Electronic Supplementary Information

Metal-Organic Frameworks at the Tip of the E-Tongue: Machine Learning-Driven Disposable Electrochemical Sensors

Ivan S. Zlobin[a], *Artem N. Sinelnikov*[a], *Nikita S. Toroptsev*[a], *Konstantin O. Titov*[a], *Evgenia Antoshkina*[a], *Igor A. Nikovskiy*[a], *Valentin V. Novikov*[bc], *Yulia V. Nelyubina*[a*]

[a]Nesmeyanov Institute of Organoelement Compounds, Russian Academy of Sciences, 119334, Vavilova Str., 28, bld. 1, Moscow, Russia

[b]Departament de Química Inorgànica i Orgànica, Universitat de Barcelona, 08007, Gran Via de les Corts Catalanes, 585, Barcelona, Spain

[c]Institute of Nanoscience and Nanotechnology, University of Barcelona, 08007, Gran Via de les Corts Catalanes, 585, Barcelona, Spain

[*] Corresponding author email: yulia.v.nelyubina@gmail.com

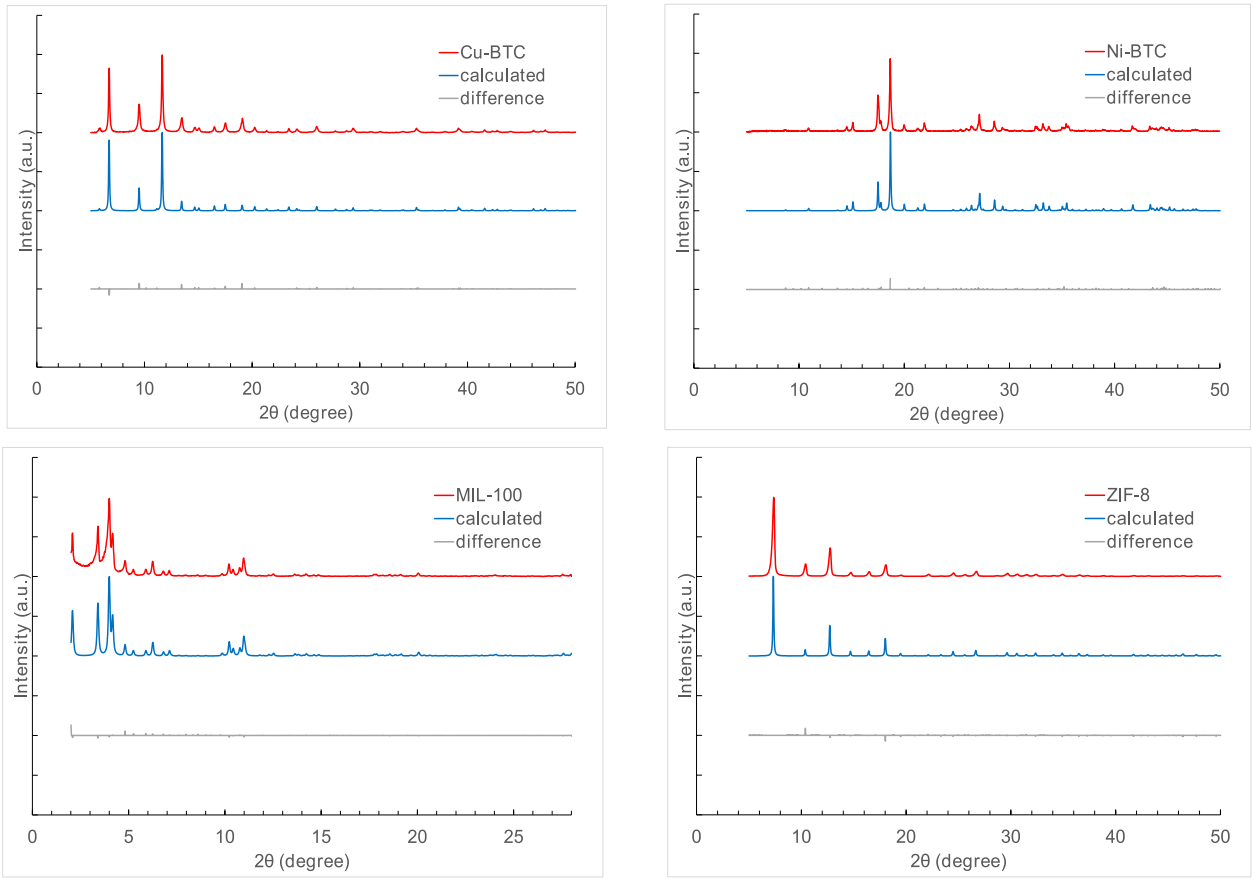


Figure S1. X-ray diffractograms collected (red line) and calculated (blue line) for MOFs (Cu-BTC, Ni-BTC, MIL-100(Fe) and ZIF-8) and their difference (black line).

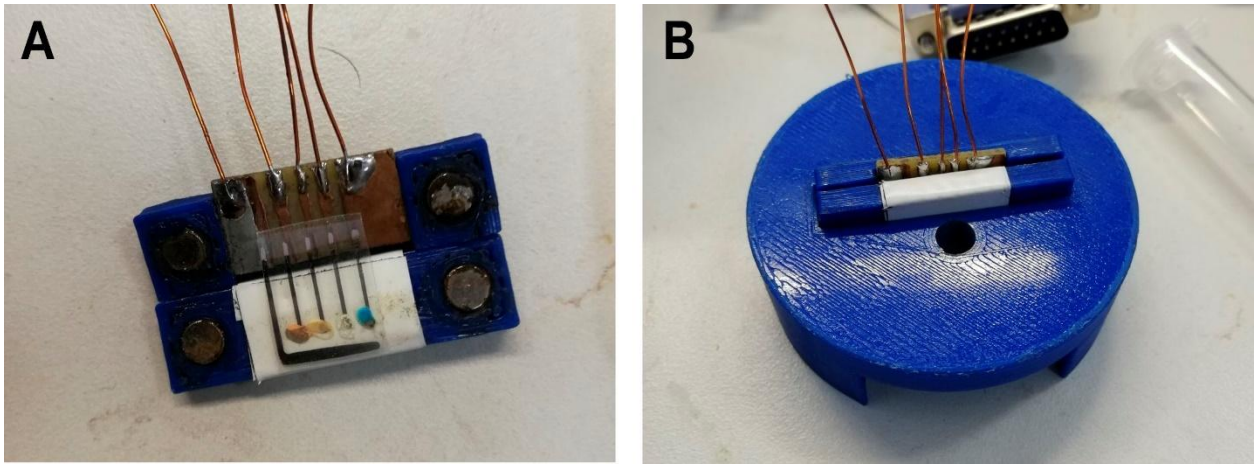


Figure S2. Adapter with a contact pad for the electrode connection (A) and a fabricated adapter holder (B).

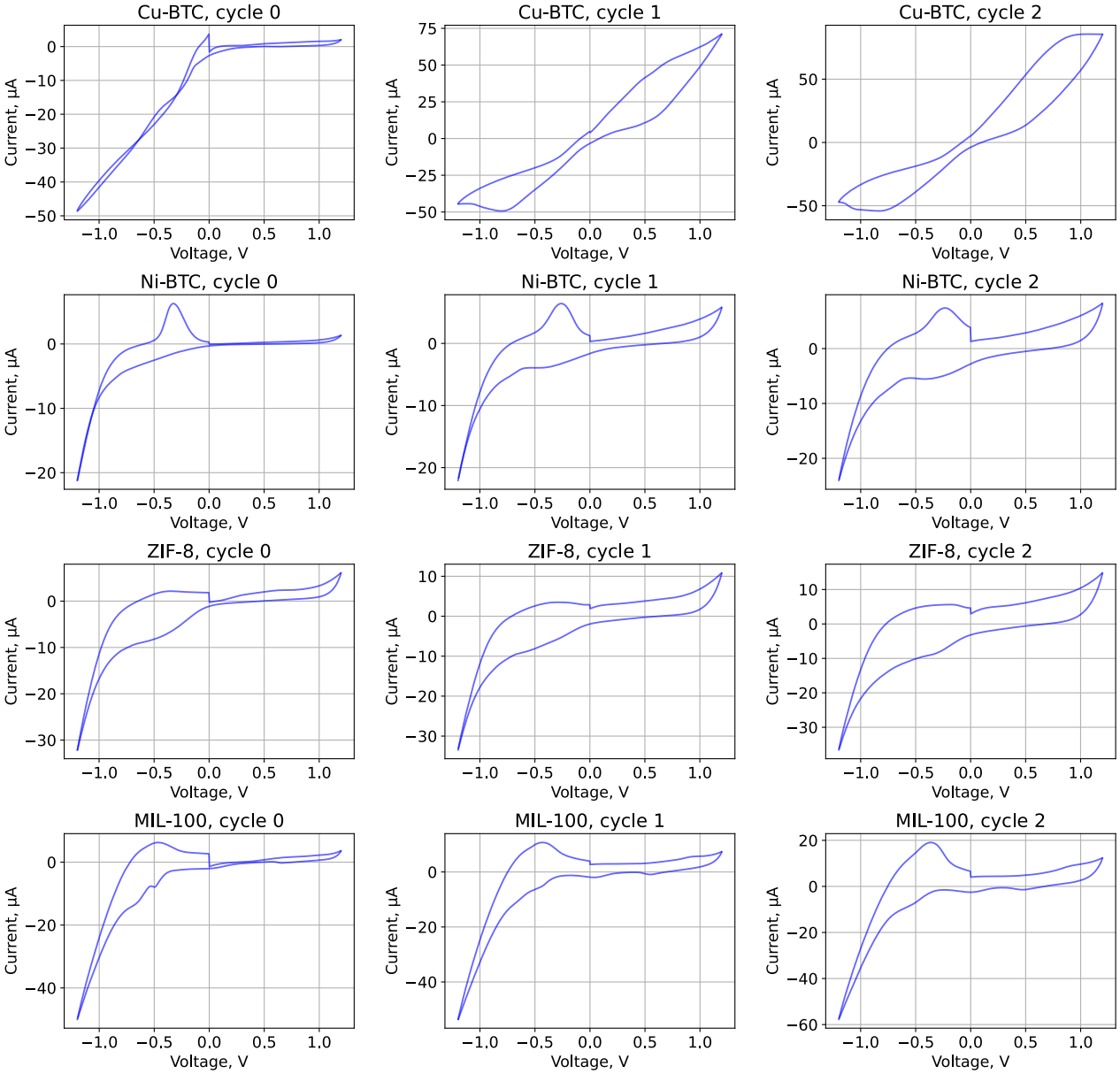


Figure S3. A characteristic fingerprint of an XZ tea sample, which includes three voltage scanning cycles for decreasing and increasing voltage for each MOF-modified electrode. The observed difference between Cycle 0 and Cycle 1 for Cu-BTC arises from its Cu^{2+}/Cu^{+} redox transformations, which may induce irreversible structural changes (such as ligand dissociation) during the initial electrochemical activation. ZIF-8, Ni-BTC and MIL-100(Fe) feature higher stability under the same conditions.

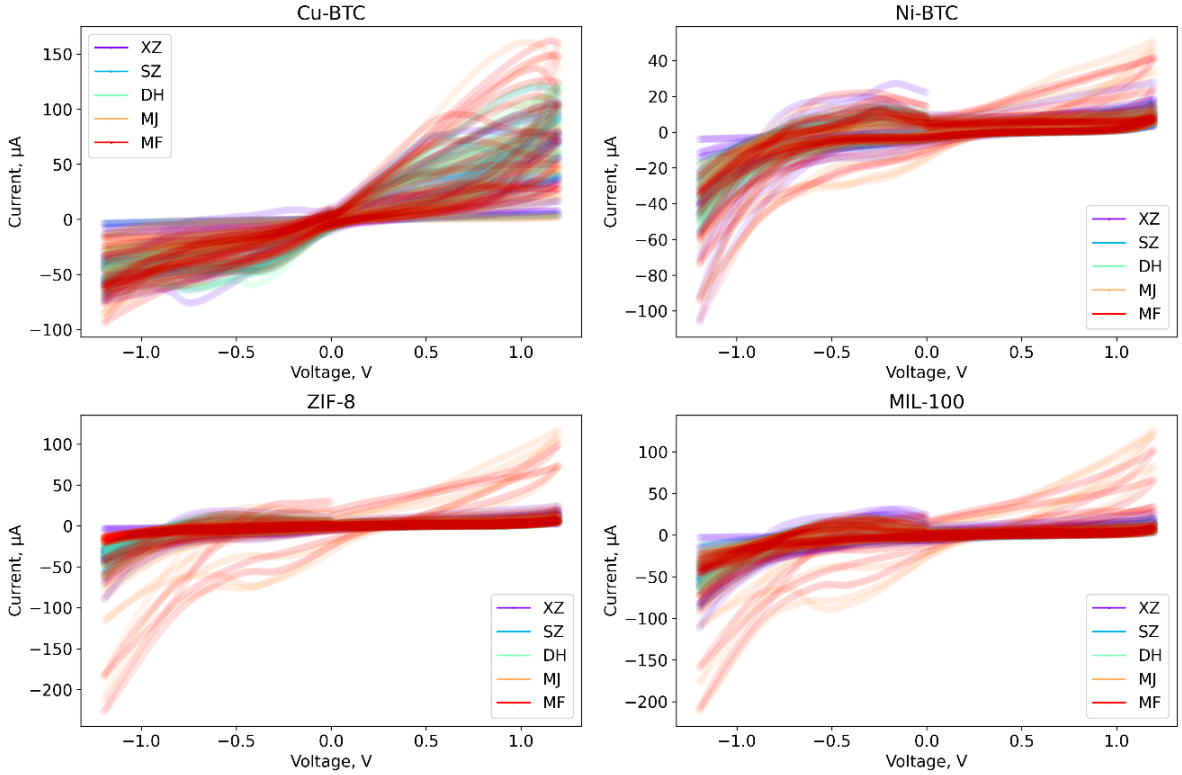


Figure S4. CV curves for the tea samples of five tea varieties as recorded by different working electrodes; only the second scanning cycle is shown for readability.

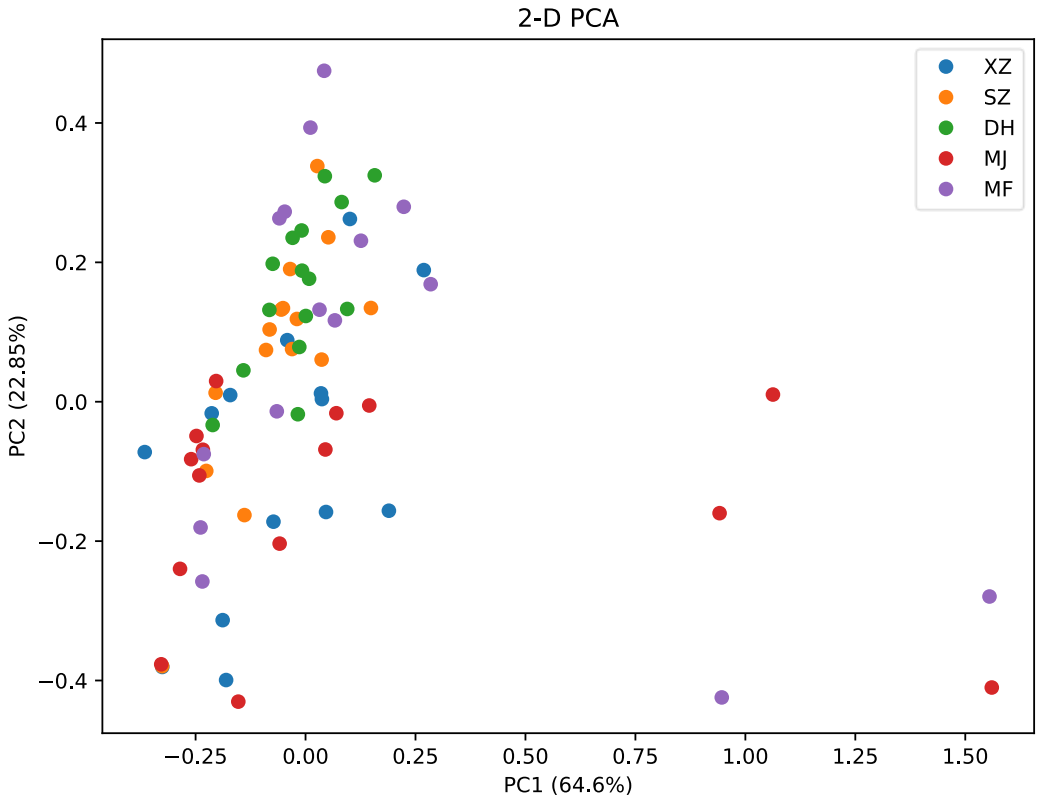


Figure S5. PCA score plots from the characteristic fingerprints of the samples collected by the fabricated e-tongues. In brackets, the explained variance for each principal component is given. Despite a rather high explained variance (87.32%), no visual separation of tea varieties is observed or any intrinsic pattern identified.

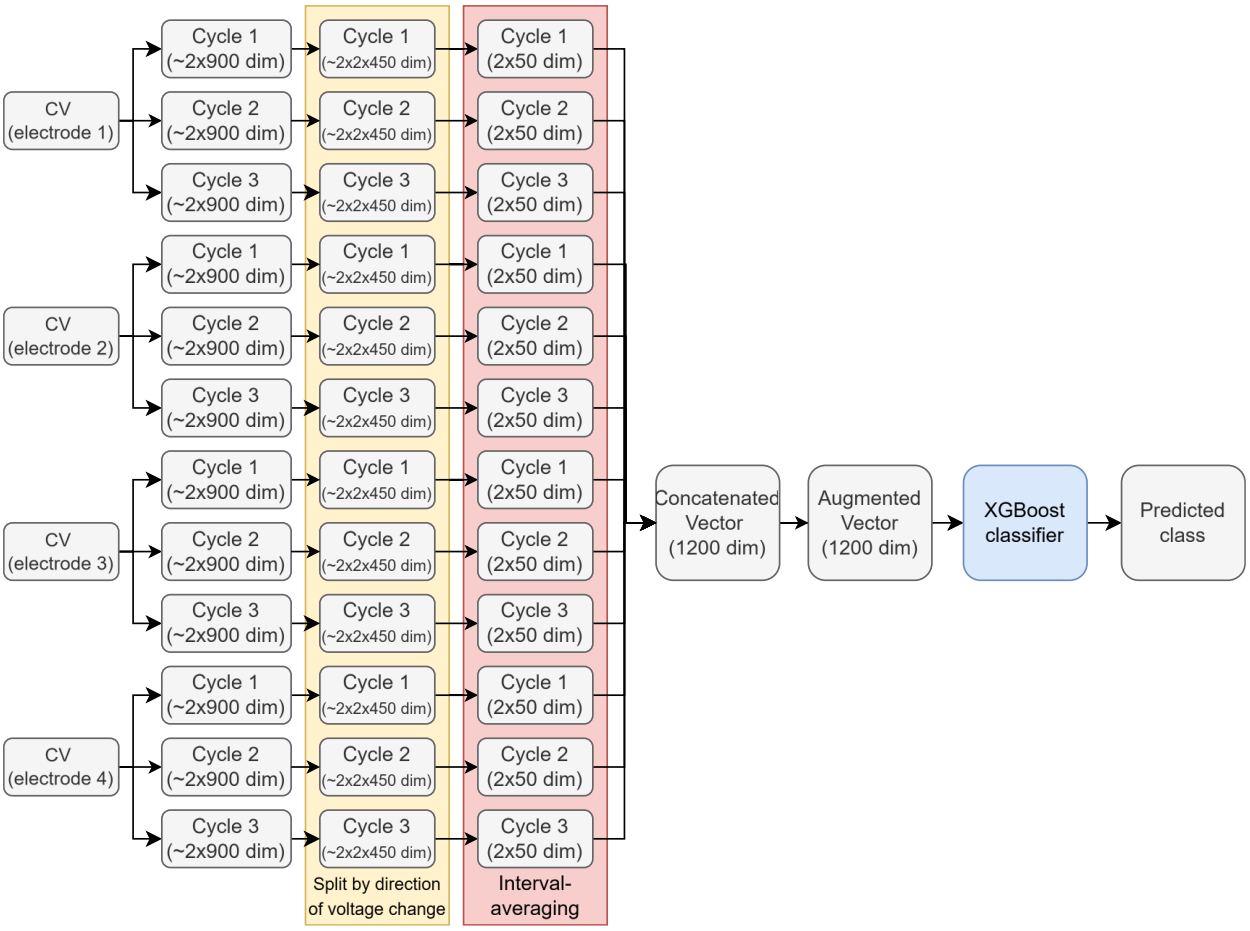


Figure S6. Schematic representation of the CV data preprocessing and classification pipeline using the XGBoost algorithm.

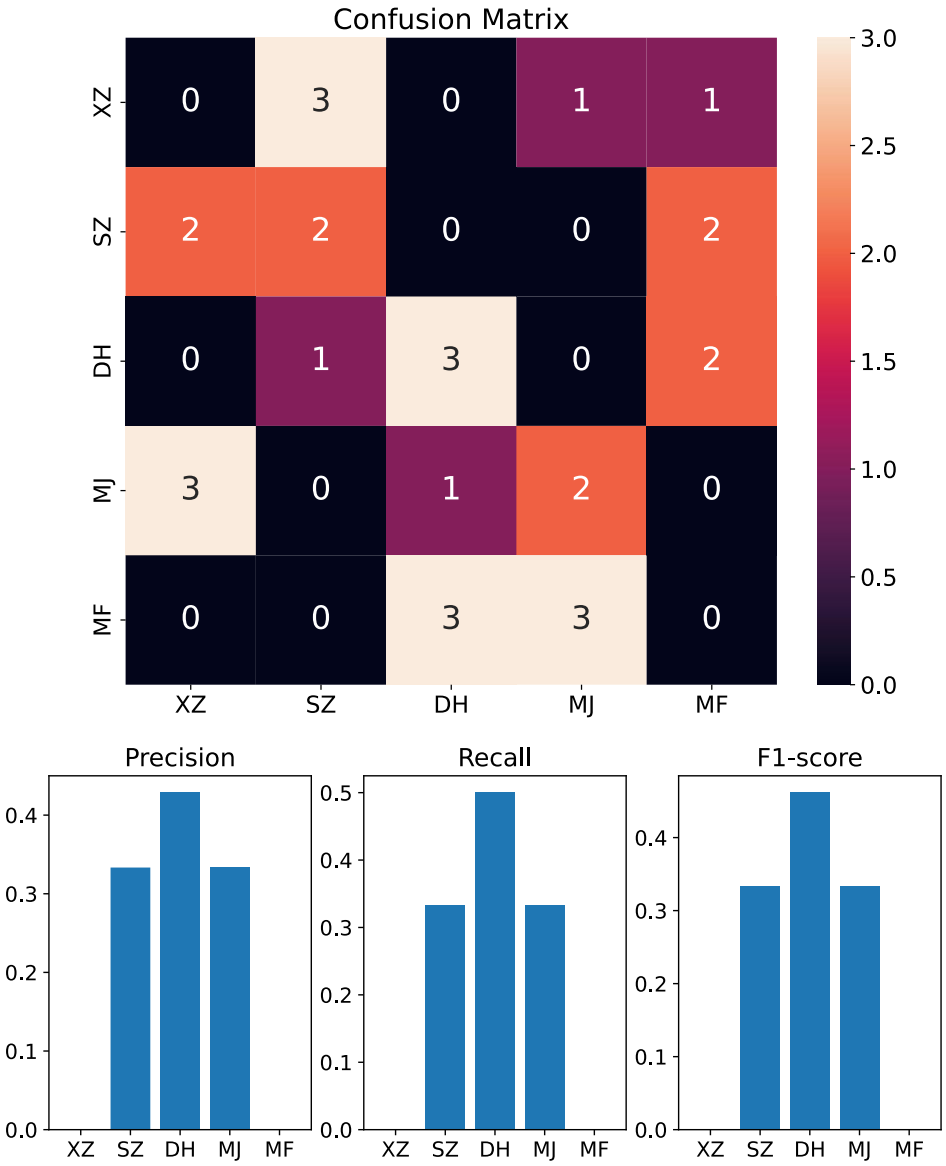


Figure S7. Classification results of the characteristic fingerprints for the five tea varieties collected by the e-tongues with three MOF-modified working electrodes (**Ni-BTC**, **MIL-100(Fe)** and **ZIF-8**) as obtained by the XGBoost algorithm.

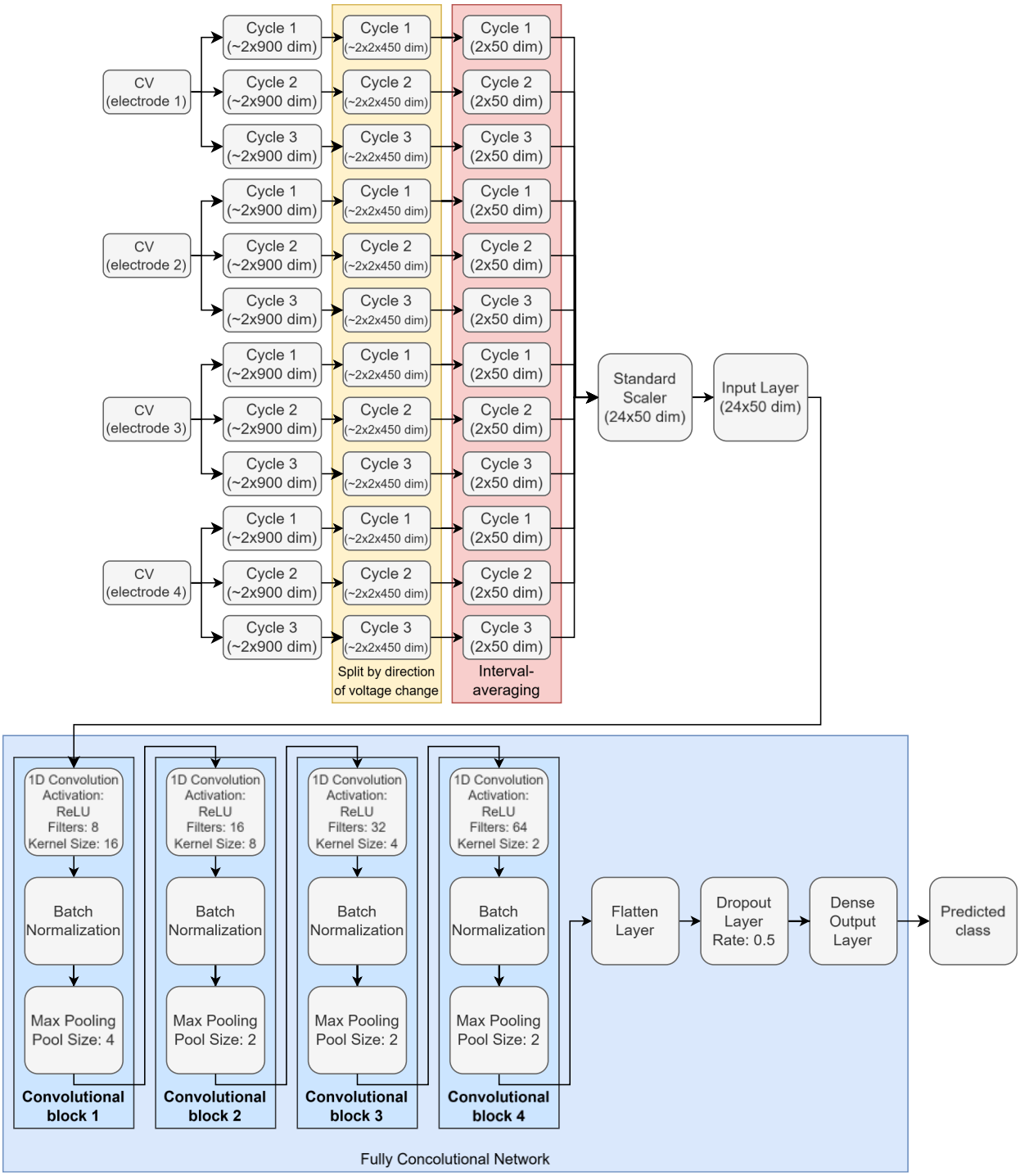
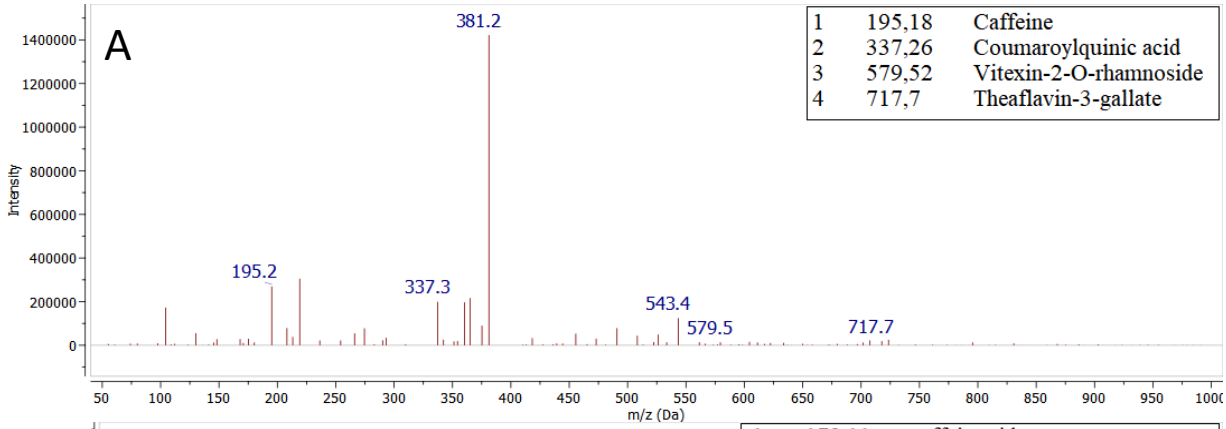
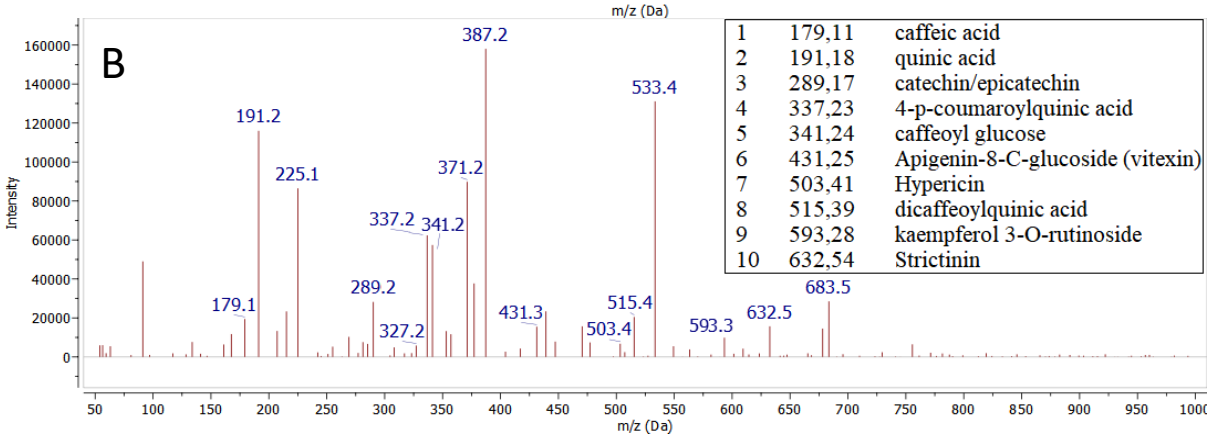
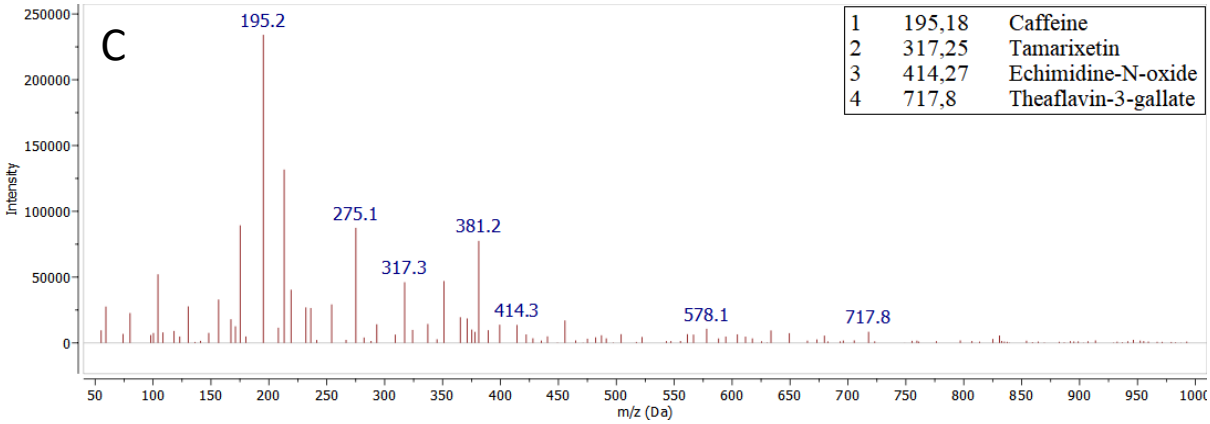
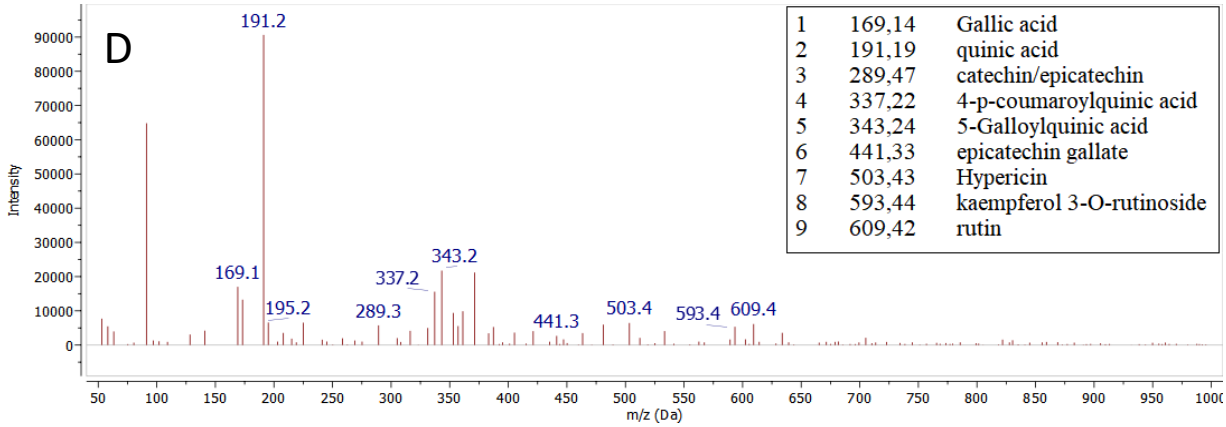


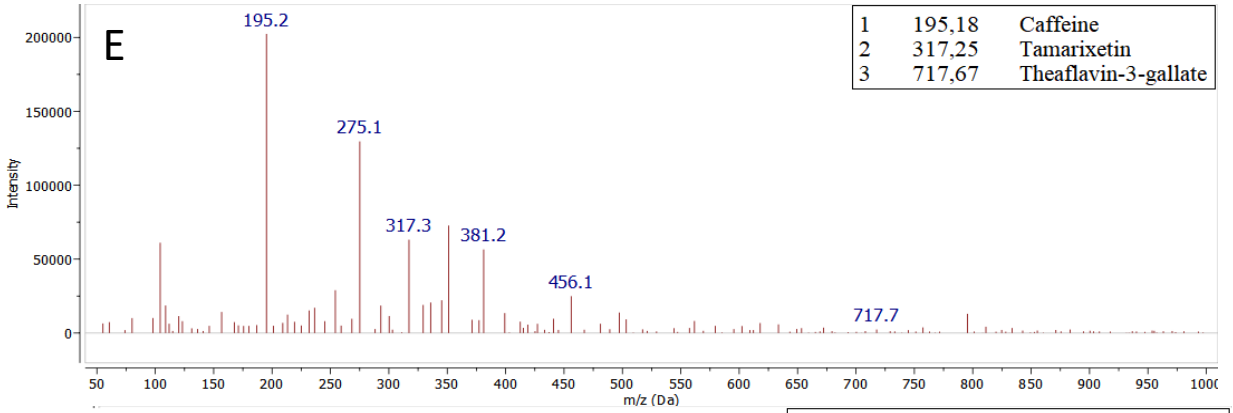
Figure S8. Schematic representation of the CV data preprocessing and classification pipeline using Fully Connected Network (FCN).

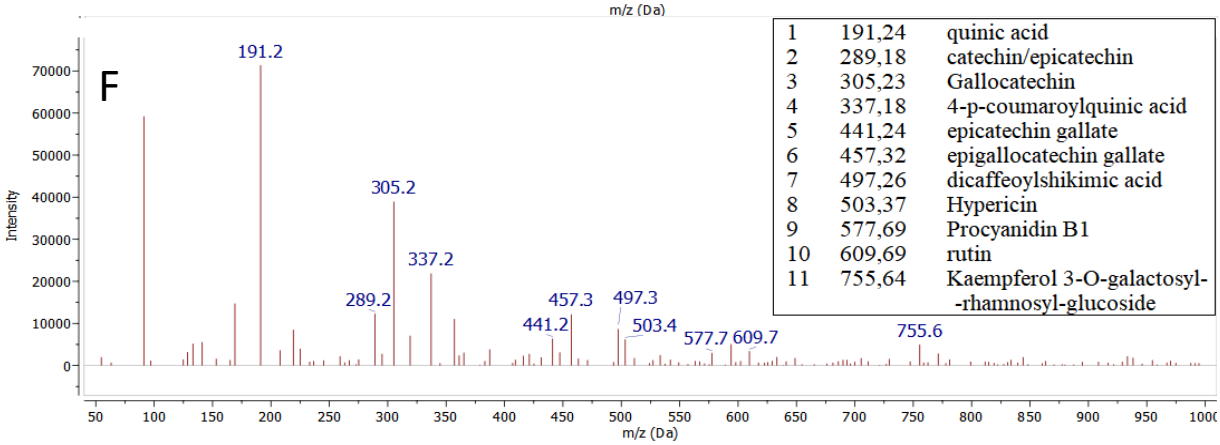

A
381.2
195.2
337.3
543.4
579.5
717.7
Intensity
m/z (Da)
1 195,18 Caffeine
2 337,26 Coumaroylquinic acid
3 579,52 Vitexin-2-O-rhamnoside
4 717,7 Theaflavin-3-gallate

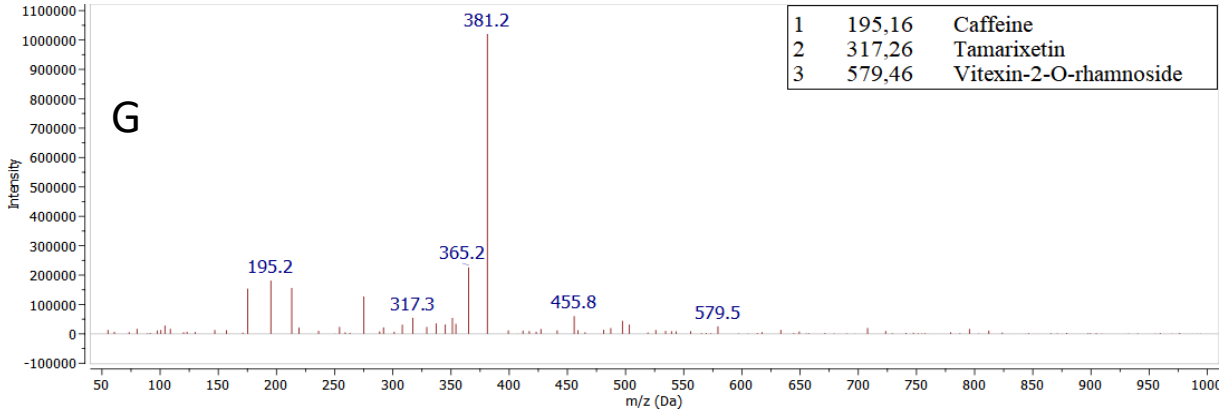

B
387.2
533.4
191.2
225.1
371.2
337.2
341.2
289.2
179.1
327.2
431.3
515.4
503.4
593.3
632.5
683.5
Intensity
m/z (Da)
1 179,11 caffeic acid
2 191,18 quinic acid
3 289,17 catechin/epicatechin
4 337,23 4-p-coumaroylquinic acid
5 341,24 caffeoyl glucose
6 431,25 Apigenin-8-C-glucoside (vitexin)
7 503,41 Hypericin
8 515,39 dicaffeoylquinic acid
9 593,28 kaempferol 3-O-rutinoside
10 632,54 Strictinin

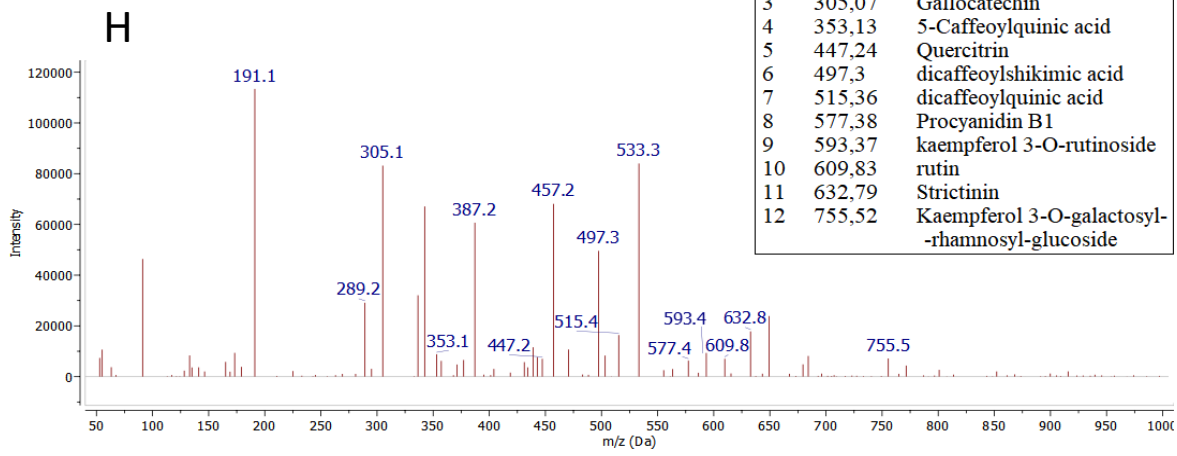

C
195.2
275.1
381.2
317.3
414.3
578.1
717.8
Intensity
m/z (Da)
1 195,18 Caffeine
2 317,25 Tamarixetin
3 414,27 Echimidine-N-oxide
4 717,8 Theaflavin-3-gallate


D
191.2
169.1
343.2
337.2
195.2
289.3
441.3
503.4
593.4
609.4
Intensity
m/z (Da)
1 169,14 Gallic acid
2 191,19 quinic acid
3 289,47 catechin/epicatechin
4 337,22 4-p-coumaroylquinic acid
5 343,24 5-Galloylquinic acid
6 441,33 epicatechin gallate
7 503,43 Hypericin
8 593,44 kaempferol 3-O-rutinoside
9 609,42 rutin


E
195.2
275.1
317.3
381.2
456.1
717.7
Intensity
m/z (Da)
1 195,18 Caffeine
2 317,25 Tamarixetin
3 717,67 Theaflavin-3-gallate


F
191.2
305.2
337.2
289.2
457.3
497.3
441.2
503.4
577.7
609.7
755.6
Intensity
m/z (Da)
1 191,24 quinic acid
2 289,18 catechin/epicatechin
3 305,23 Gallocatechin
4 337,18 4-p-coumaroylquinic acid
5 441,24 epicatechin gallate
6 457,32 epigallocatechin gallate
7 497,26 dicaffeoylshikimic acid
8 503,37 Hypericin
9 577,69 Procyanidin B1
10 609,69 rutin
11 755,64 Kaempferol 3-O-galactosyl--rhamnosyl-glucoside


G
381.2
195.2
365.2
317.3
455.8
579.5
Intensity
m/z (Da)
1 195,16 Caffeine
2 317,26 Tamarixetin
3 579,46 Vitexin-2-O-rhamnoside


H
191.1
305.1
533.3
457.2
387.2
497.3
289.2
515.4
593.4
632.8
353.1
447.2
577.4
609.8
755.5
Intensity
m/z (Da)
1 191,11 quinic acid
2 289,21 catechin/epicatechin
3 305,07 Gallocatechin
4 353,13 5-Caffeoylquinic acid
5 447,24 Quercitrin
6 497,3 dicaffeoylshikimic acid
7 515,36 dicaffeoylquinic acid
8 577,38 Procyanidin B1
9 593,37 kaempferol 3-O-rutinoside
10 609,83 rutin
11 632,79 Strictinin
12 755,52 Kaempferol 3-O-galactosyl--rhamnosyl-glucoside

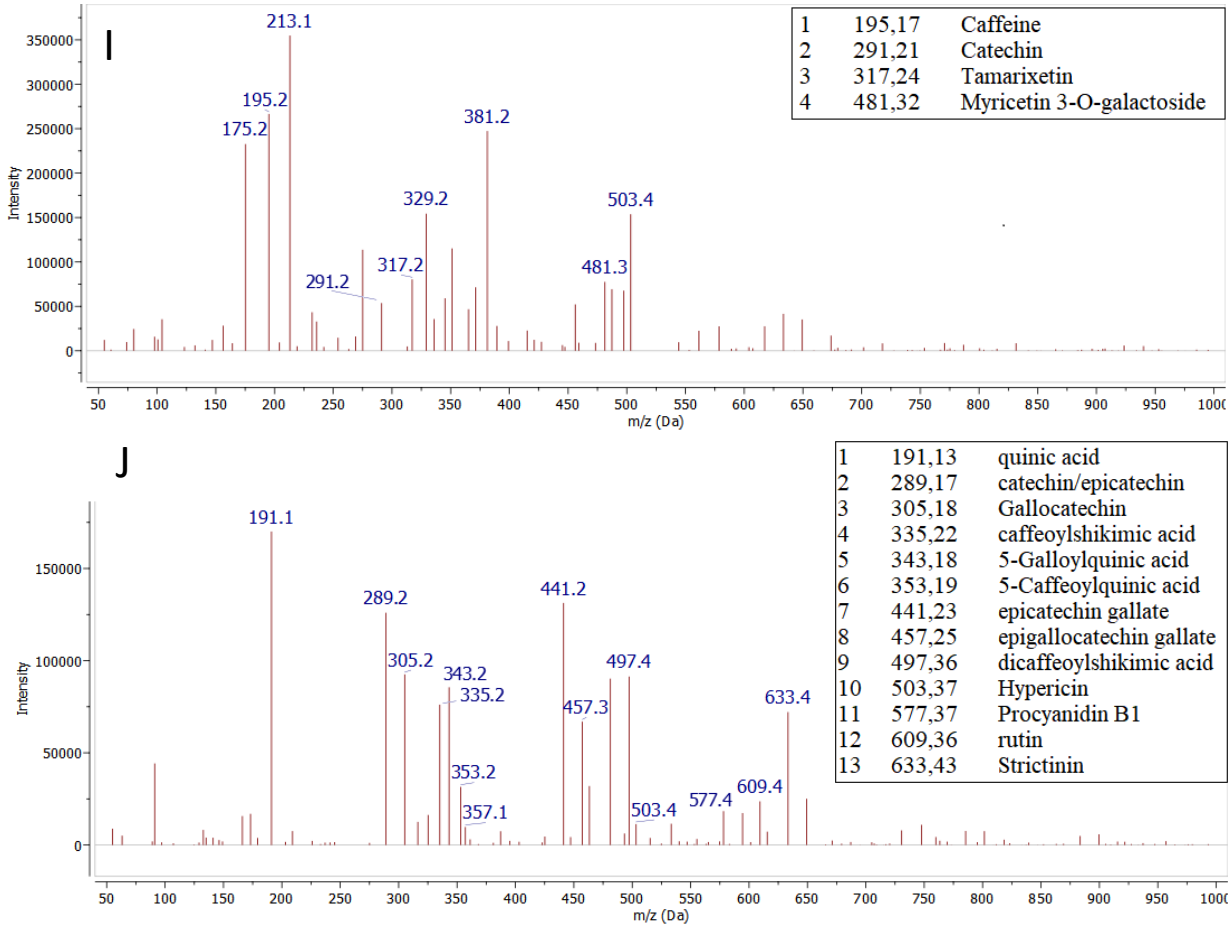


Figure S9. ESI-mass spectra of the tea samples XZ (A, B), SZ (C, D), DH (E, F), MJ (G, H) and MF (I, J).

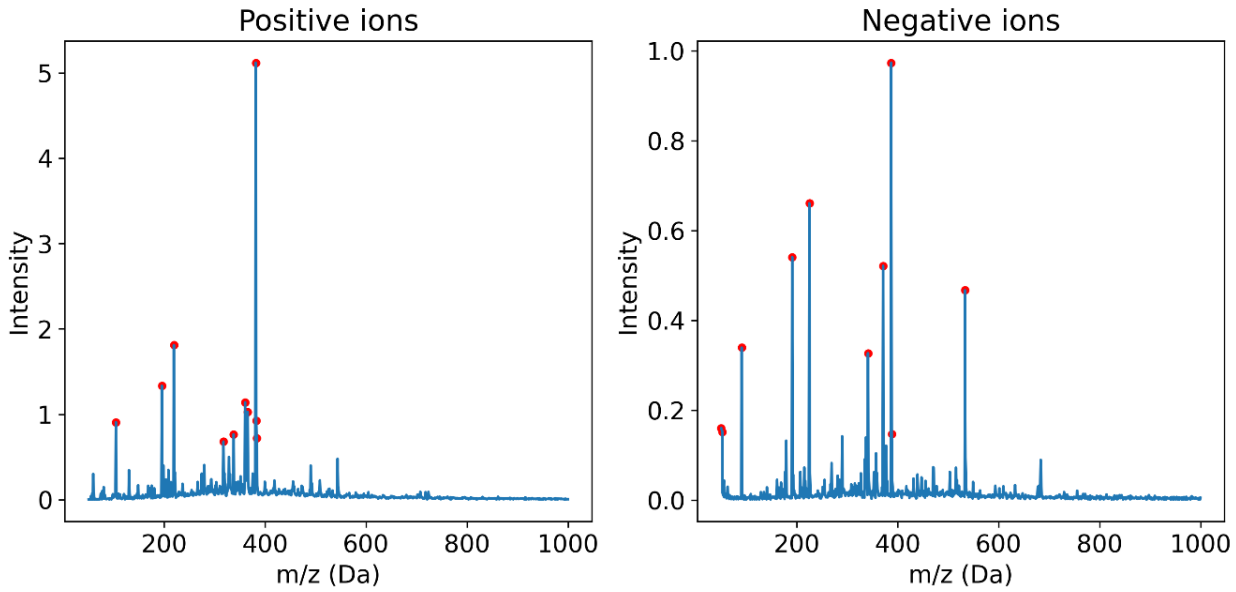


Figure S10. Representation of the XZ tea mass spectra at the input of the FCN model. The dimensionality of the mass spectra of the XZ tea sample (blue) is reduced to 10 most intensive peaks (red).

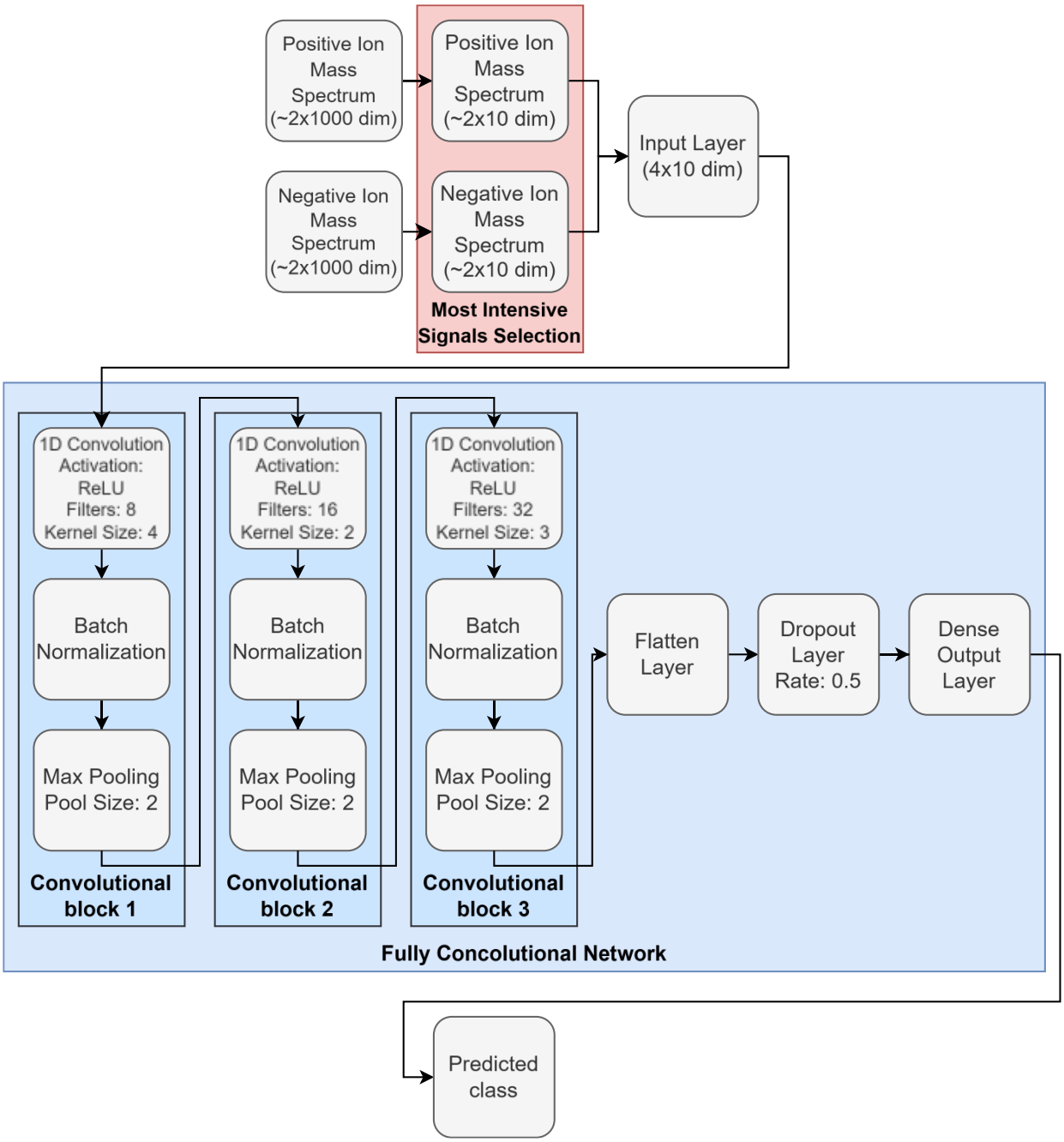


Figure S11. Schematic representation of the MS data preprocessing and classification pipeline using Fully Connected Network (FCN).

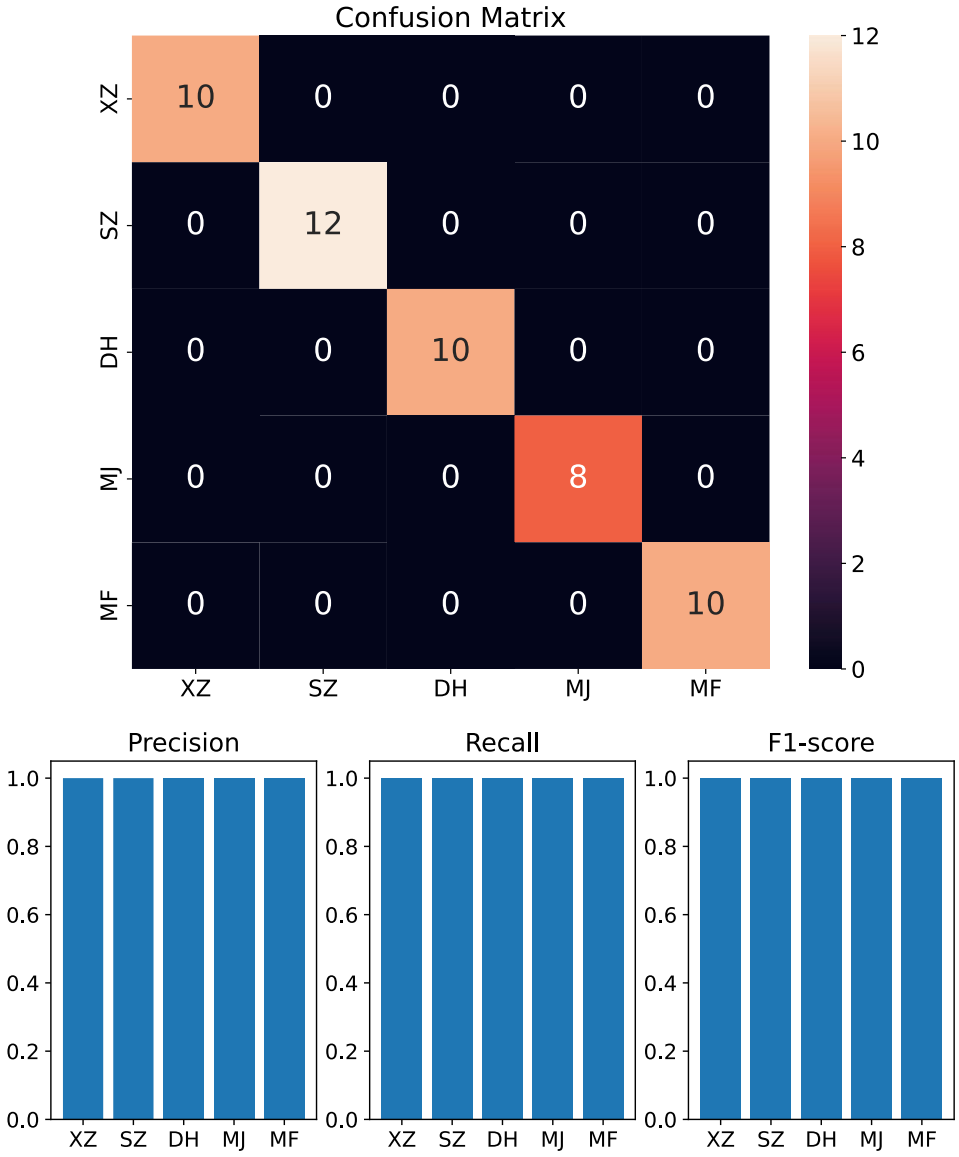


Figure S12. Classification results of the mass spectra for the five tea varieties as obtained by the FCN model.

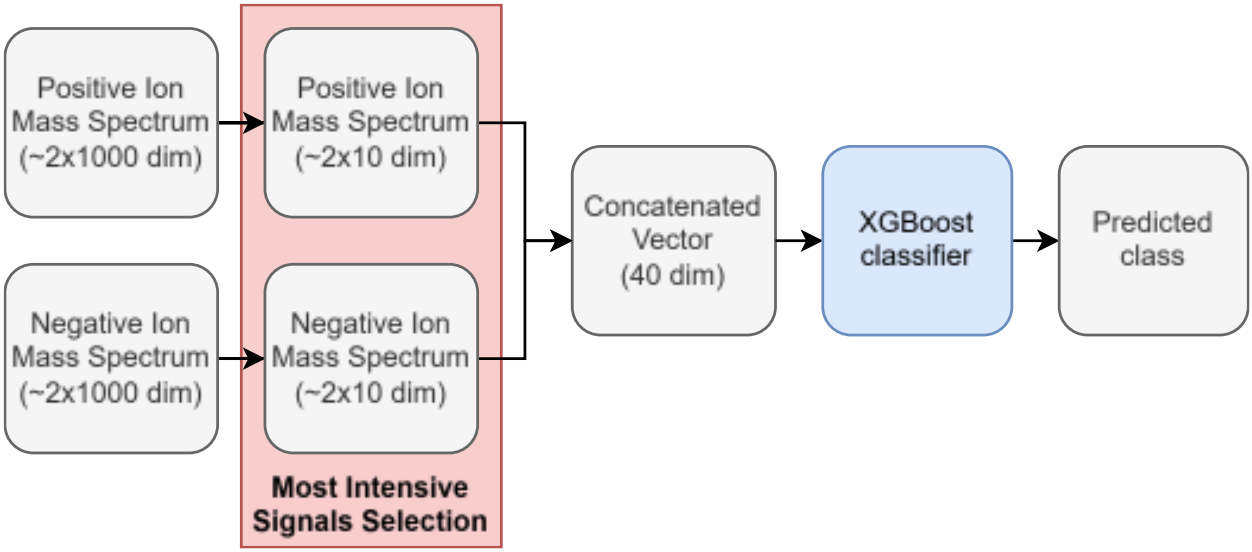


Figure S13. Schematic representation of the MS data preprocessing and classification pipeline using the XGBoost algorithm.