and samples below the cut-point are classified as confirmed pulmonary active TB samples. **b,** ROC curve of model with mathematically optimal Youden's cut-point (black arrow) and the modified (green arrow). **c**, boxplot of ANN estimation score of the model for subpopulation with QFT positive status. **d**, boxplot of ANN estimation score of the model for subpopulation with QFT positive and HIV negative statuses. **e**, boxplots of ANN estimation scores for selected confounding factors. QFT- QuantiFERON-TB Gold test.

This manuscript has been submitted for publication

Conclusions

Artificial intelligence can lead to reliable algorithms to distinguish breath patterns from groups of patients versus controls for a wide set of illnesses, as well as identify potential biomarkers.

Still so much work to be done...



… the best results are yet to come …

AlgoReach
Analyzing Complex Systems

UNIVERSIDAD COMPLUTENSE MADRID

Thank you very much

Artificial Intelligence in the early diagnosis of diseases through Breath analysis

Dr. **José S. Torrecilla,** MSc, MBA
Department of Chemical Engineering
Complutense University of Madrid
Madrid, Spain
jstorre@ucm.es