Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Ai-Driven Coffee Tasting: Predicting Taste Attributes From Electrochemical Imprints

Rodion Golovinsky
ITMO University
Ilya Ivanov
ITMO University
Aleksandr Aglikov
ITMO University
Timur Aliev
ITMO University
Ivan Timofeenko
Siberian Federal University
Alexander Novikov
ITMO University
Mikhail Zayats
Belarusian State University
Mariia Ashikhmina
ITMO University
Olga Orlova
ITMO University
Maria Masalovich
msmasalovich@itmo.ru
ITMO University
Ekaterina Skorb
ITMO University

Research Article

Keywords: coffee, taste, quality, cyclic voltammetry, machine learning (ML), electrochemical imprint

Posted Date: October 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10842069/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

The present study focuses on the problem of taste recognition of coffee by the means of electrochemical imprints. Cyclic voltammetry facilities and machine learning (ML) techniques made it possible to develop a combined method of discrimination of taste profile operating with such sensory categories as sweetness, bitterness, acidity as well as overall quality. The electrochemical responses of four different electrodes in nearly 200 different samples of coffee contributed to the essential databases used for training ML-models via supervised algorithms. The best performance of 'quality' recognition was achieved with the help of a copper and gold electrodes with weighted F1 score of 0.84 for Cu (XGBClassifier) and 0.89 for Au (LogisticRegresssion). Meanwhile, nickel and carbon were the most efficient in the recognition of 'low' and 'middle' acidity with weighted F1 scores of 0.87 for Ni (GradientBoostingClassiffier) and 0.75 for C (LogisticRegression). Gas chromatography-mass spectrometry (GC-MS) identified key volatiles, while Density functional theory (DFT)/docking simulations confirmed effective binding of caffeine, acetic acid and furanmethanol with the electrode surface supporting electrochemical fingerprints as taste proxies.

Introduction

Coffee is one of the most popular beverages in the world, cherished for its rich aroma, complex flavor, stimulating effects as well as antioxidant properties and biological activity of its components [1–6].

Coffee can comprise more than 1,500 chemical substances, with approximately 850 being volatile and 700 soluble [7]. Moreover, such processes as ripening, fermentation, roasting, brewing and blending also balance concentrations of different components creating an unrepeatable flavor of every cup of coffee [8–14]. The taste of coffee ranges from delicate and fruity to bold and chocolatey. Each sensory profile is built on six main attributes: sweetness, acidity, bitterness, body, aroma, and finish.

Chromatographic approaches combined with mass spectrometry (MS) or ultraviolet-visible spectroscopy made it possible to identify substances contributing to the taste [15–18]. Various acids such as malic, citric, tartaric, acetic, lactic, and phosphoric acid were found to be responsible for sourness of coffee [19–21]. Substances producing sweetness are numerous carbohydrates and some amino acids such as glycine [22, 23]. Alkaloids and polyphenolic compounds such as caffeine and chlorogenic acids, respectively, as well as products of caramel pyrolysis after roasting stand for bitterness [24–26].

Up to the present moment, cup tasting produced by qualified personnel has remained a common practice to discover coffee flavor and aroma and hence prescribe certain organoleptic characteristics to a coffee beverage.

Digital approaches have been recently introduced into the analysis of sensory characteristics to avoid subjectivity in appraisal of taste and quality [27–29]. Thus, machine learning combined with molecular docking has demonstrated high efficiency in predicting the intensity of umami taste and bitterness [30–

32]. In addition, the use of computer vision and machine learning methods made it possible to distinguish white and pink wines [33].

Electronic noses and tongues have also gained popularity owing to rapidly growing opportunities for artificial intelligence (AI). In general, they are non-specific electrochemical sensors utilizing signals from electrochemical measurements and converting them into some sensory categories with the help of pattern recognition instruments such as versatile machine learning models [21, 34–41]. The main constraint of the AI-driven method is moderate accuracy and reproducibility caused by relatively small datasets on the one hand, and lack of specific information used in the electrochemical responses, on the other. Potentiometric sensors respond to charged ionic species, leaving uncharged molecules completely undetected while voltammetric sensors utilize signals based on redox activity providing more relevant structure to property dependent information.

In our previous research cyclic voltammetry signals from a multielectrode sensor were successfully processed with ML tools to predict a series of antibiotics in milk [42].

The present research demonstrates an easy and prompt method to receive digital imprints of coffee taste utilizing cyclic voltammetry as one of the simplest electrochemical methods and non-specific electrodes to achieve an accuracy comparable to expert cupping. The results obtained in the study present a new approach for digitizing such taste attributes of coffee as acidity, bitterness, sweetness, and overall quality [43–46].

Experimental section and methods

Tasting method

196 samples of ground coffee were granted by the manufacturers TastyCoffee https://tastycoffee.ru/, Rockets Coffee https://rockets.coffee/, Coffemania https://coffeemania.ru/, Akademia Coffee https://akademiacoffee.com/, and The Whelder Catherine https://theweldercatherine.ru/. Roasting and grinding characteristics of some samples are shown in Table 1. The full list of coffee varieties used in the present work is shown in the repository at https://zenodo.org/records/20840464 (panelists_scores_EN.xlsx).

11 g of coffee were weighed (grind size — coarser than espresso, close to pour-over or aero press) to prepare samples for tasting and measuring. Hot water with temperature of about 92 ºC was poured into cups with coffee in a circular motion. A “cap” of coffee was to be formed on the surface. Coffee samples were waiting for four minutes to brew. The water for brewing had a mineralization of Na^+ 20–200, Ca^{2+} 0–20, Mg^{2+} 0–5, SO_4^{2-} 5-100, Cl^- 5-100 and HCO_3^- 30–450 mg/l. The water composition was controlled using an Aquaphor filtration system to ensure reproducible extraction conditions. When the coffee was brewed, a spoon was immersed in the cup at an angle of 70–80 degrees and with three or four movements, the “cap” was moved to the far wall of the cup. The formation of a “cap” — a floating layer of

coffee grounds — is a standard feature of the Specialty Coffee Association (SCA) cupping protocol. The cap traps volatile aromatic compounds during steeping, enabling their evaluation during the “break” step. The consistent formation of the cap across all samples confirmed reproducible brewing conditions for both sensory assessment and electrochemical measurements.

Table 1
Roasting and grinding description of coffees

№	Name	Roasting type, grinding type	№	Name	Roasting type, grinding type
1	Tasty coffee- Brazil Eduardo Costa	Roasted for a filter coffee maker, roasted fine grind	15	Tasty coffee- Costa Rica San Jose	Espresso roast, roasted fine grind
2	Tasty coffee - Gudji Ethiopia	Espresso roast, roasted fine grind	16	Tasty coffee- Brazil Mogiana	Espresso roast, roasted fine grind
3	Tasty coffee - Nutty Blend	Coffee for dairy drinks, roasted fine grind	17	Tasty coffee- Brazil Lagonya	Roasted for a filter coffee maker, roasted fine grind
4	Tasty coffee - Fuity Blend	Espresso roast, roasted fine grind	18	Tasty coffee- Floral	Espresso roast, roasted fine grind
5	Tasty coffee- Ethiopia Irgacheffe Nat	Espresso roast, roasted fine grind	19	Tasty coffee- Colombia Lime Santuario	Roasted for a filter coffee maker, roasted fine grind
6	Tasty coffee- Colombia Bogota	Espresso roast, roasted fine grind	20	Tasty coffee- Bolivia Mañana	Espresso roast, roasted fine grind
7	Tasty coffee- Ethiopia Oromia	Espresso roast, roasted fine grind	21	Tasty coffee- Bolivia Karanavi	Roasted for a filter coffee maker, roasted fine grind
8	Tasty coffee - Colombia Buenavista	Roasted for a filter coffee maker, roasted fine grind	22	Tasty coffee- Berry	Coffee for dairy drinks, roasted fine grind
9	Tasty coffee- China Simao	Espresso roast, roasted fine grind	23	Tasty coffee- Colombia La Experanza Rojo	Roasted for a filter coffee maker, roasted fine grind
10	Tasty coffee- Brazil Cerrado	Espresso roast, roasted fine grind	24	Tasty coffee- Colombia El Paraiso	Roasted for a filter coffee maker, roasted fine grind
11	Tasty coffee- Colombia Cherry Santuario	Roasted for a filter coffee maker, roasted fine grind	25	Tasty coffee- Candy	Coffee for dairy drinks, roasted fine grind
12	Tasty coffee- Funky	Espresso roast, roasted fine grind	26	Coffemania.ru- Peru La Chirimoya	Medium-roasted

№	Name	Roasting type, grinding type	№	Name	Roasting type, grinding type
13	Coffemania.ru-Espresso	Medium-roasted	27	Rockets coffee-Brazil, Uganda Robusta	Espresso
14	Welder Catherine-Ethiopia Lalisa	Espresso roast	28	Rockets coffee-Bolivia Java Los Rodrigiez	Filter
№	**Name**	**Roasting type, grinding type**	**№**	**Name**	**Roasting type, grinding type**
29	Welder Catherine-Colombia Rafael Aya	Roasted for a filter coffee maker	43	Rockets coffee- Peru, Brazil, Uganda Robusta	Espresso
30	Welder Catherine-Kenya Rungeto G	Espresso roast	44	Rockets coffee-Kenya Chorongi AA	Filter
31	Welder Catherine-Java golden	Roasted for a filter coffee maker	45	Welder Catherine-Ethiopia Sakacha	Roasted for a filter coffee maker
32	Welder Catherine-Brazil Benicio	Roasted for a filter coffee maker	46	Welder Catherine-Peru Keromarca	Roasted for a filter coffee maker
33	Welder Catherine - Costa Rica La Lupa	Roasted for a filter coffee maker	47	Welder Catherine-Brazil Jaguar Peaberry	Espresso roast
34	Welder Catherine - Costa Rica Juanra Montero	Roasted for a filter coffee maker	48	Coffemania.ru-Ethiopia Raro	Medium-roasted
35	Welder Catherine-Brazil Jaguar Peaberry	Coffee for dairy drinks	49	Coffemania.ru-Bolivia Trapiche	Medium-roasted
36	Welder Catherine-Indonesia Java, washed	Roasted for a filter coffee maker	50	Rockets coffee-Colombia El indio	Filter
37	Welder Catherine-Kenya Tiririka Ndundu AA	Roasted for a filter coffee maker	51	Rockets coffee-Ethiopia Arbegona	Filter/ Espresso
38	Welder Catherine-Guatemala Pablo Godinez	Roasted for a filter coffee maker	52	Rockets coffee-Honduras Joel Vasguez	Filter
39	Rockets coffee-Ethiopia Arbegona	Filter/ Espresso	53	Rockets coffee-Ethiopia Arbegona	Filter

№	Name	Roasting type, grinding type	№	Name	Roasting type, grinding type
40	Rockets coffee-Dekaf	Filter	54	Welder Catherine-Bolivia Capacabana	Roasted for a filter coffee maker
41	Rockets coffee-Rwanda Muteteli	Filter	55	Rockets coffee-Brazil, Ethiopia	Espresso
42	Rockets coffee-Ethiopia, Colombia	Espresso	56	Rockets coffee-Burundi Butegana	Filter

The degustation process of coffee samples was conducted in accordance with SCA cupping method by qualified experts https://www.scith.coffee. Each coffee was independently evaluated by two to four qualified panelists — members of the Russian Specialty Coffee Association (RSCA). A total of 196 coffees were assessed by a pool of four panelists (526 individual evaluations) with 193 samples receiving at least two independent assessments (98 by two, 53 by three, and 42 by all four panelists).

Coffee was evaluated for seven sensory components – aroma, bouquet, aftertaste, body, acidity, bitterness and sweetness. Afterwards, cupping sheets were filled in with taste specifications (berries, nuts, grapes etc.), intensity of the sensory components and assessment scores (panelists_scores_EN.xlsx at https://zenodo.org/records/20840464).

Two different scales were used for the evaluation of coffee taste. The first scale was aimed at ranking the intensity of the listed sensory components and comprised five degrees (low, below middle, middle, above middle, high). The second scale was to rate relative quality of the particular sensory attribute based on the panelists' perception. This appraisal was carried out on a 16-point numerical scale representing levels of quality in quarter point increments between 6 and 9, where 6 is good and 9 is outstanding. The values below 6 are out of specialty grade. 87.6% of scores differed by no more than 1.0 point across all pairwise comparisons. The median score across evaluating panelists was used for each sensory attribute. Samples were excluded when panelists' assessments were incomplete or when any pairwise score difference exceeded 1.5 points — a difference sufficient to place scores in different categories.

Electrochemical measurements and database collection

Cyclic voltammetry in a three-electrode electrochemical cell was chosen for both: obtaining imprints of coffee and studies of sensors in a background electrolyte. Copper, nickel, gold, and glassy carbon were used as working electrodes in consequence, i.e. played a role of voltammetric sensors. They were successively polished with P1200 and P2500 grade sandpaper after each measurement and then rinsed with water. A silver chloride electrode was used as the reference. A platinum disc electrode was used as the counter. The working electrodes and the counter had an active surface of 7.065 mm^2. Potentiostat-galvanostat R-20X (Electrochemical Instruments, Russia) was used as a measuring device.

4 ml of 1M sodium phosphate buffer solution and 0.202 g of potassium nitrate were added to 16 ml of brewed coffee. 16 ml of water for brewing instead of brewed coffee were taken to obtain a blank electrolyte solution. The analytes were stirred with a glass rod for 1 minute. The measurements were conducted in the potential range from − 2.0 to + 2.0 V at a sweep rate of 50 mV/s. 6 voltammetric cycles per sample of brewed coffee were run with the first cycle being idle to obtain the database of imprints.

196 different varieties of coffee were measured totally. Each variety was measured four times in accordance with the number of sensors (Cu, Au, Ni and C) to form four separate databases. The cyclic voltammograms were saved as .edf files representing columns of current-voltage-time values (available at https://zenodo.org/records/20840464).

Machine learning methods

Preprocessing sensory data and label encoding

Three key sensory categories: sweetness, acidity, and bitterness were utilized to address the problem of taste prediction. The intensities prescribed by the panelists were categorized into three classes: low, middle, and high, middle class being comprised of lower than middle, middle and higher than middle subclasses together to simplify ML task.

The categories of ‘good’ and ‘bad’ were used to address the problem of quality prediction. The score scale of taste attributes was utilized for this purpose (panelists_scores_EN.xlsx at https://zenodo.org/records/20840464). The average score for the perceptions of acidity, sweetness and bitterness was calculated, and a threshold value of 7.5 was established. The coffee samples with an average score above this threshold were classified as ‘good’, while the samples with a score equal to or below 7.5 were classified as ‘bad’.

The distribution of the target features is shown in Fig. 1. It can be seen, the initial distribution is non-uniform throughout all the sensory categories (Fig. 1a, c, e). ‘Middle sweetness’ class three times exceeds ‘low sweetness’ class (Fig. 1a), meanwhile ‘middle acidity’ class is twice as big as ‘low acidity’ one (Fig. 1c). In case of bitterness, ‘low’ category is, on the opposite, the most numerous (Fig. 1e), and only ‘good/ bad' dataset looks quite uniform (Fig. 1g). To minimize the non-uniformity of the data in the sensory categories, the high classes were combined with middle ones (Fig. 1b, d, f) followed by subtraction of ‘lower than middle’ values and merging them with low classes. Then the problem of binary classification was subsequently solved for sweetness, acidity and bitterness with the classes encoded as ‘low’ and ‘middle’ and for quality with the classes of ‘good’ and ‘bad’.

Preprocessing electrochemical data

The electrochemical data was filtered in accordance with the fullness of information provided by the panelists. In the case of lack of information concerning panelists’ assessments or in case of opposite or

significantly different expert opinions, the electrochemical data for the corresponding samples of coffee were deleted. For this reason, 162 samples of coffee were involved in the classification based on the sensory categories of acidity, sweetness and bitterness, and only 92 samples were used for the detection of quality.

There were two alternative approaches to preprocessing electrochemical data. The first one is based on the sequences of experimental current values from .edf files. The second approach suggests extraction of new features from the current sequences.

As received tabular data from cyclic voltammograms, i.e. the arrays of current per time values (voltage is time-dependent via scan rate) was successfully used for the quality classification (‘good/ bad’).

The strategy with feature extraction was performed for the taste classification (‘low/ middle’). The features were extracted via correlation analysis from the data rows of current using tsfresh library https://tsfresh.readthedocs.io/en/latest/ with default parameters of the extract_features function and without an additional manual selection. They comprised such metrics as mean values, standard deviation, FFT (fast Fourier transformation) coefficients, autocorrelation function, number of local peaks etc. with the total number of about 100. Furthermore, StandardScaler was applied to transform each feature (column) to have zero mean and unit variance. This normalization step ensured that all features comparably contributed to the learning process and improved the stability and convergence of the machine learning models.

Data splitting

The datasets were split into training and testing arrays comprising 80 and 20% of electrochemical data, respectively, followed by tuning the representation of each class to equal proportion in both arrays.

Model design

Classical machine learning models were chosen along with gradient boosting models, the latter performing well with small and sparse datasets. The following predictive algorithms for classification imported from the scikit-learn library https://scikit-learn.org/stable/about.html [47] were used: MLPClassifier https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html, LogisticRegression https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html, GradientBoostingClassifier https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html, SVC https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html, RandomForestClassifier https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html and KNeighborsClassifier https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html. Additionally,

boosting models such as CatBoost https://github.com/catboost/catboost [48] and XGBoost (XGBoostClassifier, XGBClassifier) https://github.com/dmlc/xgboost/ [49] were imported from the corresponding libraries.

Hyperparameter tuning

Cross-validation with a five-fold split was performed to analyze the stability of model performance across different data partitions. Hyperparameter tuning for all machine learning models was performed using a combination of GridSearch and Optuna. GridSearch was applied to explore predefined parameter grids, ensuring systematic coverage of key hyperparameter values. In addition, Optuna was used for more efficient optimization via adaptive sampling and pruning strategies which iteratively explored the hyperparameter space and discarded unpromising trials at early steps. This approach allowed for faster convergence to optimal configurations compared to exhaustive search improving both computational efficiency and model performance.

Evaluation of ML models

38 classification models were obtained overall. The models were directly evaluated on the testing set calculating accuracy and weighed F1 Score.

Gas chromatography–mass spectrometry analysis

To perform gas chromatography, the analyte was prepared as follows: 1.34 g of ammonium sulfate was added to 2 ml of coffee. The solution was shaken until the salt dissolved. Then 2 ml of diethyl ether were added to the solution. The analyte was shaken for 2 min. After settling for 5–10 min, the analyte was transferred to 2 ml vials. GC-MS analysis was performed using a 5890 Series II gas chromatograph with a mass selective detector 5972 (Hewlett Packard, Germany). The GC was equipped with a 50 m×0.2 mm i.d., 0.4 μm film, HP-INNOWAX (Crosslinked polyethylene glycol) column, HP Part No. 4.293.415 (Hewlett Packard, USA).

2 μL of a sample were manually injected in split/ splitless injector in split mode using a split ratio of 30. The injector temperature was maintained at 250°C.

Helium of A grade quality (99.995%) was used as carrier gas at the constant pressure of 194 kPa (which corresponds to 0.9 ml/min at 50°C). The oven temperature was maintained at 50 ºC for 3 min after the injection and then programmed to increase at 3 ºC min^{-1} to 140 ºC, followed by isothermal holding for 10 min, then the increase at 7 ºC min^{-1} to 200 ºC, isothermal holding for 10 min, and the final increase at 10 ºC min^{-1} to 225 ºC and maintaining for 55 min.

The transfer line to the MS was maintained at 280°C. Mass spectrometry was performed with electron-impact (EI) ionization (electron energy was 70 eV) at an ion source temperature of 178°C. Data were

acquired in total ion current (TIC) mode, after a solvent cut between 7.5 and 120 min. The Chemstation software (Hewlett Packard, USA) was used for instrument control and data handling.

Alternatively, GS-MS chromatography was carried out at GC-MS Shimadzu GC-MS QP2010 Ultra, equipped with Supelco Carboxen/DWB/PDMS column and CombiPAL autosampler. Brewed coffee was transferred to a 20 ml vial with a 10 ml dispenser. Then samples were thermostatted for 15 minutes at 60°C. The following procedure was as described above.

DFT calculation computational details

The density functional theory (DFT) calculations including full geometry optimization procedure for all model structures were performed in the Orca 6.0.0 program [50] package using the ωB97X-D4 [51] functional, the def2-QZVPPD basis sets for Cu, Ni, Au atoms and the def2-TZVPD [52] basis sets for all other atoms. No symmetry constraints were imposed during optimizations. To confirm the absence of imaginary frequencies, Hessian matrices were computed for all structures, ensuring that optimized geometries correspond to true energy minimum on the potential energy surface. Thermodynamic parameters were calculated at 298.15 K and 1.00 atm. The spin multiplicities were considered 2 for Cu, 3 for Ni, and 2 for Au. Cartesian atomic coordinates of all optimized structures are provided as xyz-files in Supplementary Information.

Molecular docking computational details

AutoDock Vina v. 1.2.0 was used for all dockings in this study [53]. The docking parameters were kept at their default values. The size of the docking grid for copper was 18 × 40 × 40 Å. The results of molecular docking were visualized by using PyMol 2.5 software. 3D structures of ligands were downloaded from PubChem database and were used as received [54].

Results and discussion

The workflow of collecting experimental data for machine learning modelling is shown in Fig. 2. Roasted coffee seeds were processed to fine grind by the manufacturers. The ground coffee was brewed by the hot steeping technology in the laboratory under ambient conditions and then filtered to remove the solid waste. The coffee beverages were then evaluated by the panelists, and corresponding cupping sheets were filled in. The assessments concerning the intensity of acidity, sweetness, bitterness as well as the individual perception of these features were then used as the ground truth to make inputs in ML models.

Figure 2 Workflow chart depicting the procedure of obtaining experimental data for prediction of coffee taste and quality. The input of ML-models comprised the panelists' assessments of coffee taste and quality as well as the voltammetric data derived from cyclic voltammograms. Predictions of taste and quality were the output of the models

Electrochemical data obtained by the voltammetric tongues were used to train the ML-models. For this purpose, potassium nitrate and phosphate buffer solution were added to coffee samples to obtain the

corresponding concentrations of 0.1 M and 0.2 M, respectively, and played a role of supporting electrolyte. Cyclic voltammograms of copper, nickel, gold or glassy carbon were measured in the brewed coffee in consequence and saved as .edf files.

Each measurement was considered a fingerprint of the coffee sample comprising such features as sweetness, bitterness, acidity and overall quality based on average panelists' perception. Cyclic voltammetry response was then referred to the ground truth while training and the predictions were finally obtained. The models were taught to predict the intensity of three categories of taste ('low or middle') and quality ('good or bad').

Electrochemical studies

Electrochemical studies were performed with the use of cyclic voltammetry. Gold and glassy carbon electrodes were scanned in a three-electrode cell in a brewed coffee and a blank electrolyte based on the water used for coffee extraction, respectively. The experiments were aimed at obtaining specific features relevant to the electrochemical imprint of coffee. Figure 2a depicts cyclic voltammograms of gold electrode in a sample of brewed coffee compared with blank electrolyte. The difference in current values is evident as well as the appearance of some intense peaks on the cyclic voltammograms measured in the brewed coffee. Anodic current maxima in the potential region from + 0.4 to + 0.6 V (Fig. 2a, b) are probably attributed to caffeine and phenolic compounds [55–60]. In this potential range, OH^- groups are expected to reversibly oxidize, whereas the peak at a higher potential of 1.4 V can be associated with irreversible oxidation of caffeine.

The reduction peak at the gold electrode observed at a potential of + 0.6 V (Fig. 2a) obviously stands for reverse transformation from ketone groups to enol ones in phenols. Current increase in the negative range of potentials at around − 1.0 V at the gold electrode and − 0.6 V at the glassy carbon electrode (Fig. 2b) can be associated with protonic region [56] correlated with acidity. The negative range below − 1.4 V at gold and − 1.5 V at glassy carbon is responsible for hydrogen evolution reaction.

Figure 3 Cyclic voltammograms of a gold (a) and a glassy carbon (b) electrode in coffee samples (1) containing 0.2 M PBS and 0.1 M KNO_3, and blank electrolyte solutions (2) without coffee. The inset in (a) duplicates cyclic voltammograms of gold in the blank solution in a shortened range of potentials. Coordinate scale is the same for both plots in (a). The potentials are measured versus a silver chloride electrode as reference. Platinum wire is used as a counter electrode. The scanning rate is 50 mV s^{-1}

Cyclic voltammograms perform a superposition of all electrochemically active species and thus, may be considered a digital imprint of a coffee sample.

Machine learning

Figure 4 demonstrates the confusion matrixes revealing accuracy of the most effective ML-models. A typical series of voltammograms measured for the copper database and used for machine learning is

shown in Fig. 4a. In fact, two basic approaches were used to train ML models. According to the first one, implemented for the classification task of the quality of coffee, the input consisted of arrays of time dependent current values. The second one comprised only statistical metrics such as mean values, standard deviation, FFT (fast Fourier transformation) coefficients, autocorrelation function, number of local peaks, which had previously been derived from the cyclic voltammograms, and was implemented for the problem of taste classification.

Datasets from four voltammetric tongues, i.e. copper, nickel, gold and glassy carbon electrodes, were used to build the corresponding supervised ML-models with the use of a series of algorithms. Copper sensor exhibited moderate results in the sweetness classification (Fig. 4b) with accuracy of prediction of the middle class of 0.88, while the low class was correctly recognized only in 0.57 cases. As for acidity, three sensors out of four showed satisfactory results: gold, nickel and glassy carbon, highlighting nickel as the most effective among them (Fig. 4c-e). The corresponding confusion matrix for the nickel sensor revealed the accuracies of 0.75 and 0.91 for low and middle classes, respectively (Fig. 4d).

Bitterness turned out to be the least recognizable sensory category for the classification of coffee samples. The most relevant results were demonstrated by XGBClassifier using the gold sensor (Fig. 4f), although the predicted middle class was confused with the low one in more than 50% of cases.

The quality category with the predictions ‘good/ bad’ was successfully implemented for the classification. In general, the accuracy of correct detection achieved 100% for good and 80% for bad coffee (Fig. 4g, h). As mentioned above, in the case of quality prediction, the distribution of classes was close to uniform, which simplified the classification task, reduced the need to use balancing methods and made it possible to achieve high metrics including F1 scores. The F1 score of 0.89 was demonstrated by one of the simplest LogisticRegression models on the voltammetric data of a gold electrode, while for a copper electrode, the best F1 score was shown by the XGBoostClassifier model (0.84).

On comparing the performance of the sensors, it should be noted that nickel and carbon were the most efficient in the recognition of low and middle acidity with weighted F1 scores of 0.87 for Ni (GradientBoostingClassiffier) and 0.75 for C (LogisticRegression). Meanwhile, ‘quality’ prediction was successfully accomplished with copper and gold electrodes demonstrating F1 score of 0.84 for Cu (XGBClassifier) and 0.89 for Au (LogisticRegresssion). The results obtained confirm the hypothesis that the taste and quality of coffee can be accessed via voltammetric imprints.

GC-MS

To evaluate chemical composition several samples of coffee were analyzed with gaseous chromatography. It revealed acetic acid, 2-furanmethanol and caffeine possessing the most prominent peaks (Table 22).

Table 2
Gas chromatography results for dominant compounds (average of three samples)

Substance	Retention time [min]	Area of peaks [mm^2]	Height of peaks [mm]
Acetic acid	21.10	5.94	9.27
2-Furanmethanol	29.98	5.83	20.03
Caffeine	89.98	36.89	84.70

Alternative GC-MS revealed 59 components, among which cis-1,2-dihydrocatechol, a phenolic compound, as well as carbamic, hexanoic and acetic acids were detected. The corresponding chromatogram is shown in Fig. SI 1 (Supplementary information), and all the peaks are characterized in Table SI 1. It can be assumed that caffeine, 2-furanmethanol and cis-1,2-dihydrocatechol stand for bitterness, whereas organic acids are responsible for acidity. It can also be assumed that sweetness is caused by the presence of some alcohols like ethanediol and heptane-2,5-diol as well as aldehydes like 2-methylpropanal and 2-methybutanal.

DFT calculations results

Quantum chemical calculations were performed to evaluate the thermodynamic favorability of associations between caffeine, 2-furanmethanol, acetic acid or acetate anion with Cu, Ni, Au and C model species. These systems simulate adsorption processes of the molecular species on the corresponding electrode surfaces.

For 2-furanmethanol, three bonding modes were analyzed: (i) coordination through the oxygen atom of the hydroxyl group, (ii) bonding via the carbon atom, and (iii) interaction with the π-system of the furan ring. Similarly, for caffeine, two bonding configurations were examined: (i) interaction with the π -system of the pyrimidine ring and (ii) coordination through the N-atom of the imidazole ring. The calculated enthalpies, entropies, and Gibbs free energies (in Hartree) for pristine structures and their associates are shown in Table SI 2. The Gibbs free energies in kilocalories per mole for the stable associates are presented in Table SI 3.

The results on the copper support are summarized in Fig. 5a indicating stable configurations at the coordination of copper with the N-atom in caffeine (ΔG = -6.0 kcal/mole) and with the π-system of the furan ring in 2-furanmethanol (ΔG = -1.7 kcal/mole). Among all the considered molecular species, associates of copper with an acetate anion exhibited the highest thermodynamic stability (ΔG = -25.0 kcal/mole), while associates with acetic acid were the least stable (ΔG = 5.7 kcal/mole), as reflected by the positive value of Gibbs free energy. This suggests that acetic acid must undergo deprotonation prior to bonding with the electrode surface. Caffeine formed the second-most thermodynamically stable associate.

The comparative analysis of Gibbs free energies revealed that the stability of molecular associations followed the order: Ni > Cu > Au, with Ni forming the most stable associates and Au the least stable.

Affinity energies were alternatively calculated with AutoDock Vina v. 1.2.0 using copper as a support (Fig. 5b, c). The tendency of the increase of the values from caffeine (ΔG = -4.4 kcal/mole) to 2-furanmethanol (ΔG = -2.9 kcal/mole) is shown the same as by the DFT studies, confirming the overall thermodynamical favorability of the adsorption of these molecules on the surface of copper.

Conclusions

An AI-driven tasting method of coffee based on cyclic voltammetry response from non-modified metallic electrodes and glassy carbon was successfully elaborated. Nearly 200 samples of coffee were tested by the qualified panelists to attribute each sample a tasting profile including the intensity of sweetness, bitterness, acidity and relative quality of these sensory categories. At the same time, voltammetric measurements were carried out to obtain arrays of current values unique for each sample of coffee and the chosen electrode material. Cyclic voltammograms and the electrochemical data derived from them were considered the unique digital imprints of the coffee samples. Four data frames referring to the corresponding type of voltammetric sensor were used to solve the problem of binary classification via ML-algorithms: low or middle classes in case of the sensory categories of taste, and good or bad classes in case of quality. The copper and gold electrodes demonstrated the best efficiency in terms of quality prediction of ‘good’ and ‘bad’ with F1 score of 0.84 for Cu (XGBClassifier) and 0.89 for Au (LogisticRegresssion). The nickel and carbon electrodes were quite effective in acidity classification of ‘low’ and ‘middle’ with weighted F1 scores of 0.87 for Ni (GradientBoostingClassiffier) and 0.75 for C (LogisticRegression). Boosting algorithms along with LogisticRegression showed the best performance to solve these problems. Additional studies by GC, DFT, molecular docking and cyclic voltammetry were carried out to support the concept of a voltammetric imprint as the superposition of redox and adsorption curves affected by a series of substances present in coffee. The results obtained in the study present a new idea for digitizing the taste characteristics of beverages using coffee as an example, combining machine learning approaches and electrochemical analysis methods to predict its taste and quality.

From a practical standpoint, the proposed electrochemical approach offers several advantages for real-world implementation. The assay consisting of a potentiostat-galvanostat and standard metallic electrodes is relatively inexpensive: approximately $3,000–15,000 for commercially available potentiostat systems such as Metrohm Autolab, Gamry, or PalmSens, based on manufacturer catalog prices, and from $1,500 for portable models compared to NIR spectrometers ($20,000–50,000) or GC-MS/HPLC instruments ($50,000–200,000). Running costs are low: consumables consist of electrolyte solutions (phosphate buffer, KNO_3), electrode polishing materials, and water; metallic electrodes can be reused after mechanical polishing for hundreds of measurement cycles. A single measurement comprises 6 voltammetric cycles in the range from − 2.0 to + 2.0 V at 50 mV/s, requiring approximately 16 minutes per electrode. The electrochemical approach could similarly benefit from parallel multi-

electrode setups or screen-printed disposable electrodes in future implementations. The method is best suited as a rapid screening tool for quality control at production stages, especially as a first-pass filter, where speed, objectivity and cost-efficiency are paramount.

Data and Software availability

The datasets generated and analyzed during this study are available in the Zenodo repository at https://zenodo.org/records/20840464 under Creative Commons Attribution Non Commercial 4.0 International license. The data are provided exclusively for non-commercial research purposes and are not available for commercial use. For additional inquiries regarding the software or data access, please contact the corresponding author.

The code used in this study is available in the GitHub repository at https://github.com/RodionGolovinsky/digital_coffee_tester. This repository contains the source code. The software developed for this study is available in the GitHub repository at https://github.com/RodionGolovinsky/coffee_tester/tree/main under MIT License. The repository contains the source code, documentation and usage instructions. Researchers interested in replicating or extending this study can access the repository to review the code implementation and perform further analyses.

Declarations

Supplementary Information

Fig. SI 1 (GC chromatogram, PDF), Table SI 1 (GC data, PDF), Table SI 2 (DFT-calculated enthalpies, entropies, and Gibbs free energies, PDF), Table SI 3 (DFT-calculated values of Gibbs free energies of association reactions, PDF), xyz-files (DFT coordinates).

Conflicts of Interest

The authors declare no competing financial or personal interests that could have appeared to influence the work reported in this paper.

Author Contribution

Conceptualization: A.A. (lead), I.T. (lead), E.S. (supporting); Data curation: R.G. (lead), I.I. (lead), T.A. (equal); Formal analysis: R.G. (equal), I.I. (equal); Funding acquisition: A.A. (equal), I.T. (equal), M.A. (supporting); Investigation: A.A. (lead), I.D. (lead), I.T. (lead), A.N. (equal), M.Z. (lead), M.M. (lead); Methodology: A.A. (lead), I.T. (equal), E.S. (lead); Project administration: A.A. (lead), I.T. (lead), E.S. (lead); Resources: I.T. (lead), E.S. (lead); Software: R.G. (lead), T.A. (lead); Supervision: O.O. (supporting), E.S. (lead); Validation: R.G. (equal), I.I. (equal), T.A. (equal), M.A. (equal); Writing – original draft: R.G. (lead),

I.I. (supporting), M.M. (lead); Writing – review & editing: R.G. (equal), I.I. (equal), A.A. (lead), M.M. (lead), E.S. (lead).

Acknowledgement

The authors express their gratitude to the Blue Sky Research competition for infrastructural support. We sincerely thank Denis Baigozin for conducting the GC-MS measurements of coffee samples, supported by Aquaphor https://www.aquaphor.ru/. Special thanks are extended to Tasty Coffee, Rockets Coffee, Coffemania.ru, Akademia Coffee, and Whelder Catherine for providing ground coffee samples. We also gratefully acknowledge the Russian Specialty Coffee Association https://scarussia.ru/ members – Vladimir Evdokimov, Tatyana Leonova, Andrey Nesterov and Anton Pechurkin – for their participation in the sensory evaluation.

Data Availability

Data and Software availabilityThe datasets generated and analyzed during this study are available in the Zenodo repository at https://zenodo.org/records/20840464 under Creative Commons Attribution Non Commercial 4.0 International license. The data are provided exclusively for non-commercial research purposes and are not available for commercial use. For additional inquiries regarding the software or data access, please contact the corresponding author.The code used in this study is available in the GitHub repository at https://github.com/RodionGolovinsky/digital_coffee_tester. This repository contains the source code. The software developed for this study is available in the GitHub repository at https://github.com/RodionGolovinsky/coffee_tester/tree/main under MIT License. The repository contains the source code, documentation and usage instructions. Researchers interested in replicating or extending this study can access the repository to review the code implementation and perform further analyses.Supplementary InformationFig. SI 1 (GC chromatogram, PDF), Table SI 1 (GC data, PDF), Table SI 2 (DFT-calculated enthalpies, entropies, and Gibbs free energies, PDF), Table SI 3 (DFT-calculated values of Gibbs free energies of association reactions, PDF), xyz-files (DFT coordinates).

References

1. Ding F, Huang G, Sun Y et al (2026) Ultrasound Pretreatment Modifies the Physicochemical Properties and Enhances the Antioxidant Activity of Coffee Leaf Kombucha. Food Bioscience. https://doi.org/10.1016/j.fbio.2025.108158
2. Morais SL, Rede D, Ramalhosa MJ et al (2023) Assessment of the Antioxidant Capacity of Commercial Coffee Using Conventional Optical and Chromatographic Methods and an Innovative Electrochemical DNA-Based Biosensor. Biosensors. https://doi.org/10.3390/bios130908401
3. Socała K, Szopa A, Serefko A et al (2021) Neuroprotective Effects of Coffee Bioactive Compounds: A Review. Int J Mol Sci. https://doi.org/10.3390/ijms22010107

4. Samsonowicz M, Regulska E, Karpowicz D et al (2019) Antioxidant Properties of Coffee Substitutes Rich in Polyphenols and Minerals. Food Chem. https://doi.org/10.1016/j.foodchem.2018.11.057
5. Tang J, Xu Z, Wang F et al (2025) Caffeine Inhibits Tau Aggregation and Destabilizes the Fibril Associated with Chronic Traumatic Encephalopathy: A REST2 and Conventional MD Simulations Study. J Chem Inf Model. https://doi.org/10.1021/acs.jcim.4c02353
6. Kodape A, Kodape A (2024) Exploring Coffee's Impact: Aflatoxins, Phytochemicals, and Public Health Considerations. Food Bioscience. https://doi.org/10.1016/j.fbio.2024.104991
7. Sharma H (2020) A Detail Chemistry of Coffee and Its Analysis. IntechOpen. 10.5772/intechopen.91725
8. Amaral EP, Entringer TL, Veloso TGR et al (2026) Microbial Dynamics and Impact of Controlled Fermentation in Coffea Arabica over 288 h. Food Bioscience. https://doi.org/10.1016/j.fbio.2025.108081
9. Ferreira LJC, Casé IN, Bertarini PLL et al (2024) Impact of Immature Coffee Fruits and Water Addition during Spontaneous Fermentation Process: Chemical Composition and Sensory Profile. Electron J Biotechnol. https://doi.org/10.1016/j.ejbt.2024.04.001
10. Gulshan A, Yiasmin MN, Zeb M et al (2025) Influence of Brewing Method on the Nutritional Composition and Sensory Attributes of Coffee by Arabica Beans of Different Roasting Degrees. Food Chem. https://doi.org/10.1016/j.foodchem.2025.147716
11. Osorio Pérez V, Matallana Pérez LG, Fernandez-Alduenda MR et al (2023) Chemical Composition and Sensory Quality of Coffee Fruits at Different Stages of Maturity. Agronomy. https://doi.org/10.3390/agronomy13020341
12. Yildirim S, Gok I, Demir E et al (2022) Use of Electrochemical Techniques for Determining the Effect of Brewing Techniques (Espresso, Turkish and Filter Coffee) and Roasting Levels on Total Antioxidant Capacity of Coffee Beverage. J Food Process Preserv. https://doi.org/10.1111/jfpp.16626
13. Hackler M, Särchen V (2025) COFFEE PROCESSING. In: The Art and Science of Coffee Fermentation. Springer, Cham. https://doi.org/10.1007/978-3-031-91599-4_5
14. Soares EA, Bento GP, Verruma-Bernardi MR et al (2026) Ranking Descriptive Analysis and Preference of Cold Brew Coffee: Impact of Roasting Degree. Food Measure. https://doi.org/10.1007/s11694-026-04500-w
15. Amalia F, Irifune T, Takegami T (2023) Identification of Potential Quality Markers in Indonesia's Arabica Specialty Coffee Using GC/MS-based Metabolomics Approach. Metabolomics. https://doi.org/10.1007/s11306-023-02051-5
16. Choe W, Woo JH, Hong M et al (2023) Determination of Aromatic Profiles of Coffee Beans According to Different Roasting Times by SPME/GC-MS Analysis. J Appl Pharm Sci. http://doi.org/10.7324/JAPS.2023.143318
17. Jeszka-Skowron M, Zgoła-Grześkowiak A, Grześkowiak T (2015) Analytical Methods Applied for the Characterization and the Determination of Bioactive Compounds in Coffee. Eur Food Res Technol.

https://doi.org/10.1007/s00217-014-2356-z

18. Kvirencova JR, Bechynska K, Hrbek V et al (2026) LC-MS–Based Multimarker Strategy for Quantification of Arabica and Robusta in Coffee Blends. Food Anal Methods 19:247. https://doi.org/10.1007/s12161-026-03173-x
19. Aurum FS, Zaman MZ, Purwanto E et al (2023) Coffee Authentication via Targeted Metabolomics and Machine Learning: Unveiling Origins and Their Discriminating Biochemicals. Food Bioscience. https://doi.org/10.1016/j.fbio.2023.103122
20. Rune CJB, Giacalone D, Steen I et al (2023) Acids in Brewed Coffees: Chemical Composition and Sensory Threshold. Curr Res Food Sci. https://doi.org/10.1016/j.crfs.2023.100485
21. Zhou R, Zhu J, Li Y et al (2025) The Evaluation of Interactions and Molecular Mechanisms between Three Acidic and Sweet Taste Substances with Sensory Analysis, Electronic Tongue, and Molecular Dynamics Simulation. Food Bioscience. https://doi.org/10.1016/j.fbio.2025.107915
22. Da Silva Dias T, Bruneli Silva F, Ferreira EJ et al (2026) Characterization and Correlation of Roasted and Ground Coffee Monosaccharides for Fraud Detection. Food Chem. https://doi.org/10.1016/j.foodchem.2026.147992
23. Mei S, Ding J, Chen X, Identification of Differential Volatile and Non-volatile Compounds in Coffee Leaves Prepared from Different Tea Processing Steps Using HS-SPME/GC–MS and, HPLC-Orbitrap (2023) -MS/MS and Investigation of the Binding Mechanism of Key Phytochemicals with Olfactory and Taste Receptors Using Molecular Docking. Food Res. Int. https://doi.org/10.1016/j.foodres.2023.112760
24. Bayer S, Mayer AI, Borgonovo G et al (2021) Chemoinformatics View on Bitter Taste Receptor Agonists in Food. J Agric Food Chem. https://doi.org/10.1021/acs.jafc.1c05057
25. Schwarzmann ET, Washington MP, Rao NZ (2022) Physicochemical Analysis of Cold Brew and Hot Brew Peaberry Coffee. Processes. https://doi.org/10.3390/pr10101989
26. Tamanna N, Mahmood N (2015) Food Processing and Maillard Reaction Products: Effect on Human Health and Nutrition. Int J Food Sci. https://doi.org/10.1155/2015/526762
27. Sagita D, Mardjan SS, Widodo S et al (2025) Emerging Trends and Developments in Coffee Quality Assessment and Characterization Techniques: A Systematic Review and Bibliometric Analysis. Food Measure. https://doi.org/10.1007/s11694-025-03680-1
28. Setyaningsih W, Prajna D, Larasati ID et al (2026) A Non-Destructive Classification Method of Peaberry and Non-Peaberry Coffee Using Near Infrared Spectroscopy and Machine Learning Algorithms. Food Measure. https://doi.org/10.1007/s11694-026-04281-2
29. Rocha RAR, Cruz MAD, Silva LCF et al (2024) Evaluation of Arabica Coffee Fermentation Using Machine Learning. Foods. https://doi.org/10.3390/foods13030454
30. Ashikhmina MS, Zenkin AM, Pantiukhin IS et al (2024) Uncovering the Taste Features: Applying Machine Learning and Molecular Docking Approaches to Predict Umami Taste Intensity of Peptides. Food Bioscience. https://doi.org/10.1016/j.fbio.2024.105358

31. Ashikhmina MS, Nesterov PV, Volkova O et al (2025) Computational Profiling of Food-Derived Bitter Compounds as Ligands of TAS2R Receptors. Food Bioscience. https://doi.org/10.1016/j.fbio.2025.107907
32. Shityakov S, Skorb EV, Nosonovsky M (2025) TAS2R Taste Receptor Clustering Suggests That Bitter Wine Taste Perception Forms a 2D Dataspace. J Comput Biophys Chem. https://doi.org/10.1142/S2737416524500844
33. Aliev T, Korolev I, Yasnov M et al (2025) Rosé or White, Glass or Plastic: Computer Vision and Machine Learning Study of Cavitation Bubbles in Sparkling Wines. RSC Adv. https://doi.org/10.1039/D5RA00046G
34. Alessio P, Constantino CJL, Daikuzono CM et al (2016) Academic Press,., Chap. 17 - Analysis of Coffees Using Electronic Tongues, in Electronic Noses and Tongues in Food Science, ed. By M.L. Rodríguez Méndez https://doi.org/10.1016/B978-0-12-800243-8.00017-2
35. Almario Á, Cáceres R (2014) Electronic Tongue and Neural Networks, Biologically Inspired Systems Applied to Classifying Coffee Samples. Am. J Anal Chem. https://doi.org/10.4236/ajac.2014.54033
36. Arrieta AA, Nuñez YE (2020) Mendoza. Mini-electronic Tongue Used to Discriminate between Coffee Samples of Different Geographical Origin. Int J Technol. https://doi.org/10.14716/ijtech.v11i2.3225
37. Brito-Pereira R, Policia R, Ribeiro C et al (2025) Hydrogel In-Tape Electronic Tongue. ACS Appl Electron Mater. https://pubs.acs.org/doi/ 10.1021/acsaelm.4c02059
38. Chang YT, Hsueh MC, Hung SP et al (2021) Prediction of Specialty Coffee Flavors Based on Near-Infrared Spectra Using Machine- and Deep-Learning Methods. J Sci Food Agric. https://doi.org/10.1002/jsfa.11116
39. Cho S, Moazzem MS (2022) Recent Applications of Potentiometric Electronic Tongue and Electronic Nose in Sensory Evaluation. Prev Nutr Food Sci. https://doi.org/10.3746/pnf.2022.27.4.354
40. Gabrieli G, Muszynski M, Thomas E et al (2022) Accelerated Estimation of Coffee Sensory Profiles Using an AI-assisted Electronic Tongue. IFSET. https://doi.org/10.1016/j.ifset.2022.103205
41. Tahara Y, Toko K (2013) Electronic Tongues–A Review. IEEE Sens J. https://doi.org/10.1109/JSEN.2013.2263125
42. Aliev T, Belyaev VE, Pomytkina AV et al (2023) Electrochemical Sensor to Detect Antibiotics in Milk Based on Machine Learning Algorithms. ACS Appl Mater Interfaces. https://pubs.acs.org/doi/ 10.1021/acsami.3c12050
43. Timofeenko IA, Khantimirov AG, Dranichnikova EA et al Machine Learning Model for Coffee Flavor Identification (RUS). RU 2023683061 software, 2023-11-02 Bull. 11. https://new.fips.ru/registers-doc-view/fips_servlet?DB=EVM&DocNumber=2023683061&TypeFile=html
44. Timofeenko IA, Khantimirov AG, Dranichnikova EA et al ML Model for Identifying Quality of Coffee (RUS). RU 2023682692 software, 2023-10-30 Bull. 11. https://new.fips.ru/registers-doc-view/fips_servlet?DB=RUPATAP&DocNumber=2023127045&TypeFile=html
45. Timofeenko IA, Khantimirov AG, Dranichnikova EA et al Method for Evaluating Taste of Bean Coffee and Ground Coffee. RU2023127045 patent, 2025-03-04, Bull. 7

46. Timofeenko IA, Khantimirov AG, Dranichnikova EA et al Database of Device Readings on the Taste Characteristics of Coffee (RUS). RU2023624498, 2023-12-08, Bull. 12
47. Pedregosa F, Varoquaux G, Gramfort A et al (2011) Scikit-learn: Machine Learning in Python. J Mach Learn Res 12(85):2825–2830
48. Dorogush AV, Ershov V, Gulin A (2018) CatBoost: Gradient Boosting with Categorical Features Support, arXiv preprint, arXiv:1810.11363 (cs.LG). https://arxiv.org/abs/1810.11363
49. Chen T, Guestrin C (2016) XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785
50. Neese F (2025) Software update: The ORCA Program System—Version 6.0. Wiley Interdisciplinary Reviews: Computational Molecular Science. https://doi.org/10.1002/wcms.70019
51. Caldeweyher E, Bannwarth C, Grimme S (2017) Extension of the D3 Dispersion Coefficient Model. J Chem Phys 147:3
52. Weigend F, Ahlrichs R (2005) Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. Phys Chem Chem Phys. https://doi.org/10.1039/B508541A
53. Trott O, Olson AJ (2010) AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. J Comput Chem. https://doi.org/10.1002/jcc.21334
54. Kim S, Chen J, Cheng T et al (2025) PubChem 2025 update. Nucleic Acids Res. https://doi.org/10.1093/nar/gkae1059
55. Bennett JA, Socash T, Khamis N et al (2024) Electrochemical Investigation of the Oxidation of Phenolic Compounds Using Hypervalent Iodine Oxidants. ACS Omega. https://doi.org/10.1021/acsomega.4c09236
56. Bumbaugh RE, Pennington DL, Wehn LC et al (2026) Direct Electrochemical Appraisal of Black Coffee Quality Using Cyclic Voltammetry. Nat Commun. https://doi.org/10.1038/s41467-026-71526-5
57. Honda T, Takemura K, Matsumae S et al (2024) Quantification of Caffeine in Coffee Cans Using Electrochemical Measurements, Machine Learning, and Boron-Doped Diamond Electrodes. PLoS ONE. https://doi.org/10.1371/journal.pone.0298331
58. Rustemova A, Korkmaz Ş, Bosnali W et al (2026) Electrochemical Sensing of Caffeine in Turkish Coffee via a Natural Flavonoid-Modified Pencil Graphite Electrode: A Rapid and Cost-effective DPAdSV Approach. Food Anal Methods 19:287. https://doi.org/10.1007/s12161-026-03199-1
59. Kushwaha A, Singh G, Gaur UK et al (2024) Highly Sensitive and Selective Electrochemical Detection of Caffeine, Theophylline and Guaiacol in Green Tea, Green Coffee, Coffee, and Tea Extracts Using a SnS2 Nanoflake-Modified Electrode. Mater Adv. https://doi.org/10.1039/D3MA00561E

60. Petrucci R, Zollo G, Curulli A et al (2018) New Insight into the Oxidative Mechanism of Caffeine and Related Methylxanthines in Aprotic Medium: May Caffeine Be Really Considered as an Antioxidant? Biochim. Biophys Acta - Gen Subj. https://doi.org/10.1016/j.bbagen.2018.05.011

Figures

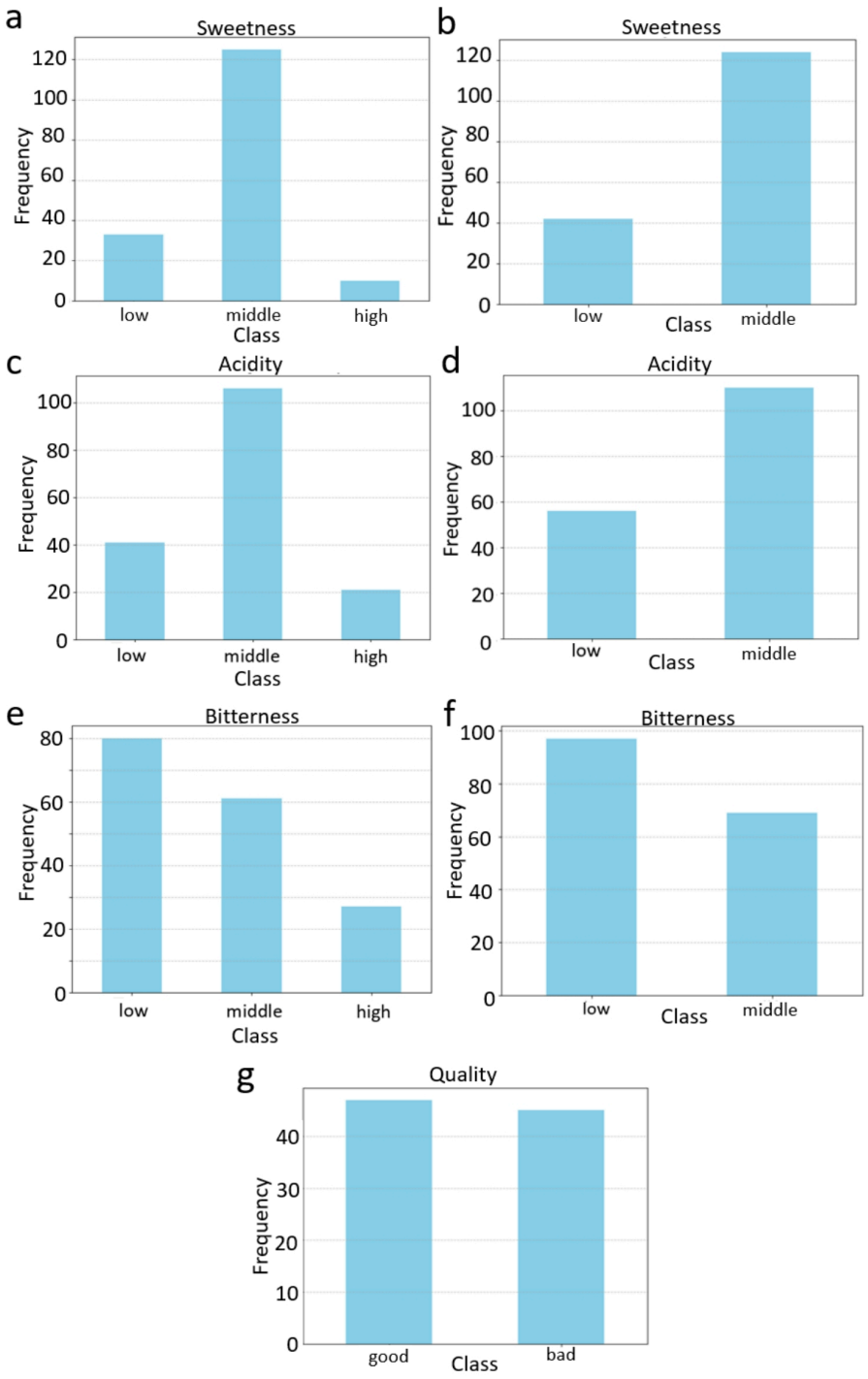


Figure 1

The distribution of target variables within three sensory categories: sweetness (a, b), acidity (c, d), bitterness (e, f); and quality category (g) before (a, c, e, g) and after rearrangement (b, d, f). The initial distribution shows the non-uniformity of the data sets of sensory categories, where high class is the least abundant. The target variables are received from the panelists' assessments of the intensity of the sensory categories (a-f) and quality perception of these sensory categories (g). Classes are redistributed for uniformity and feasibility for the ML-input and the highest accuracy of predictions

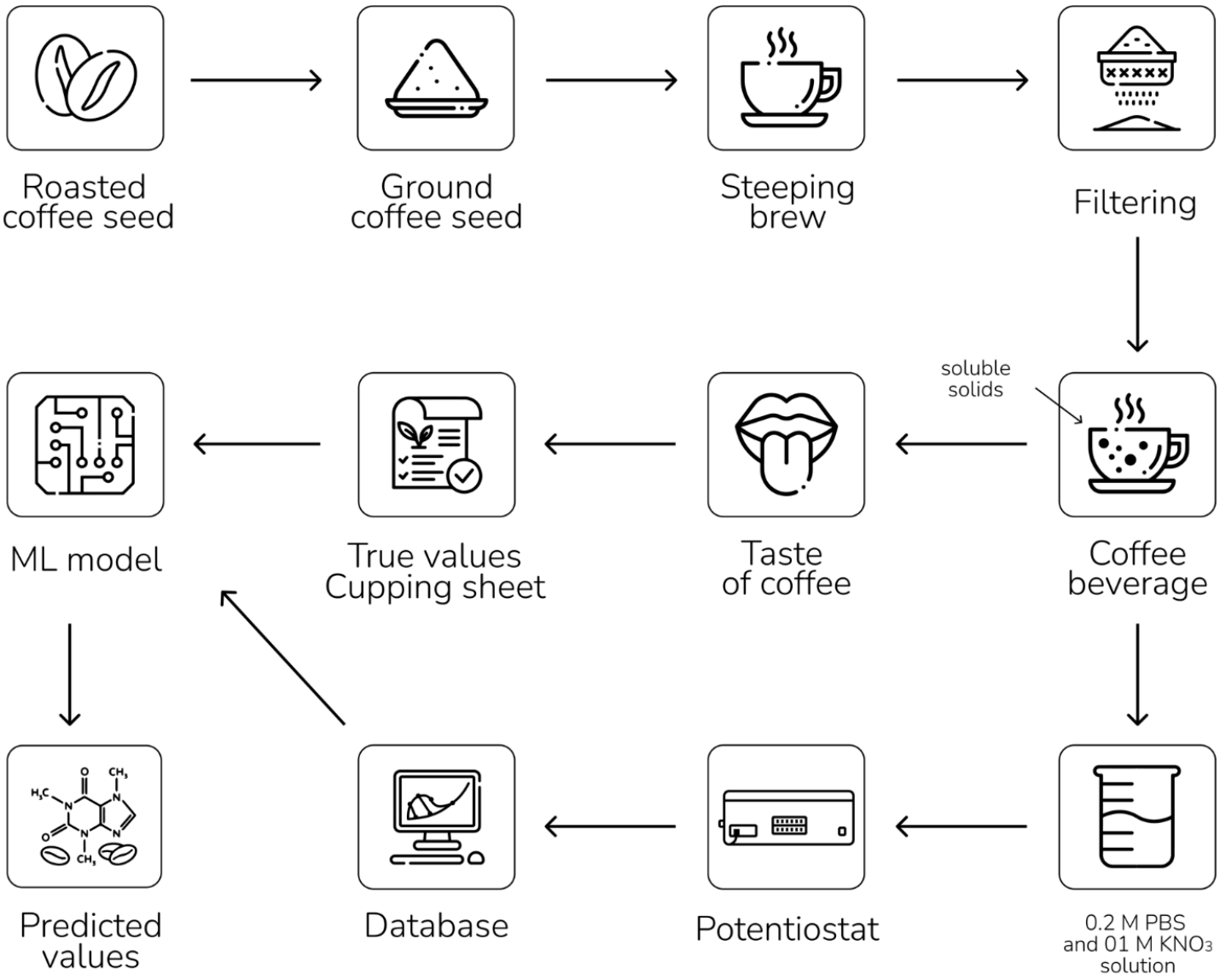


Figure 2

Workflow chart depicting the procedure of obtaining experimental data for prediction of coffee taste and quality. The input of ML-models comprised the panelists' assessments of coffee taste and quality as well as the voltammetric data derived from cyclic voltammograms. Predictions of taste and quality were the output of the models

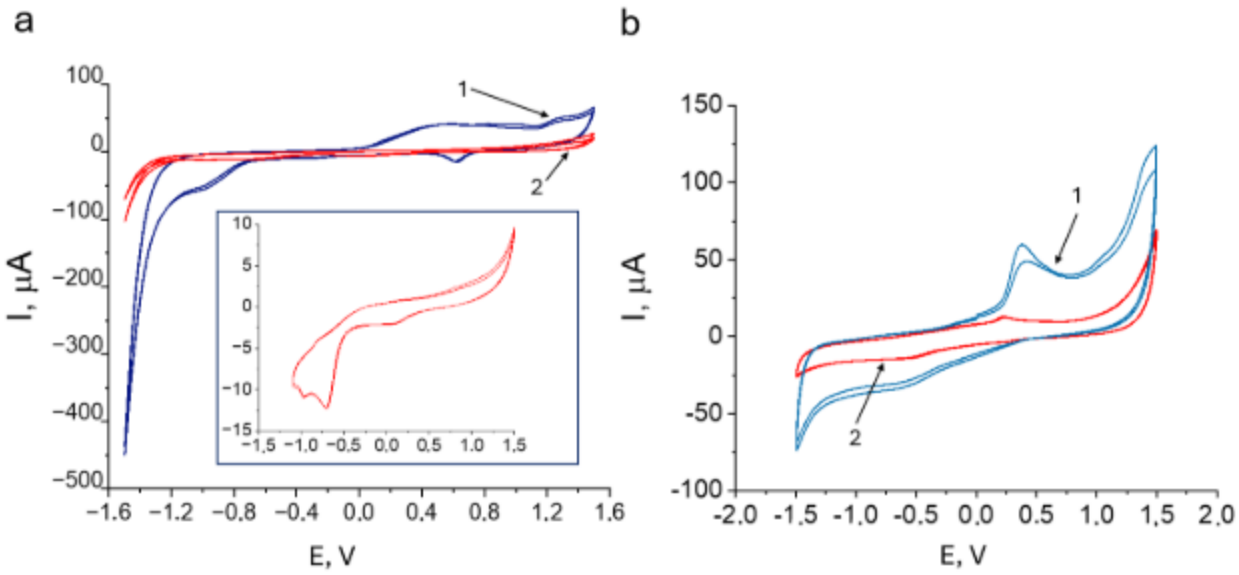


Figure 3

Cyclic voltammograms of a gold (a) and a glassy carbon (b) electrode in coffee samples (1) containing 0.2 M PBS and 0.1 M KNO_3, and blank electrolyte solutions (2) without coffee. The inset in (a) duplicates cyclic voltammograms of gold in the blank solution in a shortened range of potentials. Coordinate scale is the same for both plots in (a). The potentials are measured versus a silver chloride electrode as reference. Platinum wire is used as a counter electrode. The scanning rate is 50 mV s^{-1}

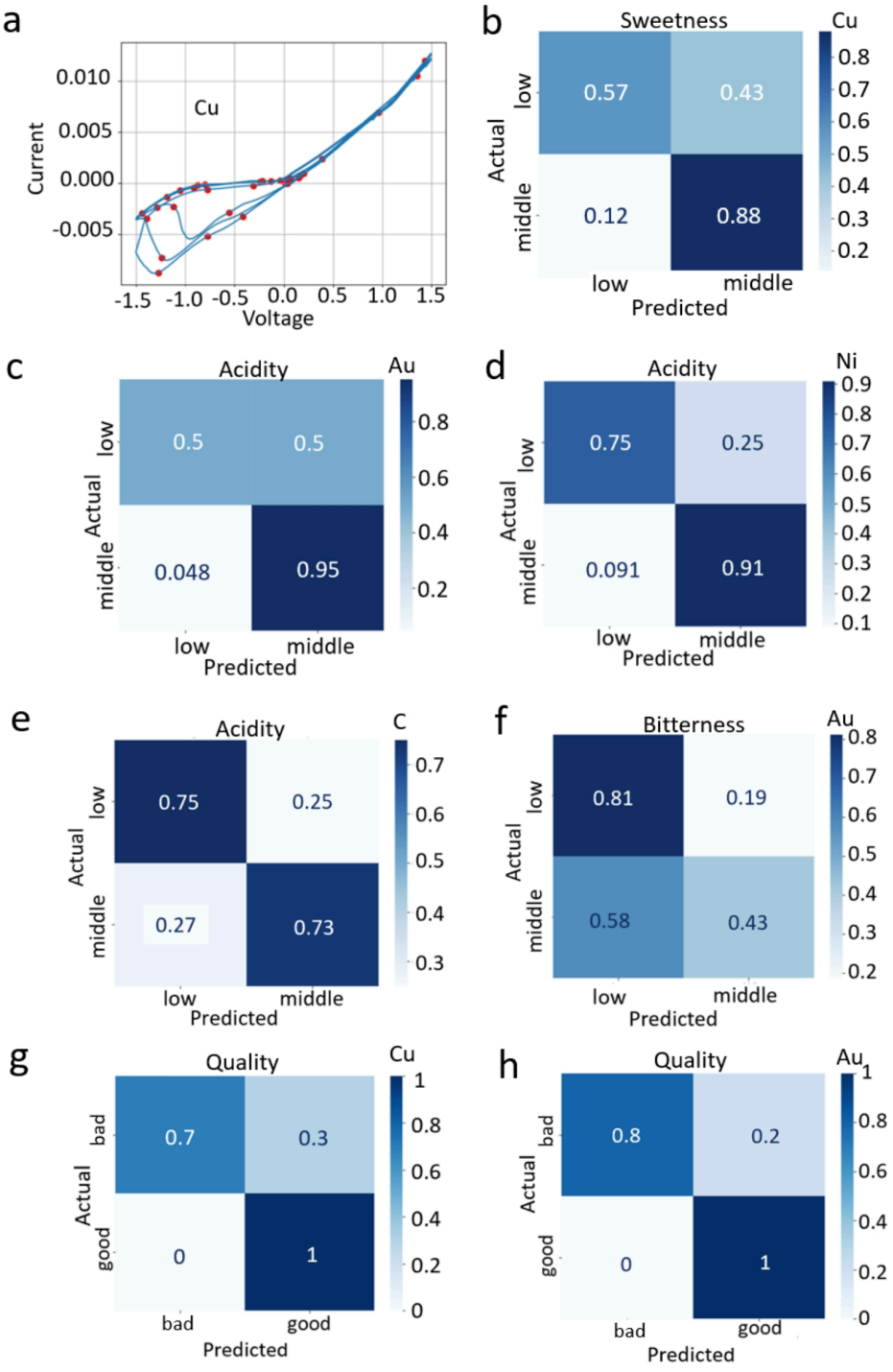


Figure 4

Machine learning results with the best performance: a) a series of typical voltammograms of a copper electrode in brewed coffee used for machine learning (dots denote the most important current values affecting final accuracy metrics of ML models used for taste classification (g); b-h) confusion matrixes for binary ML-models depicting prediction accuracy for the intensity of taste (b-f) and quality (g, h) by means of versatile ML algorithms:

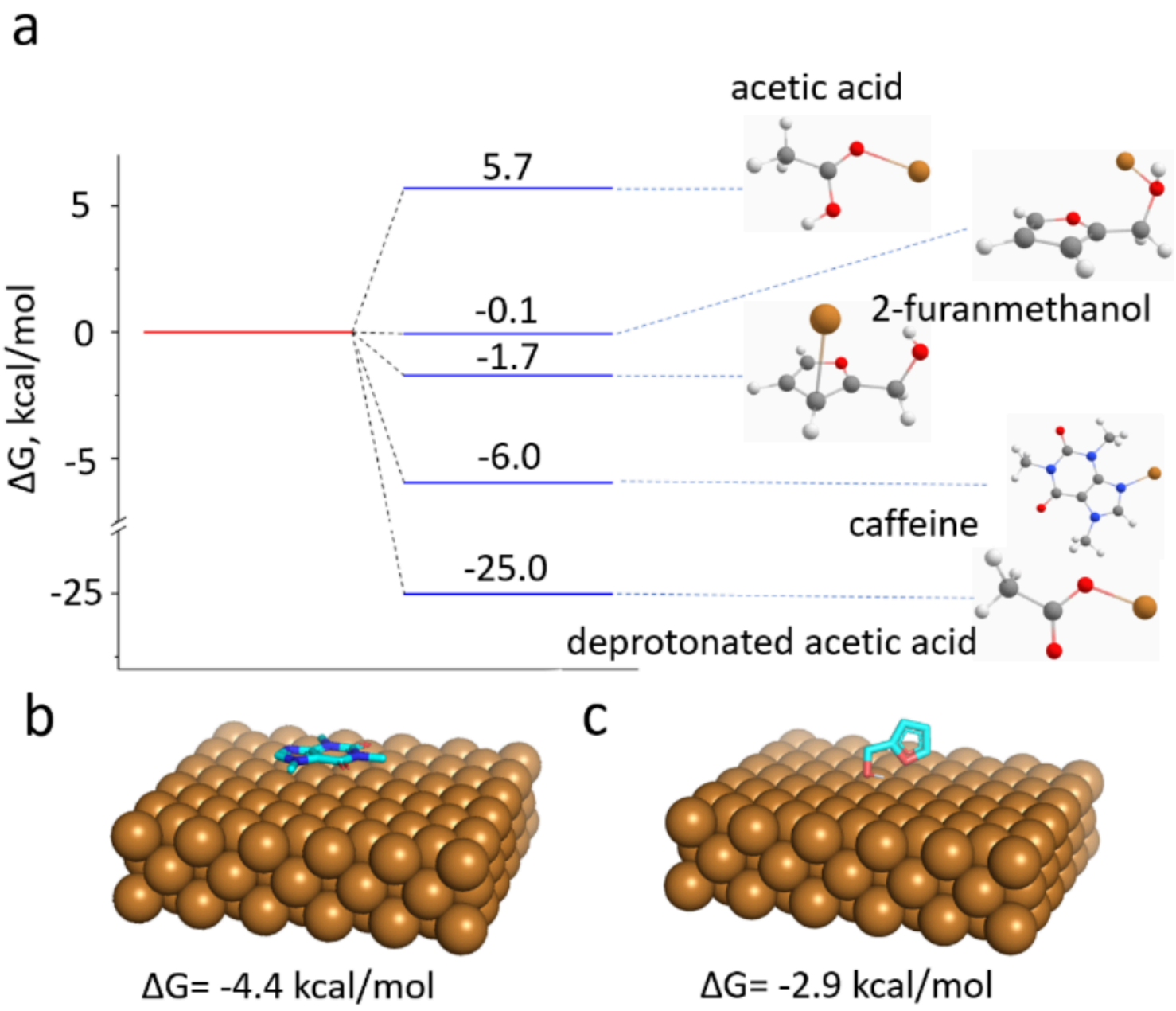


Figure 5

Computation models of the interaction between copper and selected individual substances of coffee and corresponding affinity energies obtained with DFT method (a) and molecular docking method for caffeine (b) and 2-furanmethanol (c)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SIcoffeeFoodAnalyticalMethods.docx
- furanmethanolOCu.xyz
- acetate.xyz

- aceticCu.xyz
- aceticNi.xyz
- caffeineCu.xyz
- caffeine.xyz
- aceticAu.xyz
- acetateCu.xyz
- caffeineAu.xyz
- furanmethanolAu.xyz
- furanmethanolONi.xyz
- caffeineNi.xyz
- acetateNi.xyz
- furanmethanol.xyz
- furanmethanolCu.xyz
- furanmethanolNi.xyz
- acetateAu.xyz
- acetic.xyz