Machine Learning Based Urinary pH Sensing Using Polyaniline Deposited Paper Device and Integration of Smart Web App Interface : Theory to Application

Souvik Biswas[1], Arijit Pal[1], Pratip Chakraborty[2], Koel Chaudhury[1], and Soumen Das[1*]

[1]School of Medical Science and Technology, Indian Institute of Technology, Kharagpur, WB, 721302, INDIA.

[2]Division of Assisted Reproduction, Department of Infertility, Institute of Reproductive Medicine, Kolkata, WB, INDIA

*Corresponding Author:
Prof. Soumen Das
Email: sou@smst.iitkgp.ac.in

Abstract— The present study investigates the electron transport properties of polyaniline after exposure to acidic and alkaline pH employing density functional theory-based first principle calculation. In-situ deposited polyaniline based paper device maintains emeraldine salt form while it is exposed to acidic pH and converts to emeraldine base when it is subjected to alkaline pH solutions. These structural changes at acidic and alkaline pH are validated experimentally by Raman spectra. Furthermore, the computed Raman spectra from density functional theory are also validated with the experimental one. The changes in the theoretical energy band gap of polyaniline obtained from first principle calculations were correlated with the changes in the experimental impedimetric response of the sensor after exposure to acidic and alkaline solutions. Finally, the impedimetric responses were used to predict urine pH through a machine learning-based smart and interactive web application. Different machine learning-based regression models are implemented to acquire best possible outcome. Gradient Boosting Regressor with least square loss model was selected as it showed lowest mean square, mean absolute, and root mean square error than other models. The smart sensing platform successfully predicts the unknown pH of urine samples with an average accuracy of more than 98%. The locally deployed smart web app can be accessed within a local area network by the end-user which holds promise towards effective detection of urinary pH.

1.   Introduction

Polyaniline is one of the widely used conducting polymers for different sensing applications as the electrical properties of polyaniline (PANI) and its derivatives can be easily manipulated according to the need of specific application. PANI mainly consists of phenyl rings (C_6H_6) along with nitrogen heteroatoms, allowing it to possess different oxidative states (Epstein and Macdiarmid, 1991). Lower ionization potential and/or large electron affinity of PANI produce higher conductivity (Bredas and Street, 1985).

The pH of human urine generally lies within the range of 5-8 (Martín-Gutiérrez et al., 2016). There are two standard methods for measuring urine pH; one involves a glass pH electrode with a pH meter, and the other includes paper strip-based colorimetric measurement. However, these methods require a bulk amount of biofluids and involve complex chemistry. Moreover, the colorimetric paper-based dipstick method lacks the precision required to measure pH in the desired biospecimen. Thus, the need to develop a low-cost, disposable, and robust paper-based impedimetric sensor for reliable and accurate quantification of pH is well realized.

A research group from the Federal University of Goiás proposed a novel smartphone based pH dection method to detect pH values within the range of 1-12, which does not require any machine learning based models (Nogueira et al., 2017). However, the obtained pH values of vinegar acid samples lack sensitivity, as the standard deviation of the obtained results through smartphone-based detection is relatively higher than the standard method. Several research groups have studied different methods for the detection of pH through machine learning based classification model. Most of the studies are mainly focused on classification models and do not provide a suitable predictive regression framework. Availability of a robust regression model will help in quantitative estimation of pH from different analytes, which is far more superior than machine learning

classification based semi quantitative method. Regression models are mainly required to predict pH as they correctly estimate the pH values from a set of trained data sets. Whereas, in the case of the classification model, it can only classify selected pH class already defined in the trained model. A comparative study of different low cost paper based sensor and machine learning based pH detection methods are tabulated in Table I.

The present study proposes a proof of concept for the development of a smart paper-based impedimetric urine pH sensing platform incorporating a web app interface. The selected range for urine pH measurement is 5-8 as it is clinically significant ("Laboratory Assessment of Kidney Disease," 2011; Martín-Gutiérrez et al., 2016). Although, several research groups have used density functional theory (DFT) based calculations to highlight the electron transport properties in different structures of PANI (Cavazzoni et al., 2006, 2004; Reis et al., 2017), effect of pH on material and electronic properties of PANI has not been extensively investigated. The novelty of the present study lies within the fact that it explains the underlying science of electron transport through first principle DFT calculation and further correlates the theoretical calculations with the experimental Raman and Impedimetric analysis, which bridges the gap between experimental validation of first principle calculation. Furthermore, a hyperparameter optimized robust gradient boosting regressor framework is applied to predict urine pH from unknown samples through an interactive web app interface, which also highlights the uniqueness of this study. A set of three optimizable regression algorithms (i.e. optimizable tree, support vector machine, and ensemble regressor) is used and gradient boosting ensemble regressor model is chosen as it shows best suitable performance metrics. The developed web app is deployed within a local area network and can be accessed through connected devices. This proof of concept will enhance the idea of smart sensing within a primary healthcare facility. The hyperparameter optimized regression algorithm

also helps to enhance the resolution of the developed sensing platform and produces more accurate results compared to the pH values obtained manually from the sensors.

The article is arranged as follows. Section II explains the materials and methods related to experimentation. Section III includes the results obtained from the first principle calculation and experimental validation of theoretical calculations by impedance and Raman analysis. Section IV elucidates the development of implemented regression framework for smart sensing. Section V provides conclusion of the study.

Table I.
Comparison of low-cost paper sensors and machine learning based pH detection methods.

Material used	Method	Substrate used	pH Range	Solution volume required	Response time (s)	Accuracy /Error	Reference
Iridium Oxide deposited conductive fibre	Potentiometric	Textile fabric	4-8	Bulk	-	4% error	(Zamora et al., 2018)
Ion-sensitive field-effect transistor	Semi quantification through classification model and quantification through regression model both implemented	-	5-9	Not mentioned	125ms	Prediction error <1.5%	(Hsu et al., 2019)
H_2O_2 test for pH detection	Smartphone based machine learning classification from colorimetric results	Paper strip	2–9	Bulk	1-5 Sec	90 % classification accuracy	(Solmaz et al., 2018)
Fluorescent sensor array	pH classification through machine learning classifier model implemented with fluorescent-images	Printed paper	1-12	0.7 µl/each spot	-	93.36% classification accuracy	(Kim et al., 2021)
Myrciaria califlora peel extract as natural pH indicator	Smartphone based detection	Wax printed paper microzones	1-12	20 µl	300 Sec	95% accuracy	(Nogueira et al., 2017)
In situ deposited polyaniline on the paper substrate	Impedimetric	Chromatography Paper	5-8	2µl	~3 Sec, ~37 Sec	95.53% accuracy	Our previous study (Biswas et al., 2020)
Polyaniline deposited paper substrate	Impedimetric analysis with machine learning based regression model	Chromatography Paper	5-8	2µl	~3 Sec, ~37 Sec	Manual = 96.35 %; machine learning based predictive regression model = 98.41 %	Present study

2.   Materials and Methods

2.1. Chemicals and Reagents

Aniline extrapure AR, 99.5% (Product ID. 22285), and ammonium persulfate (APS) AR, ACS, ExiPlus, 99.5% (APS, Product ID. 29534) were bought from Sisco Research Laboratories Pvt. Ltd., Mumbai, India. Chromatography paper (3MM CHR, CAT No. 3030-614) was purchased from GE Healthcare UK Limited, Buckinghamshire, UK. The developed sensor was tested with Sorensen's buffer. The buffer was prepared with di-sodium hydrogen phosphate anhydrous (Na_2HPO_4) and potassium dihydrogen phosphate (KH_2PO_4) both procured from Merck Life Sciences, India. Sorensen's buffer mimics extracellular fluid characteristics and also immune to drastic changes in pH with varying temperatures. Thus, Sorensen's buffer was used to test the developed sensor at varying pH. Hydrochloric acid (HCl, procured from Merck Lifesciences, India) was used to prepare PANI. All the solutions for experimental work was prepared with Milli-$Q^®$ water.

2.2. Sensor Device Fabrication

The sensor device was fabricated, as reported in the previous study (Biswas et al., 2020). At first, pieces of chromatography paper with a dimension of 10 mm $\times$ 5 mm were prepared. These pieces were then submerged into a 2ml centrifuge tube containing a solution of 0.1M aniline (prepared in 1M HCl) and 0.1M aqueous solution of APS in a 1:1 molar ratio. The submerged paper was then incubated at 0°C for two hours. In-situ polymerization of PANI over paper substrate was achieved by chemical oxidative method where APS acts as an oxidizing agent with HCl as dopant medium (Stejskal and Gilbert, 2002). After successful polymerization, the PANI modified paper substrate was thoroughly cleaned with de-ionized water and air-dried overnight. Two silver electrodes were drawn with a gap of 9mm using a stencil over the air-dried PANI modified paper substrate. A schematic representation of sensor fabrication steps is depicted in

supplementary Fig. S1.

2.3. Computational Details

The ground-state DFT calculations of protonated polyaniline emeraldine salt (ES) and emeraldine base (EB) were carried out using Quantum-ESPRESSO package (Giannozzi et al., 2009). The structures of the different PANI states were prepared and visualized using VESTA (Momma and Izumi, 2011). All computations were carried out with Generalized Gradient Approximation (GGA) of Perdew-Burke-Ernzerhof (PBE) for exchange-correlation energy (Perdew et al., 1996). The core electron characteristics were expressed using Ultrasoft pseudopotentials generated with Rappe Rabe Kaxiras Joannopoulos (RRKJ) method (Rappe et al., 1990). Plane-wave basis function were considered for all the calculations having kinetic energy cut off of 60 Ry. k-space sampling of 8×8×8 has been implemented for construction of Monkhorst–Pack net with Tetrahedron method (Blöchl et al., 1994) for integration of Brillouin zone. Calculation of molecular orbital and Raman tensors was carried out at the gamma point of k-space grid.

2.4. Material and Device Characterization

In-situ deposition of PANI over paper substrate was confirmed by scanning electron microscope (FESEM, Merlin, ZEISS Corp., Germany). Raman spectroscopy (Alpha 300R, WITec GmbH, Germany) was used to validate the structural changes in PANI after exposure to varying pH analyte.

2.5. Impedimetric Analysis

Impedance spectroscopy of the sensor device at varying pH was obtained through Agilent 4294a Impedance Analyzer by applying a sinusoidal excitation voltage of 100mV in the frequency range of 40Hz-110MHz. The fabricated sensor was exposed to a 2µl analyte droplet over the sensing

region between the two electrodes. It was ensured that the droplet did not interact with the electrodes and only soaked into the PANI-modified paper substrate. Three consecutive measurements were taken to achieve reliable and reproducible results.

2.6. Regression Framework for Integration of Smart Web App Interface

A set of three optimizable regression models i.e., tree, support vector machine (SVM), and ensemble based regressor models were compared to select best regression framework according to their performance metrics. Initially, a set of 39 (13 different pH values each measured with 3 separate devices) impedimetric responses of the sensor following exposure to varying pH buffer solution was considered. The same data set was also used to tune the hyperparameters and train the optimized model. The best performing regression framework is used to predict the pH values from impedance data for accurate and reliable estimation of urine pH. The selected optimized regression model was then tested with a python-based web app interface to predict the pH values of 11 urine samples. The web app was deployed over the local area network and accessible to all
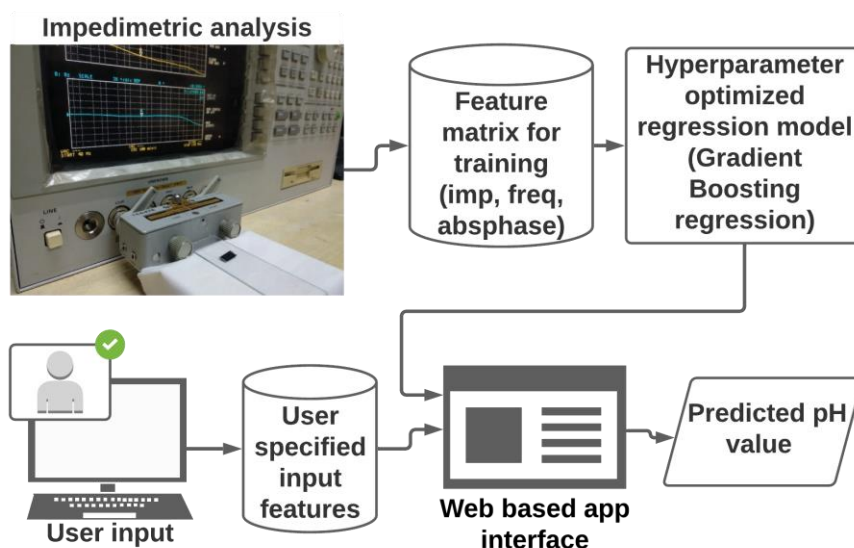

Fig. 1. Schematic representation of implemented machine learning based gradient boosting regressor model.

connected devices through an assigned IP address. The user provides the input as impedance and phase value obtained for a urine sample of unknown pH at 1 kHz frequency, and the smart web app predicts the pH value accurately. In the subsequent section, algorithm for tuning the hyperparameters and development of python-based smart web app has been thoroughly discussed. A simple schematic representation of the regression framework to predict pH values of urine samples through an interactive web app interface is depicted in Fig. 1.

3. Results and Discussion

3.1. Details of Optimized Geometry

Bipolaronic form of polyaniline was considered to mimic the protonation state of emeraldine salt (ES). At lower acidic pH the bipolaronic activity increases and with higher acidic pH it gradually decreases (Biswas et al., 2020). In PANI, Franck-Codon like vertical ionization process takes place, which requires energy of E_{IPV} (supplementary Fig. S2). The energy required for geometry relaxation in an ionized state is E_{rel} (supplementary Fig. S2). To further achieve a relaxed ionization state from the ground state, the geometry of the ground state must be distorted; this requires energy of E_{dis} (supplementary Fig. S2). In the oxidation process, an electron is removed from the polymeric chain of PANI; this causes a reduction in ionization energy by $\Delta\varepsilon$. In case of PANI-ES, $\Delta\varepsilon$ is always greater than E_{dis}; as a result, localized lattice distortion within the polymeric chain takes place, which promotes polaron formation (Bredas and Street, 1985). Furthermore, removal of another electron from PANI polymeric chain generates a pair of equivalent charges with strong local lattice deformation, termed as bipolaron (Bredas and Street, 1985). The distortion energy (E_{dis}) required to form the bipolaron is nearly similar to that of two polaron formation. Furthermore, reduction in ionization energy or relaxation energy (E_{rel}) is much greater for bipolaron formation than for two polaron formation. Due to this phenomenon, bipolaron
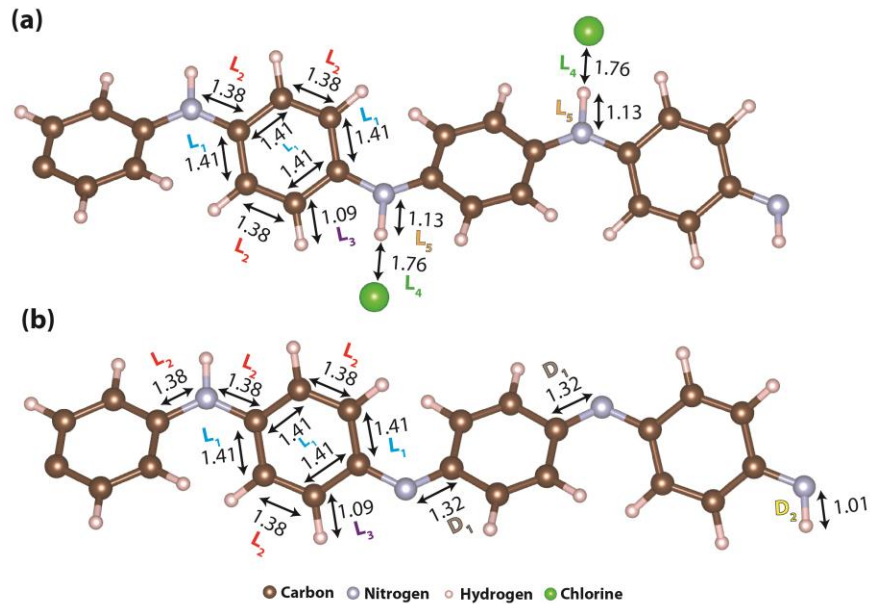
Fig. 2. Optimized structure for PANI-ES and PANI-EB.

is thermodynamically more stable than two polarons, despite the presence of Columbic repulsion between two equivalent charges (Brédas et al., 1984; Bredas and Street, 1985; Cavazzoni et al., 2006). Thus, bipolaronic form of PANI-ES has been considered for further analysis. On the other hand, PANI-EB state is considered after exposure to basic pH. Higher the alkalinity of the analyte solution, conversion of ES to EB increases significantly. Thus, EB state is considered at basic pH to mimic the deprotonation state of PANI. Finally, the geometry relaxation is considered for PANI-ES (bipolaronic form) and PANI-EB structure to mimic the exposed sensor device to acidic and basic pH, respectively.

Most of the experimental studies related to HCl doped PANI in aqueous solution suggested a Triclinic P structure of PANI (Cavazzoni et al., 2004; Pouget et al., 1991; Winokur and Mattes, 1998). Thus, a Triclinic P unit cell was considered for geometry relaxation. The optimized structure of PANI ES and EB are depicted in Fig. 2 (a) and (b), respectively. The relaxed unit cell of PANI-ES consists of a repeating aniline unit doped with two chlorine atoms along the z-axis

having optimized sides of a = b = 12.55 Å. After geometry relaxation, cell length along the direction of PANI-axis was found to be 20.56 Å. The optimized angles between the three axes were, α= 87.47°, β= 89.41°, and γ= 91.04°. Fig. 2(a) shows different optimized bond lengths for PANI-ES structure. It is evident from the results that distance between the carbon atom bound to nitrogen and neighboring carbon atom in the same ring (L_1 in Fig. 2 (a)) was 1.41 Å. On the other hand, optimized bond distance between other carbon atoms of the same ring, as well as nitrogen and neighboring carbon atom (L_2 in Fig. 2 (a)) was 1.38 Å. The measured bond length between chlorine-hydrogen atom (L_4 in Fig. 2(a)) and hydrogen-nitrogen atom (L_5 Fig. 2(a)) was 1.76 Å and 1.13 Å, respectively. Also, the optimized bond distance between carbon and hydrogen atom (L_3 in Fig. 2(a)) was 1.09 Å.

The relaxed unit cell of PANI-EB consists of repeating aniline units with alternating amine and imine nitrogen (Biswas et al., 2020). The optimized coordinates for the unit cell were, a = 12.63 Å, b = 12.30 Å, c = 20.33 Å and angles between the three axes were α= 89.26°, β= 90.51°, γ= 90.47°, respectively. The calculated values of L_1, L_2, and L_3 (Fig. 2 (b)) were found to be 1.41, 1.38, and 1.09 Å, respectively. The optimized bond distance between the imine nitrogen and the neighbouring carbon atom (D_1 in Fig. 2 (b)) was 1.32 Å. Also, the bond length of amine nitrogen and hydrogen (D_2 in Fig. 2 (b)) was observed to be 1.01 Å.

3.2. Validation of Optimized Geometry with Raman analysis

Experimental Raman spectra of PANI after exposure to acidic (i.e., pH 5) and alkaline pH (i.e., pH 8) buffers were obtained to compare it with the computed theoretical spectra. Alpha 300R Raman Spectroscopy (WITec GmbH, Germany) was used to measure the Raman spectra using a laser source of 633 nm. The major assigned Raman bands for PANI ES and EB are tabulated in Table S1 and Table S2 of supplementary information, respectively. Furthermore, Raman tensors

of optimized PANI-ES and PANI-EB were computed at the gamma point of *k*-space grid. Experimental and computed Raman spectra are represented in Fig. 3. The in-situ deposition of PANI over cellulose fibers of the paper substrate is also visualized by FESEM depicted in inset of the supplementary Fig. S1.

A visible Raman band at ~1383 cm^{-1} confirms the presence of polarons with small conjugation length in PANI matrix following exposure to acidic pH (Engert et al., 1994) (Fig. 3(a)). A stretching of C─N is visible at ~1404 cm^{-1}. Raman band for C─C stretching and C─H bending of quinoid ring is also observed at ~1488 cm^{-1}. Finally, a high-intensity peak at ~1514 cm^{-1} corresponding to N─H bending confirms increased of bipolaronic activity in PANI structure after exposure to acidic pH (i.e., PANI-ES) (Louarn et al., 1996). The computed spectra (Fig. 3(b)) show similar band assignment for polaron, C─N stretching, C─C stretching along with C─H bending in quinoid ring, and N─H bending at ~1369, ~1398, ~1487, and ~1503 cm$^{-1,}$ respectively. A peak at ~1476 cm^{-1} in computed Raman spectra is visible due to the presence of alkaline group at para
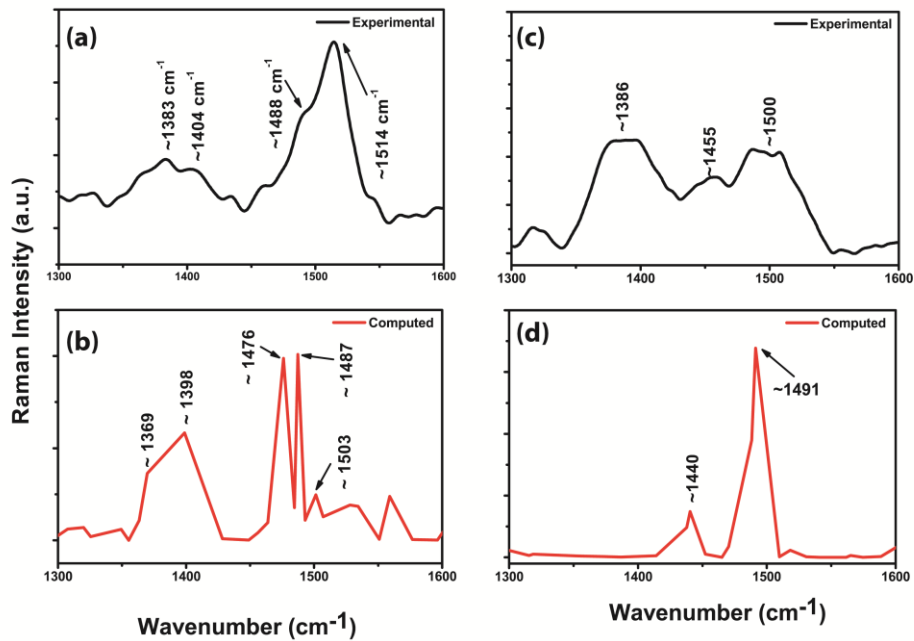

Fig. 3. Comparison of experimental Raman spectra with computed spectra. (a)-(b) PANI-ES and (c)-(d) PANI-EB Raman spectra.

position in PANI-ES structure. A comparative study between experimental and computed Raman bands after exposure to acidic pH (ES structure) is provided in supplementary Table S1.

Raman spectra of PANI after exposure to alkaline pH buffer show a peak at ~1386 cm^{-1} due to the presence of a few polarons. A peak at ~1500 cm^{-1} with medium intensity is observed, which corresponds to N─H bending of amine nitrogen in EB structure of PANI (Fig. 3(c)). C=N stretching of quinoid ring in EB structure is also visible at ~1455 cm^{-1} (Quillard et al., 1994). The computed spectra for PANI-EB obtained from DFT study demonstrate similar peaks at ~1440 and ~1491 cm^{-1} which corresponds to C=N stretching of quinoid rings and N─H bending, respectively. The peak at ~1386 cm^{-1} corresponding to polarons is not visible in the computed spectra (Fig. 3(d)) as the DFT study considers only fully deprotonated state of PANI. In a practical scenario, it is impossible to completely deprotonate the PANI matrix within the sensing region. The comparison between experimental and computed Raman peaks for after exposure to alkaline pH (EB structure) is tabulated in supplementary Table S2.

3.3. Energy Band Diagram and Density of States Analysis

Periodic PANI unit doped with two chlorine atoms forms the PANI-ES structure, whereas a unit of PANI with benzenoid amines (─NH─) and quinoid imines (=N─) forms the EB structure. The corresponding band structure for PANI-ES and PANI-EB in the equilibrium state is depicted in Fig. 4 and Fig. 5, respectively. The optimized PANI ES and EB geometry from relaxation step is used to calculate the band structure. A closed k-path (with 8×8×8 Monkhorst-pack grid) of Γ→X→Y→Z→Γ in first Brillouin zone was considered for band structure calculation. Corresponding density of states (DOS) for both the structures were also calculated in the same k-space grid and are shown in Fig. 4 & 5. The energy difference between the lowest unoccupied molecular orbital or conduction band and highest occupied molecular orbital or valence band
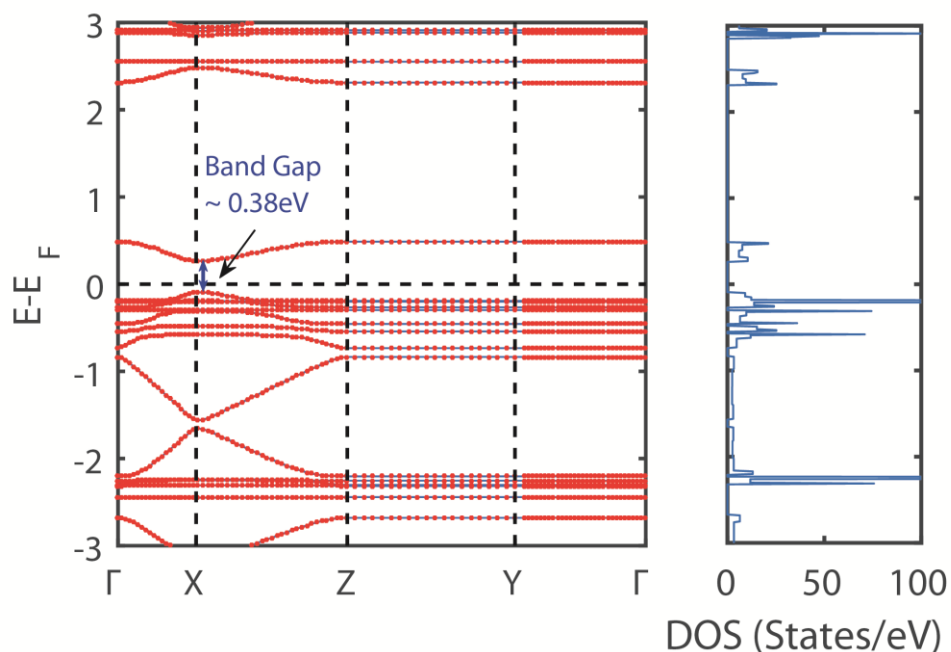
Fig. 4. Band and density of states plot for PANI-ES structure.

determines the band gap of the system. The calculated band gap for PANI-ES structure was found to be approximately 0.38 eV, whereas the band gap increased to 0.88 eV for EB structure. The obtained band gap values are similar to the previous DFT based studies (Reis et al., 2017; Wang et al., 2012). However, the experimental values of band gap are quite high (around 1.3eV for ES and 2 eV for EB) (Stafström et al., 1987). The selection of exchange-correlation functions (PBE for the present study) in DFT calculation leads to the underestimation of computed band gap. Though more realistic values can be obtained by considering hybrid functionals and many-body corrections, it is computationally expensive. Although use of PBE functional undervalues the band gap, it is a robust and computationally efficient method for quantitively exploring the fundamentals of electron transport properties in PANI ES and EB (Martin, 2004).

The change of band gap in ES and EB form can be correlated with the structural changes of PANI with varying pH. Initially, at acidic pH, the activity of bipolaron formation is much higher and facilitates higher flow of electrons through PANI matrix connected with two silver electrodes.

The higher conductivity of the system is due to the presence of doped chlorine ions in the PANI-ES structure. Thus, a narrow band gap is visible in the theoretical study owing to higher conductivity of PANI-ES at lower acidic pH. On the other hand, at basic pH, deprotonation of PANI takes place, and ES state converts to EB consisting only of amine and imine nitrogen. Thus, absence of chlorine ions in the structures broadens the band gap and reduces conductivity of the sensing device, which reduces the flow of electrons through the PANI matrix. The electronic DOS study can also correlate the changes in conductivity of PANI at ES and EB states. It is evident from the DOS plots in Fig. 4 & 5 that a higher number of electrons are present in the valence band of ES than EB. Thus, in case of ES, a greater number of electrons can jump from the valence band to the conduction band with a lower amount of energy than EB. This phenomenon also validates the rise in conductivity in PANI matrix exposed to acidic pH, which promotes a higher bipolaronic activity in PANI-ES state.

The charge densities of the valence band (bonding π state) and conduction band (antibonding
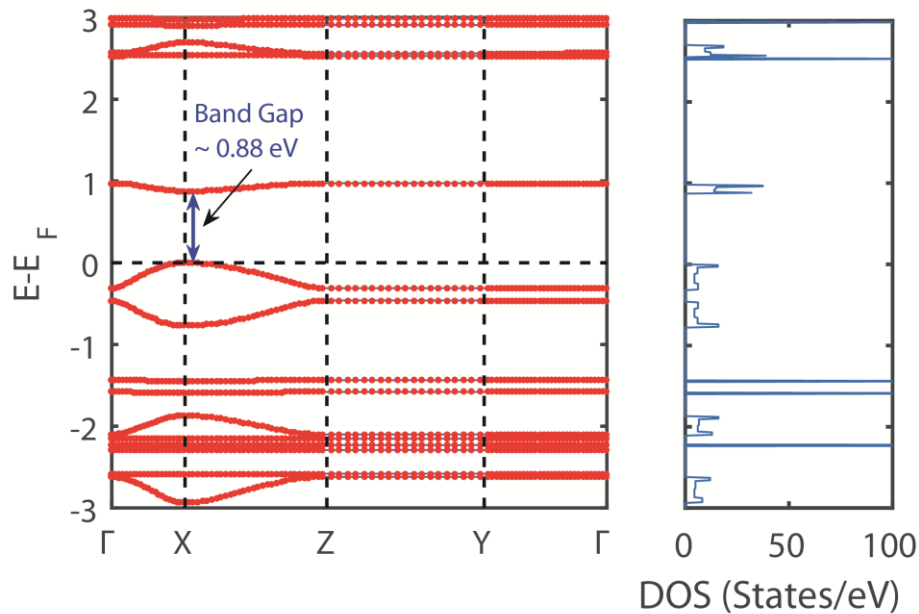

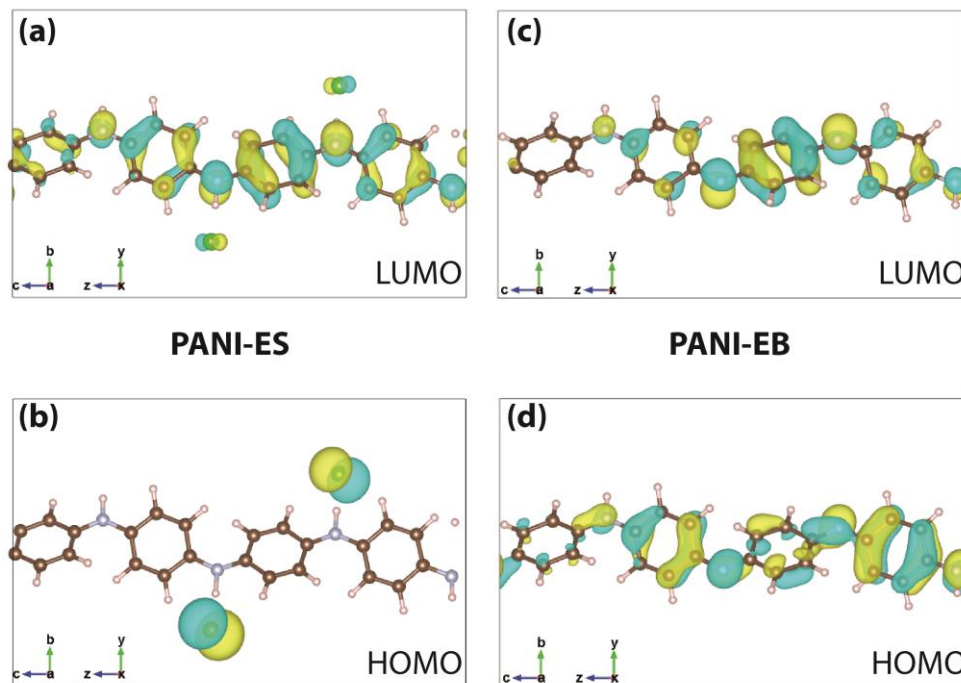Fig. 5. Band and density of states plot for PANI-EB structure.

Fig. 6. Molecular orbitals of PANI-ES and PANI-EB structure. (a) & (c) represent LUMO (conduction band) and HOMO (valence band) of PANI-ES and (c) & (d) represents LUMO (conduction band) and HOMO (valence band) of PANI-EB structure. (Here yellow and cyan marked regions depict depletion (positive charge) and accumulation (negative charge) of electrons respectively. All figures have been generated using VESTA visualization package having an iso value of 0.0004.)

π^* state) for PANI ES and EB are depicted in Fig. 6. The conduction charge densities of PANI-ES are localized around the benzenoid rings (Fig. 6(a)), whereas the valence charge density is distributed across chlorine ions (Fig. 6(b)). The protonation of PANI introduces chlorine ions into the structure, which contributes excess electrons. The orbital shapes of PANI-ES confirm the presence of p_x, p_y and p_z orbital of chlorine. Depending on the H^+Cl^- dipoles and orbital lobes the energy level order of these orbitals is $p_z > p_x > p_y$. As p_y orbital is near to the H^+ ion, electrostatic potential for p_y orbital is lowest, as a result, it has the lowest energy level. Furthermore, p_x and p_z orbitals are far away from H^+ ions, thus they have higher energy levels. Additionally, other H^+ ions might have more effect on p_x orbital which confirms the lower energy level of p_x orbital than p_z orbital. Similarly, for PANI-EB structure, conduction charge densities are delocalized across the

PANI chain with most part surrounding quinoid rings (Fig. 6(c)), whereas the valence charge density of PANI-EB is localized around the quinoid and benzenoid rings (Fig. 6(d)).

3.4. Electron Localization Function and Bader Charge Analysis

Electron localization function (ELF) measures the probability of finding an electron in the surrounding space of a reference electron with same spin situated at a given point. ELF provides the spatial localization of the reference electron and proposes a methodology to map the electron pair probability in a multielectronic system. The ELF was obtained using Quantum ESPRESSO package and visualized with Vesta data analysis software. The ELF was calculated on a 3-dimensional grid space with single determinantal wavefunction obtained from Kohn Sham orbitals φ_i (Silvi and Savin, 1994), and given as,

$$ELF = \frac{1}{1 + \left(\dfrac{D}{D_h}\right)^2} \tag{1}$$

where, $D = \frac{1}{2} \sum_i |\nabla \varphi_i|^2 - \frac{1}{8} \frac{|\nabla \rho|^2}{\rho}$, $D_h = \frac{3}{10}\left(3\pi^2\right)^{5/3} \rho^{5/3}$, and ρ is the spin density.

The ELF value ranges between 0 to 1, where 1 means perfect localization. Fig. 7 depicts the ELF of PANI ES and EB at an ELF value of 0.7, which is suitable for mapping the localized electron
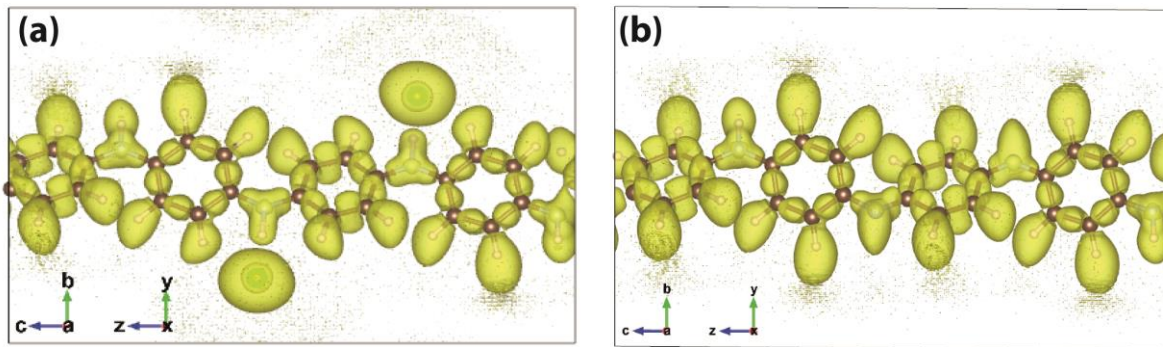

Fig. 7. 3D electron localization function plot of (a) PANI-ES and (b) PANI-EB. Yellow marked region defines the probable region of finding an electron. The scattered yellow dots in PANI-ES validates the presence of dense electron gas in bipolaron form.

pair in the multielectronic system. There are various types of localized domains known as electron basins, located surrounding an attractor in ELF plot (Fig. 7) (Frison and Sevin, 1999). The highly covalent nature of C—C bonds in PANI-ES structure can be confirmed by the presence of electron basins near C—C bonds. In case of C—N, N—H, and C—H bonds, the electron basins become delocalized. Additionally, Bader charge analysis (Tang et al., 2009) reveals an increase of approximately 0.20 |e| (electronic charge) around Cl^- ions compared to localized charge around the neighbouring nitrogen atom. This phenomenon confirms that electrons jump from Cl^- ions to neighboring nitrogen atom at excited state, which clearly correlates with the DOS plot (Fig. 4) where density of electrons near valence band is much higher than conduction band at equilibrium of PANI-ES state. Valence and conduction charge density plot (Fig. 6 (a) & (b)) also reveals a higher localization of electrons near Cl^- ions at valence band than in conduction band owing to the results obtained from Bader charge analysis. The ELF plot of PANI-ES (Fig. 7(a)) validates the fact that a dense electron cloud is present near Cl^- ions and acting as an electron donor to the PANI matrix. In the experimental study, exposure to acidic pH increases the bipolaronic activity within PANI matrix. Due to the bipolaronic activity, hole formation takes place within the PANI polymeric chain. To attain charge equilibrium Cl^- ion acts as a donor and gradually an increase in conductivity is observed.

In impedimetric experiments, conductivity of the device reduces following exposure to alkaline pH, due to deprotonation and conversion of ES to EB state. The deprotonation of PANI-ES removes the donor chlorine ion from the PANI matrix and converts to PANI-EB. Thus, a higher impedimetric response was observed. In case of PANI-EB structure localized electron basins are visible in benzenoid and quinoid rings. The electron basin for lone pair electrons of amine and imine nitrogen is visible in ELF plot (Fig. 7 (b)). Furthermore, Bader charge analysis of PANI-EB

structure shows an average charge increases of 0.21 |e| in nitrogen atom of EB compared to PANI-ES state, which confirms that absence of Cl^- ion impacts the electron flow within the PANI polymeric chain, thus a lower concentrated electron cloud is observed near N atom (Fig. 7(b)). The reduced conduction of electrons takes place mostly within the benzenoid and quinoid rings. This phenomenon can be verified by the conduction and valence band charge densities (Fig. 6 (c) & (d)). The overall charge densities of PANI ES and EB are depicted in supplementary Fig. S3.

3.5. Impedimetric Response at Varying pH Buffers

AC impedance spectroscopy method measures the current $i(t)$ with respect to applied voltage $v(t)$ to evaluate complex impedance of the system and is expressed as,

$$Z(j\omega) = \frac{F\{v(t)\}}{F\{i(t)\}} \tag{2}$$

where, $F\{\}$ denotes Fourier transform operator, and ω is the angular frequency. The applied AC voltage and current can be expressed as $v(t) = V_m \sin(\omega t)$ and $i(t) = I_m \sin(\omega t + \theta)$, respectively. Furthermore, the complex impedance $Z(j\omega)$ can be written in terms of real (Z') and imaginary (Z'') impedance as,

$$Z(j\omega) = Z' + jZ'' \tag{3}$$

where, $j = \sqrt{-1}$, $Z' = |Z|\cos(\theta)$, $Z'' = |Z|\sin(\theta)$, $\theta = \tan^{-1}\left(\frac{Z''}{Z'}\right)$ and $|Z| = \sqrt{(Z')^2 + (Z'')^2}$.

Impedimetric responses for the developed sensor was obtained at varying pH buffer solution with 100mV of AC excitation voltage in the frequency range of 40 Hz-110 MHz. 2µl of target pH buffer was dropped between the two electrodes, and impedimetric responses were obtained. A schematic representation of impedance measurement is shown in Fig. 8 (a). The observed Bode impedance and phase plots are depicted in Fig. 8(b) and (c), respectively. The impedance values are chosen at 1kHz frequency to compare the results with previously reported values (Biswas et

al., 2020). The obtained impedimetric values at 1kHz frequency were found to be 6151.28±94.51 Ω, 8053.62±36.46 Ω, 13197.76±217.95 Ω, and 20245.47±179.91 Ω (impedance ± standard deviation) for pH 5, 6, 7 and 8 respectively, which lies within the statistically significant range as reported earlier. This confirms the fact that with increasing pH conductivity of the PANI matrix decreases. At lower pH, the bipolaronic activity in PANI-ES enhances the conductivity of PANI matrix within the sensor, and with increasing pH (alkaline region), this bipolaronic activity decreases. As a result, the PANI-ES form gets deprotonated, and transitions into PANI-EB form, as shown in Fig. S4 of supplementary information. The computed band gap values also suggested a rise of 0.5 eV in PANI-EB than in PANI-ES. Theoretical energy band gap of PANI-ES is found to be 0.38 eV (Fig. 4.), whereas 0.88 eV (Fig. 5) gap is prominent in PANI-EB structure. The rise
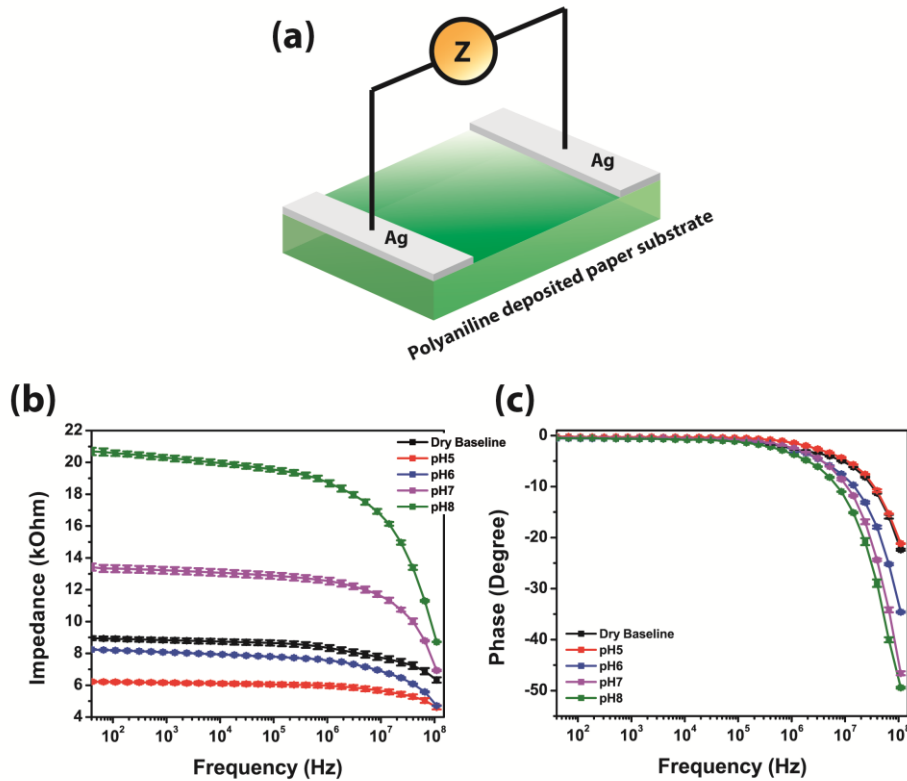

Fig. 8. (a) Schematic representation of device under test. Bode (b) Impedance and (c) phase plot of the developed sensor at varying pH analyte.

in energy band gap in case of EB structure can be correlated with the higher impedimetric values when the sensor device is subjected to alkaline pH . The device gets less conductive due to the removal of Cl⁻ ion and hence a rise in band gap is visible. The Bader charge distribution and ELF study also validates the electron transport properties in PANI-ES and EB structure which is the main sensing element of the device at acidic and alkaline pH, respectively. Supplementary Fig. S5 depicts the impedimetric responses at varying pH adopted from the previous study (Biswas et al., 2020). These results were also used to train the regression model. Sensitivity and resolution of the sensor have been discussed thoroughly in the earlier reported article (Biswas et al., 2020). It is important to mention that the sensor was unable to detect pH below the difference of 0.13. Implementation of regression framework overcomes the device limitations efficiently, which is discussed in the subsequent section.

3.6. Real Sample Analysis

The developed sensor was tested with 11 urine samples. Analysis of urine samples was performed in accordance with the ethical standards of the institutional research committee. The pH values of those samples were calculated from the obtained impedance at 1kHZ frequency
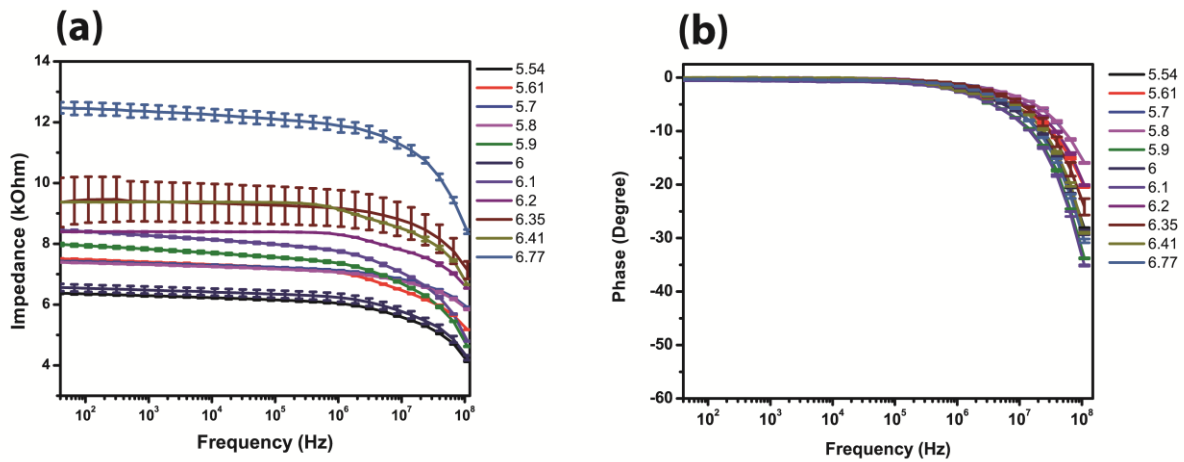

Fig. 9. Bode (a) impedance and (b) phase plot of the sensor after exposure to real urine samples.

TABLE II

Manually calculated pH from calibration curve compared with glass pH electrode reading

Sample no.	pH of urine sample measured by glass electrode	Calculated pH from calibration curve at 1kHz frequency	Accuracy for manual calculation (%)	Average accuracy for manually calculated pH (with standard deviation) (%)
1	6.35	6.39	100.62	
2	5.61	5.44	96.96	
3	6.41	6.34	98.90	
4	5.9	5.61	95.08	
5	6.2	6.03	97.25	
6	5.80	5.61	96.72	96.35 ± 4.72
7	6.1	6.02	98.68	
8	6	5.1	85	
9	5.54	5.03	90.79	
10	6.77	6.87	101.47	
11	5.70	5.61	98.42	

manually using the previously reported calibration curve (supplementary Fig. S5 (c)) (Biswas et al., 2020). The impedance and phase values of the same 11 urine samples were then provided as input in the smart web app, which predicts urine pH through an optimized regression model. The obtained pH values from the sensor and predicted pH values from the web app were compared with measured pH values from the glass pH electrode, which is the gold standard for measuring pH. The results are tabulated in Table II. The manual calculation yields 96.35% average accuracy, whereas the optimized regression model predicts pH values with an average accuracy of 98.41% through the smart web app interface. The impedimetric responses for 11 urine samples are provided in Fig. 9 (a) and (b).

4. Regression Framework for Smart pH Sensing

Impedimetric responses of the sensor at varying pH display a gradual decrease in conductivity with increasing pH value. Furthermore, the first principle calculations and experimental results unveil the electron transport properties of the system at acidic and alkaline pH. Finally, the

impedimetric values for varying pH buffers (from pH 5 to 8) have been considered for preparing a robust regression model for the accurate quantitative detection of unknown pH values in urine samples.

Three set of optimizable regression algorithms are compared to select the best suitable regression model. Table III. tabulates the performance metrics of optimizable tree, support vector machine (SVM) and ensemble regressor.

The parameters for optimizing the ensemble model are ensemble method (*EnM* : Bagging or Least Square Boost (i.e. Gradient Boosting Regressor with least square loss)), leaf size (*Leaf$_{min}$*), number of learners (*Learn$_{min}$*), and learning rate (*R*). Similarly, the only optimizable parameters for Tree regression algorithm is *Leaf$_{min}$* and for optimizable SVM, Box Constraints (*BoxC*), epsilon (*ε*) and kernel (*K*) are available as optimizable hyperparameters. Bayesian Search Optimization is considered to predict the global mean square error (MSE) with optimized ensemble parameters.

Bayesian optimizer was used for all regression models to optimize the selected hyperparameters. The Bayesian optimization attempts to achieve minimum parameters for an objective function *f(x)* within a bounded domain. To perform the optimization, a Gaussian process model of *f(x)* is considered. Then a Bayesian update procedure is used to modify the Gaussian process model at each newly evaluated *f(x)*. Depending on the updated *f(x),* an acquisition function *a(x)* is chosen to maximize the determination of the next point *x* for a new calculation. Bayes optimization includes six different acquisition functions, i.e., *expected-improvement-per-second-plus*; *expected-improvement*; *expected-improvement-plus*; *expected-improvement-per-second*; *lower-confidence-bound*; and *probability-of-improvement*. In this study *expected-improvement-per-second-plus* acquisition function is chosen as it mostly outperforms other functions (Bull, 2011; Snoek et al., 2012). This acquisition function consists of three independent functions. The first

function, *expected-improvement* function, calculates the expected amount of improvement in the objective function by defining x_{best} as the lowest posterior mean, $\mu_J(x_{best})$ as the lowest value of the posterior mean and *J* as a posterior distribution function. Now, the expected improvement (*EI*) is given as,

$$EI(x, J) = E_J[\max(0, \mu_J(x_{best}) - f(x))] \qquad (4)$$

In some instances, time required to calculate the objective function can depend on a particular region. Thus, *per-second* function implements time-weighting in acquisition function, which helps in increased algorithm performance. While performing the Bayesian optimization, it considers another Bayesian model of objective function to calculate evaluation time as a function of position *x*. The expected improvement per second (*EIpS*) is defined as,

$$EIpS(x) = \frac{EI_J(x)}{\mu_g(x)} \qquad (5)$$

where, $\mu_g(x)$ represents posterior mean of the timing of Gaussian process model.

A *plus* function is used in the acquisition function to escape the local minima of the objective function if it starts to overexploit a region. Now overexploitation is defined by,

$$\sigma_J^2(x) = \sigma_H^2(x) + \sigma^2 \qquad (6)$$

where, $\sigma_H(x)$ is the standard deviation of the posterior objective function at *x* and σ is the standard deviation of the added noise. Bayesian optimization at each instance checks whether exploration ratio (d_σ) is greater than $\sigma_H(x)$, and if this holds true, the algorithm confirms that *x* is overexploiting. Thus, it attempts to generate an updated point *x*, that is immune to overexploiting. After performing the search, the algorithm's goal is to find minimum optimized mean square error (MSE*)* values.

During the search step, Bayesian optimizer does not choose the minimum MSE value, rather it depends on the upper confidence interval of the searched MSE points, which is termed as 'bestpoint'. To find out the lowest possible MSE, a new set of criteria is updated in the algorithm to identify minimum observed MSE point, which is termed as 'minimum error point'. Although, minimum error point gives the lowest possible value of MSE, it may not produce the most effective optimized parameters. Thus, the best point parameters are chosen for further calculations.

Initially, 39 impedimetric responses were considered to train the regression models. For the

---

Pseudo Code 1: Tuning of optimized hyperparameters for ensemble regression model

Input: x : Impedance data set; y : output class.
Output: *EnM* : Optimized ensemble method;
         Leaf$_{min}$: Optimized minimum leaf size;
         Learn$_{min}$: Optimized minimum number of learner;
         R : Learning rate;
         MSE : Optimized mean square error

Initialisation

1: Define parameters :
         params [*EnM*] ← (*Bag* or *LSboost*);
         params [*Leaf$_{min}$*] ← (1, *max*(2,*floor*(n/2)));
                 where n=number of observations;
                 floor() rounds off each element of
                 the selected input
         params [*Learn$_{min}$*] ← (10, 500);
         params [*R*] ← (0.001, 1);

2: Define evaluation:
         cv = kfold (*n_split*; *n_repeat*; *random_state*)
         Where, *n_split*: number of folds (>2);
                 n_repeats: number of times cross-
                       validator repetition;
                 random_state: generation of random states
                       (generally, equals to 1).
3: Define search:
         while *search =BayesSearch()* **do**
            Estimator ← *ensembleRegressor*();
            Search Space ← *params*;
            Cross Validation ← *cv*;
         end while
4: Perform the search:
         search.fit (*x*, *y*)
5: Calculate minimum *MSE* considering minimum observed objective function along with minimizing upper confidence interval of the optimized *MSE* among the searched points.
6: Repeat step 1 to 5 until bestpoint and minimum error *MSE* achieved.
7: **return** Output

present study, the resolution of the sensor is around 0.13 pH, thus, 39 datasets can be acquired

within the range of 5 to 8 pH, beyond this, the impedimetric responses will produce false-positive

outcomes. To overcome this problem, hyperparameters for the ensemble regressor model is

optimized so that any values within the range of pH 5 to 8 can be predicted accurately. The feature

matrix consists of 1170×3 data points where the first column represents frequency, second column

provides the values of absolute impedance and the last column contains absolute phase. Each

impedimetric response comprises of 30 data points, hence the number of rows of the feature matrix

becomes 39×30=1170.

MATLAB 2020b has been used to achieve optimized hyper parameters having a 10-fold cross

validation with repeat iteration of 3. Impedance, frequency and absolute phase values of the

obtained sensor response have been considered as features of the input (x in the Pseudo Code 1),

and corresponding pH value has been considered as the predicted class (y in Pseudo Code 1).

Supplementary Fig. S5 shows the impedimetric responses which have been considered for the

input data. Based on the performance metrics tabulated in Table III. ensemble-based regression

model out performs other regression algorithms with highest R-squared and lowest root mean

square error (RMSE) value. The ensemble-based regression model provides excellent results in

Table III. Performance metrics for different regression models

Regression Model	Optimized parameters	Metrics for regression models			
		RMSE	MAE	Best point MSE	R-squared
Optimizable Tree	Minimum Leaf Size = 4	0.1151	0.050678	0.011636	0.98
Optimizable SVM	Box constraint = 0.00221; Epsilon: 0.70918; optimized Kernel Function = Linear	0.39758	0.35206	0.0060155	0.82
Optimizable ensemble model	Number of Learners = 125, minimum leaf size = 1; optimized ensemble method = Bag	**0.07622**	**0.039763**	**0.0053131**	**0.99**

predicting unknown urine pH compared to other regression algorithms due to its better performance and reliability of handling complex impedimetric data sets (Bria et al., 2020; Dhamu and Prasad, 2020). A simple pseudo code for tuning parameters for ensemble regressor (best regression framework) is depicted in Pseudo Code 1. RMSE and mean absolute error (MAE) are mainly considered as suitable performance metrics for regression problems (Li et al., 2022; Lu et al., 2022; Wang et al., 2022). RMSE and MAE are calculated from the following equation:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(a_i - \hat{a})^2} \tag{7}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|a_i - \hat{a}| \tag{8}$$

Where, N= sample size, a_i = actual value of a at i^{th} iteration and $\hat{a}$ = predicted value.

The achieved optimized *EnM* is Least Square Boost or Gradient boosting regressor with least square loss. Other achieved optimized parameters are 1, 23, and 0.73 for *Leaf*$_{min}$, *Learn*$_{min}$, and *R* respectively. The MSE is found to be 0.0058. Fig. 10(a) depicts the minimum MSE plot for ensemble based regressor model. Light blue points show the estimated minimum MSE by the optimization process at each iteration, and dark blue points indicate the actual observed minimum MSE. The bestpoint hyperparameter and minimum error hyperparameters are represented with a red square and a yellow circle, respectively. The obtained MSE for bestpoint and minimum error hyperparameter are 0.0058 and 0.0049, respectively for optimizable ensemble regression framework. The RMSE and MAE is found to be 0.0762 and 0.039763, respectively at best point for the same framework. The obtained R-squared value is 0.99, which confirms the robustness of the ensemble regression model. Optimized parameters and performance metrics for other regression models are also tabulated in Table III. Fig. 10(b) depicts the predicted vs actual response plot for optimizable ensemble regressor. The diagonal black line shows the perfect response, points

near this line are highly accurate. The distant points from perfect prediction lines are less accurate. Minimum MSE and predicted vs actual plot for other regression models are represented in Fig. S6 of supplementary information.

   These optimized parameters are then used in a python-based code to train the gradient boosting model with least square error. The code uses a package named Streamlit to deploy the regression framework over a local area network through an interactive web app interface (Fig. 11). A set of 11 urine samples were tested with the developed sensor, and impedance spectroscopy measurements were carried out. The measured impedance and absolute phase are then provided as input in the app. The web app produces the predicted pH value by efficiently using the regression framework running in the backend of the web app. As the developed device works most efficiently at 1 KHz frequency (Biswas et al., 2020), the output function for predicting the pH is set to 1KHz frequency by default. Fig. 9 shows the sensor response as Bode impedance and phase plot for 11 different urine samples. The predicted pH values from gradient boosting regression based web
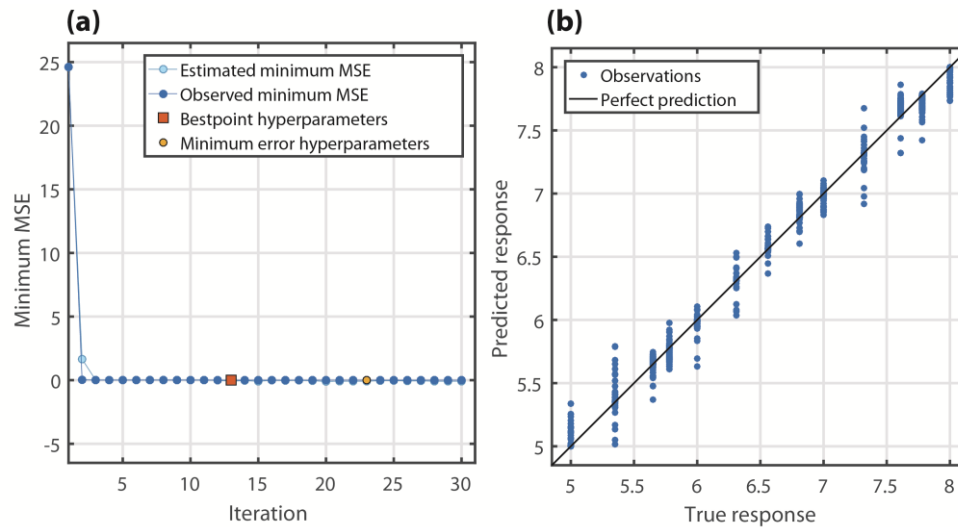

Fig. 10. Estimation of optimized hyperparameters for obtaining minimum MSE represented as (a) Minimum MSE Plot and (b) predicted vs true response plot.
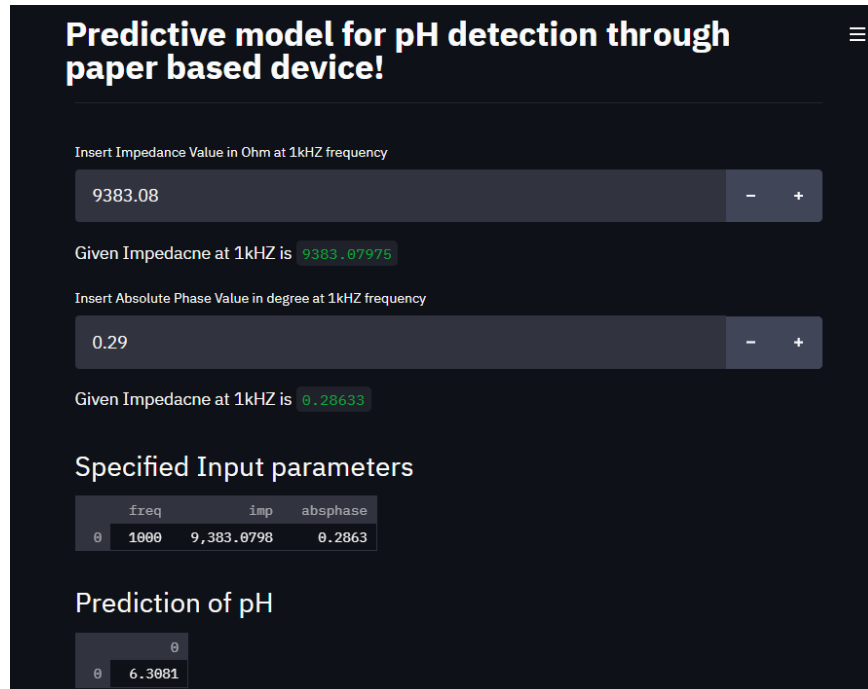
Fig. 11. Python based web app interface to predict the pH value for given feature parameters. The web app is locally deployed and can be accessed through local area network.

application are tabulated in Table IV. The average accuracy for the regression algorithm is found to be 98.41% with standard deviation of 3.19%, which is superior to the average accuracy of manually calculated pH (i.e. 96.35% with 4.72% standard deviation, depicted in Table II). The predicted pH values for real urine sample with pH 5.8 (sample no. 6) and 5.9 (sample no. 4) are

TABLE IV.
Manually calculated pH and predicted pH compared with glass pH electrode reading

Sample no.	pH of urine sample measured by glass electrode	Predicted pH from Web Application	Accuracy of Predicted pH from Web Application (%)	Average accuracy (with standard deviation) (%)
1	6.35	6.30	99.2126	
2	5.61	5.66	100.8913	
3	6.41	6.42	100.156	
4	5.9	5.85	99.15254	
5	6.2	6.30	101.6129	
6	5.80	5.78	99.65517	98.41 ± 3.19
7	6.1	5.99	98.19672	
8	6	5.41	90.16667	
9	5.54	5.29	95.48736	
10	6.77	6.78	100.1477	
11	5.70	5.58	97.89474	

found to be pH 5.78 and 5.85, which is 99.65% and 99.15% accurate, respectively, which is not distinguishable by the manually calculated pH from calibration curve. Similar phenomena can be observed in the case of pH 6.1 (sample no. 7) and 6.2 (sample no. 5). The observed values clearly show that the proposed optimized regression framework overcomes the limitation of sensor resolution efficiently.

5.   Conclusion

The present study proposes a novel smart sensing method for predicting urine pH accurately through an interactive web app interface. The first principle calculations shed light on the electron transport properties and different structural changes of PANI at acidic and alkaline pH, which is not reported earlier. The band calculation and the ELF study further gives insight into the changes of conductivity at varying pH. The impedimetric responses also validate the theoretical analysis obtained from DFT study. The rise of impedimetric responses at alkaline pH confirms the reduction in conductivity as shown in the band structure estimation of PANI-EB. In contrast, at acidic pH, the increased bipolaronic activity is validated by lower band gap values and corresponding ELF study. Incorporation of the hyperparameter optimized regression model improves the sensing accuracy to 98.41%, which overcomes any device resolution related problem reported in our previous study. The use of the sensing platform for other biofluids will be implemented in the near future with a wider range of pH. Furthermore, the developed Python-based web app interface can be deployed through a domain name server to access the sensing platform through a wide area network or WAN.

Appendix A: Supplementary Material

Supplementary tables and figures are included within the supplementary section.

References

Biswas, S., Pal, A., Chaudhury, K., Das, S., 2020. Polyaniline Functionalized Impedimetric Paper Sensor for Urine pH Measurement. IEEE Sens. J. 21, 14474–14482. https://doi.org/10.1109/jsen.2020.3013405

Blöchl, P.E., Jepsen, O., Andersen, O.K., 1994. Improved tetrahedron method for Brillouin-zone integrations. Phys. Rev. B 49, 16223–16233. https://doi.org/10.1103/PhysRevB.49.16223

Brédas, J.L., Scott, J.C., Yakushi, K., Street, G.B., 1984. Polarons and bipolarons in polypyrrole: Evolution of the band structure and optical spectrum upon doing. Phys. Rev. B 30, 1023–1025. https://doi.org/10.1103/PhysRevB.30.1023

Bredas, J.L., Street, G.B., 1985. Polarons, Bipolarons, and Solitons in Conducting Polymers. Acc. Chem. Res. 18, 309–315. https://doi.org/10.1021/ar00118a005

Bria, A., Cerro, G., Ferdinandi, M., Marrocco, C., Molinara, M., 2020. An IoT-ready solution for automated recognition of water contaminants. Pattern Recognit. Lett. 135, 188–195. https://doi.org/10.1016/j.patrec.2020.04.019

Bull, A.D., 2011. Convergence Rates of Efficient Global Optimization Algorithms. J. Mach. Learn. Res. 12, 2879–2904.

Cavazzoni, C., Colle, R., Farchioni, R., Grosso, G., 2006. HCl-doped conducting Emeraldine polymer studied by ab initio Car-Parrinello molecular dynamics. Phys. Rev. B - Condens. Matter Mater. Phys. 74, 033103. https://doi.org/10.1103/PhysRevB.74.033103

Cavazzoni, C., Colle, R., Farchioni, R., Grosso, G., 2004. Ab initio molecular dynamics study of the structure of emeraldine base polymers. Phys. Rev. B - Condens. Matter Mater. Phys. 69, 115213. https://doi.org/10.1103/PhysRevB.69.115213

Dhamu, V.N., Prasad, S., 2020. ElectrochemSENSE: A platform towards field deployable direct

on-produce glyphosate detection. Biosens. Bioelectron. 170, 112609. https://doi.org/10.1016/j.bios.2020.112609

Engert, C., Umapathy, S., Kiefer, W., Hamaguchi, H. o., 1994. Dynamic structure of charge carrier in polyaniline by near-infrared excited resonance Raman spectroscopy. Chem. Phys. Lett. 218, 87–92. https://doi.org/10.1016/0009-2614(93)E1468-V

Epstein, A.J., Macdiarmid, A.G., 1991. Structure, order and the metallic state in polyaniline and its derivatives. Synth. Met. 41, 601–606. https://doi.org/10.1016/0379-6779(91)91142-W

Frison, G., Sevin, A., 1999. A DFT/electron localization function (ELF) study of the bonding of phosphinidenes with n-heterocyclic carbenes. J. Phys. Chem. A 103, 10998–11003. https://doi.org/10.1021/jp991550q

Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G.L., Cococcioni, M., Dabo, I., Dal Corso, A., De Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R., Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A.P., Smogunov, A., Umari, P., Wentzcovitch, R.M., 2009. QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. J. Phys. Condens. Matter 21, 395502. https://doi.org/10.1088/0953-8984/21/39/395502

Hsu, W.E., Chang, Y.H., Lin, C.T., 2019. A machine-learning assisted sensor for chemo-physical dual sensing based on ion-sensitive field-effect transistor architecture. IEEE Sens. J. 19, 9983–9990. https://doi.org/10.1109/JSEN.2019.2927038

Kim, H., Lee, S., Min, J.S., Kim, Eunsu, Choi, J., Ko, J.G., Kim, Eunha, 2021. Fluorescent sensor array for high-precision pH classification with machine learning-supported mobile

devices. Dye. Pigment. 193, 109492. https://doi.org/10.1016/J.DYEPIG.2021.109492

Laboratory Assessment of Kidney Disease, 2011. . Pocket Companion to Brenner Rector's Kidney 21–41. https://doi.org/10.1016/B978-1-4160-6640-8.00002-6

Li, X., Shan, G., Shek, C.H., 2022. Machine learning prediction of magnetic properties of Fe-based metallic glasses considering glass forming ability. J. Mater. Sci. Technol. 103, 113–120. https://doi.org/10.1016/J.JMST.2021.05.076

Louarn, G., Lapkowski, M., Quillard, S., Pron, A., Buisson, J.P., Lefrant, S., 1996. Vibrational properties of polyaniline-isotope effects. J. Phys. Chem. 100, 6998–7006. https://doi.org/10.1021/jp953387e

Lu, H., Yang, L., Fan, Y., Qian, X., Liu, T., 2022. Novel simulation of aqueous total nitrogen and phosphorus concentrations in Taihu Lake with machine learning. Environ. Res. 204, 111940. https://doi.org/10.1016/J.ENVRES.2021.111940

Martín-Gutiérrez, G., Rodríguez-Beltrán, J., Rodríguez-Martínez, J.M., Costas, C., Aznar, J., Pascual, Á., Blázquez, J., 2016. Urinary Tract Physiological Conditions Promote Ciprofloxacin Resistance in Low-Level-Quinolone-Resistant Escherichia coli. Antimicrob. Agents Chemother. 60, 4252–8. https://doi.org/10.1128/AAC.00602-16

Martin, R.M., 2004. Electronic Structure. Cambridge University Press. https://doi.org/10.1017/CBO9780511805769

Momma, K., Izumi, F., 2011. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. J. Appl. Crystallogr. 44, 1272–1276. https://doi.org/10.1107/S0021889811038970

Nogueira, S.A., Sousa, L.R., Silva, N.K.L., Rodrigues, P.H.F., Coltro, W.K.T., 2017. Monitoring Acid–Base Titrations on Wax Printed Paper Microzones Using a Smartphone.

Micromachines 2017, Vol. 8, Page 139 8, 139. https://doi.org/10.3390/MI8050139

Perdew, J.P., Burke, K., Ernzerhof, M., 1996. Generalized gradient approximation made simple. Phys. Rev. Lett. 77, 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865

Pouget, J.P., Jozefowicz, M.E., Epstein, A.J., Tang, X., MacDiarmid, A.G., 1991. X-ray structure of polyaniline. Macromolecules 24, 779–789. https://doi.org/10.1021/ma00003a022

Quillard, S., Louarn, G., Lefrant, S., Macdiarmid, A.G., 1994. Vibrational analysis of polyaniline: A comparative study of leucoemeraldine, emeraldine, and pernigraniline bases. Phys. Rev. B 50, 12496–12508. https://doi.org/10.1103/PhysRevB.50.12496

Rappe, A.M., Rabe, K.M., Kaxiras, E., Joannopoulos, J.D., 1990. Optimized pseudopotentials. Phys. Rev. B 41, 1227–1230. https://doi.org/10.1103/PhysRevB.41.1227

Reis, A.S., Sanches, E.A., Frota, H.O., 2017. Energy band structure and electronic transport properties of chlorine-doped polyaniline from ab initio calculations. Synth. Met. 231, 89–94. https://doi.org/10.1016/j.synthmet.2017.07.004

Silvi, B., Savin, A., 1994. Classification of chemical bonds based on topological analysis of electron localization functions. Nature 371, 683–686. https://doi.org/10.1038/371683a0

Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical bayesian optimization of machine learning algorithms. Adv. Neural Inf. Process. Syst. 25.

Solmaz, M.E., Mutlu, A.Y., Alankus, G., Kılıç, V., Bayram, A., Horzum, N., 2018. Quantifying colorimetric tests using a smartphone app based on machine learning classifiers. Sensors Actuators B Chem. 255, 1967–1973. https://doi.org/10.1016/J.SNB.2017.08.220

Stafström, S., Brédas, J.L., Epstein, A.J., Woo, H.S., Tanner, D.B., Huang, W.S., MacDiarmid, A.G., 1987. Polaron lattice in highly conducting polyaniline: Theoretical and optical studies. Phys. Rev. Lett. 59, 1464–1467. https://doi.org/10.1103/PhysRevLett.59.1464

Stejskal, J., Gilbert, R.G., 2002. Polyaniline. Preparation of a conducting polymer(IUPAC Technical Report). Pure Appl. Chem. 74, 857–867. https://doi.org/10.1351/pac200274050857

Tang, W., Sanville, E., Henkelman, G., 2009. A grid-based Bader analysis algorithm without lattice bias. J. Phys. Condens. Matter 21, 084204. https://doi.org/10.1088/0953-8984/21/8/084204

Wang, R.X., Huang, L.F., Tian, X.Y., 2012. Understanding the protonation of polyaniline and polyaniline-graphene interaction. J. Phys. Chem. C 116, 13120–13126. https://doi.org/10.1021/jp3017612

Wang, Z., Wen, H., Su, Y., Shen, W., Ren, J., Ma, Y., Li, J., 2022. Insights into ensemble learning-based data-driven model for safety-related property of chemical substances. Chem. Eng. Sci. 248, 117219. https://doi.org/10.1016/J.CES.2021.117219

Winokur, M.J., Mattes, B.R., 1998. Structural studies of halogen acid doped polyaniline and the role of water hydration. Macromolecules 31, 8183–8191. https://doi.org/10.1021/ma980586y

Zamora, M.L., Dominguez, J.M., Trujillo, R.M., Goy, C.B., Sánchez, M.A., Madrid, R.E., 2018. Potentiometric textile-based pH sensor. Sensors Actuators B Chem. 260, 601–608. https://doi.org/10.1016/J.SNB.2018.01.002